

Development Economics Papers Summary

Contents

1	The Role of the State	2
2	Governance and Accountability	7
3	Ethnicity and Development.....	13
4	Conflict and Development	18
5	Motivation of Public Servants	23
6	Taxation and Development.....	28
7	Firms, Markets and Entrepreneurship.....	32
8	Infrastructure and Trade	44
9	Energy and Development.....	52
10	Environment and Development.....	57
11	Child Labour.....	61
12	Growth Theory / Development and Growth Accounting	62
13	Misallocation.....	65
14	Management.....	79
15	Human Capital / Education.....	79
16	Health.....	92
17	Risk Sharing & Insurance.....	107
18	Psychology and Development.....	120
19	Transfers (UCTs, CCTs, IKTs)	124
20	Institutions and Development	126
21	Microfinance.....	132
22	Technology Adoption	141
23	Pricing and Willingness-to-Pay	147
24	Methods and External Validity	153

1 The Role of the State

1.1 Papers and General Introduction

(Main) papers discussed:

- i. Djankov, La Porta, Lopez-de-Silanes & Shleifer 2002. *The Regulation of Entry*. QJE
 - Assess whether heavier regulation of entry is associated with desirable economic and governance outcomes, in order to distinguish between various theories of regulation
- ii. Besley and Burgess 2004. *Can Labor Regulation Hinder Economic Performance? Evidence from India*. QJE
 - Assess the effect of state-level labour regulations (pro-worker vs pro-employer) on manufacturing sector performance in India
- iii. Aghion et al. 2008. *The Unequal Effects of Liberalisation: Evidence from Dismantling the License Raj in India*. AER
 - Study the differential effect of the dismantlement of the License Raj (= central control on manufacturing market), on manufacturing performance, focusing on interactions with labour regulations

General theme:

- What is the **scope of government intervention** in regulating the private sector?
 - State regulations of markets justified because of
 - **Market failures** which cause a failure of the First Welfare Theorem. Theory of **second best**.
 - **Efficiency / equity trade-off**. Redistribution to satisfy welfare function or, in first best world, to reach a different PO allocation (by lump-sum transfers)
 - Reasons why state interventions may fail to be second-best efficient:
 - Djankov: **Public choice** theories of gov't: not benevolent. Entry regulations as a **rent-seeking** strategy from bureaucrats in low accountability countries
 - BB 2004: **Unintended consequences** of "pro-worker" labour regulations – reduced employment; increased urban poverty and size of informal manufacturing sector
 - Issues with **enforcement**, presence of **informal sector** can make regulations counter-productive because firms can exit (this is a theme: relate to taxation, ban on child labour)
- Focus on **regulation of entry** and **labour market regulation**
 - Entry generally beneficial to growth: reallocation through competition & selection effects, increase variety and innovation. Yet not unambiguously positive. Infant industry arguments¹
 - Labour regulations: cause misallocation by affecting firms differentially; increase unemployment and size of informal sector

¹ "relatively backward economies may switch out of the investment-based strategy too soon, so certain policies such as limits on product market competition or investment subsidies, which encourage the investment-based strategy, may be beneficial" Acemoglu, Aghion & Zilibotti 2006.

- Aghion et al. 2008: **Interactions between entry and labour regulations.** In “pro-worker” states, the removal one distortion (licence raj) made the state worse off compared to “pro-employer” states – second best considerations.

1.2 Main Ideas and Results

1.2.1 Djankov et al 2002

- Two main questions:
 - What are the consequences of regulating entry?²
 - Who collects rents from limiting entry?
- Efforts to **construct comparable measure of entry regulation**
 - Data on official regulations: number of procedures; (minimum) official cost & time required to complete process
 - Focus on “representative” firm: a domestically firm operates in largest city, performs general commercial activity, no foreign trade, some minimum capital, between 5-50 employees etc.
- **Theories of regulation:**
 - Pigou’s **public interest** theory of regulation: gov’t policy benevolently corrects market failures. Entry regulation serves to screen out low -quality products, fly-by-night-operators, produces with high externalities. Prediction: stricter regulation should be associated with better outcomes
 - **Public choice theories – regulation for rent creation** and extraction.
 - Stigler 1971: **Regulatory capture** by the industry incumbents who want to limit entry and increase mark-ups
 - **Tollbooth view:** Regulation pursued for benefit of politicians – collect bribes in exchange of permit of entry
 - **Theoretical prediction ambiguous:** benevolent social planner could finance screening through taxes rather than by imposing high cost of entry, and corrupt. Higher fees reflect *capture theory* (what if this is only “revenue-efficient” way for gov’t to raise revenue?)
- **Results**
 - Enormous variation across countries
 - Number of procedures highly correlated with time and cost of process
 - Rich countries regulate less
 - Market failures more prevalent in poor – greater need to regulate
 - Controlling for GDP p.c., regulation **negatively influences quality, positively influences** the size of the **unofficial economy** and **negatively (but insignificantly)** affects **product market competition**, and **positively influences corruption**
 - In support of the **public choice theory**, in particular **tollbooth view**

² I think the article conflates “the socially efficient level or form of regulation” with “heavy/strict regulation”. Especially 2nd paragraph p.4. A proper test of theories of regulations would require knowing: (a) the socially efficient level of regulation (unobserved); (b) the actual level of regulation put in place by bureaucrats (measured). And then comparing the “distance” between the two, and seeing whether it is significantly affected by measures of accountability, controlling for GDP/capita, etc. As of now, cannot say all regulation is rent-extortion. There are just few countries in high accountability countries that *require* extensive entry regulation.

1.2.2 Besley and Burgess 2004

- What is the **effect of labour market regulations on manufacturing performance** and on poverty, between 1958 and 1992?
- **Background & data:** Low growth of Indian manufacturing (13% in 1960 to 18% in 1995) masks large cross-state and time differences.
 - ASI data on output, fixed capital, labour, productivity, wages. NSS data on urban poverty
 - Industrial Disputes Act (applies only to registered manufacturing): code each of the 113 amendments as pro-worker (1), pro-employer (-1) or neutral (0)
- **Theoretical framework:** Stage 0: firms choose to enter registered (any size, regulations) or unregistered manufacturing sector (no regulations, but limited in size), stage 1: choose capital, stage 2: choose number of workers, and bargain over rents
 - Labour regulation can affect economic performance through 2 channels:
 - **Relative price effect:** raise cost of employment & **adjustment costs**: substitute away from labour to labour-saving inputs, reduce output because MC has increased. Make firms more likely to enter unregistered sector (also causes labour to be misallocated across firms since hiring and firing is costly).
 - **Expropriation effect:** increase in **bargaining power** of organised labour means that there is a **risk** that the workers will expropriate return from capital once investments are made³. This discourages investment. Same predictions as relative price channel, with an added emphasis on lower capital stocks (this will also reduce employment a great deal if labour and capital are complements)
- **Results: Regulation in pro-worker direction lowered registered manufacturing performance**
 - Negative effect on registered manufacturing output + positive on unregistered = overall decrease in manufacturing output, and implied magnitude considerable.
 - Lowered employment (both on the extensive and intensive margin) investment and productivity in formal manufacturing. Also **increased urban poverty**.
 - **Welfare implications:** measure % of individuals below poverty line. Expect the direct effect on poverty to depend on the extent to which the earnings of the poor are derived from registered manufacturing. Do not have this, but note that presence of urban poverty is correlated with registered manufacturing (registered manufacturing located mainly in urban areas while unregistered) is both in rural and urban areas. Prior: Pro-worker regulation is positively correlated with urban poverty because of lower output and employment in registered manufacturing.

1.2.3 Aghion et al. 2008

- Assess how effect of dismantling the Licence Raj on manufacturing performance varies across states depending on labour regulations. **Interaction between labour and product market regulations.**

³ Similar to insecure property rights for owners of capital

- **Background & data:** Licence Raj is main instrument of centrally planned industrial policy in India. Two waves: 1985, one third on industries delicensed; then 1991 delicensing imposed by IMF.
 - State-industry panel data on registered manufacturing for 3-digit industries from 1980-1997, for 16 states (ASI)
 - Important controls: From 1991, other “Washington consensus” reforms imposes: reduction in tariff barriers & liberalisation of FDI – must isolate effect of delicensing by constructing measures for tariffs and FDI and including them as controls.
 - Exploit **staggering of delicensing over a decade**. Also when an industry was delicensed, all firms in that industry were delicensed. For labour market regulation, same coding as BB 2004
- Mechanisms: licensing as a barrier to entry: should see increase in entry in delicensed versus licensed industries.
- **Result: delicensing had a stronger effect on manufacturing performance in states with pro-employer regulations** (these industries grew more quickly)
 - This regards both the **number** and the real **output of factories**

1.3 Empirical Strategy and Discussion

1.3.1 Djankov et al 2002

- Look at basic regressions for the main results
- Strengths:
 - Big picture question
 - Very consistent evidence looking at different effects.
- Concerns:
 - Concerns about omitted variable bias. Unobserved heterogeneity between countries correlated with entry regulation, e.g. governance measure which is not perfectly picked up by GDP p.c.
 - Would have to use Within-country panel data with entry regulation varying across time, and space/industry

1.3.2 Besley and Burgess 2004

- Basic regression

$$y_{st} = \alpha_s + \beta_t + \mu r_{st-1} + \xi x_{st} + \varepsilon_{st}$$

where y_{st} is logged outcome variable, r_{st} is the regulatory measure: it is lagged to capture gap between enactment and implementation of the policy. x_{st} are other exogenous variables, α_s state fixed effect (culture, geography) β_t a year fixed effect (common shocks, central government policies). Clustered standard errors to deal with serial correlation.

- Strengths: Robustness checks
 - Include many political controls to identify effect of labour regulation as opposed to other gov't policies or general climate of labour relations, union power etc.

- Robust to exclusion of West Bengal, which passed the most pro-worker amendments and experienced a large drop in manufacturing performance (but they don't try excluding Andhra Pradesh, which is the outlier on the other end..?)
- Labour regulations not correlated with total state output: not proxy for other gov't policies
- Concerns:
 - Endogeneity: for example, regression is not industry specific. However, it is well conceivable that some regions have (for geographic or other reasons) a focus on certain industries which by nature demand a certain regulation and also influence outcomes.
 - Results not robust to including state-specific time-trends (?)

1.3.3 Aghion et al. 2008

- Basic regression: state s, industry i, year t

$$y_{ist} = \alpha_{is} + \eta_{it} + \beta_{st} + \theta(d_{it})(r_{st}) + \epsilon_{ist}$$

y_{ist} is log of 3-digit state-industry real output, d_{it} dummy for delicensing: =1 from the year in which the industry is delicensed, r_{st} is measure of labour regulation as in BB 2004, α_{is} , η_{it} and β_{st} to control for industry-state, industry-time & state-year interaction unobserved effects.. coefficient of interest θ is identified by the mix of industry-year variation in delicensing interacted with state-year variation in labour regulation.

Clustered standard errors by state and year of delicensing, not only for serial correlation & heteroscedasticity, but especially **to account that for the fact that delicensing occurred nearly all in 1985 and 1991**, so that different industries within a state in these years cannot be treated as independent

- Strengths:
 - Interaction effect looked at. Even if level is wrong, this may still be correct.
 - Robustness check with all sorts of control variables.
- Concerns
 - restriction to balanced panel and industries that exist in at least 5 states each year
 - OVB: shocks to unregistered manufacturing sector
 - Do not observe flows or entry and exit, only net entry (evolution of number of firms in an industry)

1.4 Policy Conclusions

- Regulation might not necessarily always happen in the **social interest**, but in the **private interest** of bureaucrats/politicians looking for rents
 - Some cross country evidence by Djankov et al. suggests this is the case
- Especially in developing countries, some evidence by Aghion et al. and Besley & Burgess seems to suggest that **pro-labour regulation does not promote industry growth** and thus **welfare** growth
- Regulation is not only per se a concern to look at, but particularly in **conjunction with liberalisation**. While liberalisation may generally have positive or negative effect, the

debate should not neglect that the direction and extent of the effect may considerably depend on the regulatory and institutional setting in which liberalisation takes place

- Specifically, Aghion et al. suggest that liberalisation in India led to much more growth etc. in a pro-employer policy setting.

2 Governance and Accountability

2.1 Papers and General Introduction

(Main) papers discussed:

- i. Besley & Burgess 2002. *The Political Economy of Government Responsiveness: Theory and Evidence from India*, QJE.
 - Assess the role of voter's information (newspaper circulation), electoral turnout and political competition about the politician's actions on government responsiveness
- ii. Finan & Ferraz 2008. *Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcome*, QJE.
 - Use of randomised municipal audits in Brazil in 2003 to assess the difference between pre-election and post-election audits on election outcomes.
 - Interaction with # of radio stations
- iii. Duflo & Chattopadhyay 2004. *Women as Policy Makers: Evidence from a Randomized Policy Experiment in India*, Econometrica.
 - Asks whether the identity of political representatives matters
 - Studies the impact of women's leadership on the type of public goods provided

General Theme:

- Importance of **information**, **political competition** and **identity** for accountability
 - FF 2008: Influence of information and its publication (media coverage) on accountability in the sense of **type distinction** between politicians and ability to punish/reward certain types with better politicians as a consequence
 - BB 2002: Influence of media presence on accountability in the sense of **government responsiveness**
 - CD 2004: Influence of the politician's identity on accountability in the sense of **representation** of different social groups

2.2 Main Ideas and Results

2.2.1 Besley & Burgess 2002

- Using data from 1958 to 1992 (34 years!) for 16 Indian states to analyse **links between information (newspaper circulation), electoral turnout, political competition and government responsiveness** (measured by public food distribution and calamity relief expenditures in response to low food grain production and flood damage)
- Focus on how **vulnerable minority** can profit from **media coverage** and political factors.
- Model to motivate empirics: two time periods, utility from office Ω , vulnerable population fraction $\gamma < 0.5$, a fraction β of which experiences a shock ($\beta\gamma$ are the needy), effort

e can be invested by politician to help needy population, and m is mass media activity. $q(e,m)$ is the fraction of non-needy vulnerable who are informed about the incumbent's effort. It increases in both e and m . $p(e,m) > q(e,m)$ is the fraction of the needy vulnerable informed. σ is electoral turnout among the vulnerable and $v \sim U[a, 2b - a]$ (around expected support b with noise $|b-a|$) is the fraction of non-vulnerable voters voting for the incumbent for ideological reasons. The incumbent wins if:

$$\gamma\sigma[p(e,m)\beta + (1-\beta)q(e,m)] + (1-\gamma)v > \frac{1}{2}$$

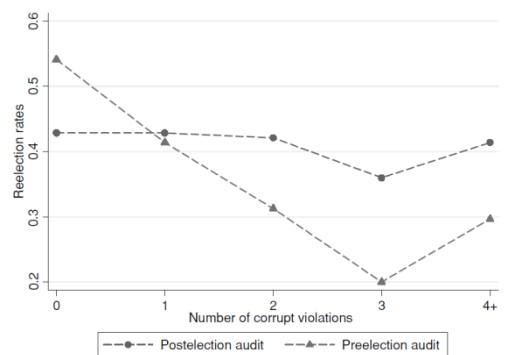
- From this, probability to win can be computed and subsequently $\max_e P(\cdot) \times \Omega - e$ can be solved for an interior solution to yield FOC:

$$\frac{\gamma\sigma}{2(b-a)(1-\gamma)} [p_e(e^*, m)\beta + (1-\beta)q_e(e^*, m)]\Omega = 1$$

- From this, it follows that the incumbent's effort increases in the parameters m , σ and γ and decreases in the expected support b (assumption: p_{em} and q_{em} are positive)
- In empirics, **newspaper** has a sign. positive effect on government responsiveness per se and in interaction with the crises (low food grain production and flood damage for public food distribution and calamity relief expenditure respectively).
 - Making a distinction between languages of newspapers, the most important effect and only significant one comes from **local languages**. Makes sense as *state* governments are analysed and for many vulnerable people, English or Hindi is not the lingua franca.
- Controlling for newspaper circulation, **political competition** and **electoral turnout** also have sign. positive effects on responsiveness, especially in interaction with the crises.
 - Political competition effect stronger for public food distribution than for calamity relief. That's probably because the former is much **more visible** and serves a broader range of the population, it can be highly politicised in election times.
 - Linking back to theory, **media activity increases the marginal value of effort**
- **Shared vulnerability** of a minority and **mass media** increase government responsiveness, i.e. have an influence on policy choices
- Responsiveness is only **weakly related** to factors of **economic development**.

2.2.2 Finan & Ferraz 2008

- 2004 municipal elections analysed with regards to how information on corruption activities before the election affect the re-election probability relative to information release after the election
- Information release probably **depends on priors**: if more corruption than expected, re-election probability should fall; if less than expected it should rise
 - They find a clear **downwards trend of the re-election rate** in the number of corrupt violations (in numbers: incumbent's likelihood of re-election decreased by 17% for at least two corruptions and pre-election audits)



- Note that the re-election rate is *higher* for pre-election audits in the case of zero corruption (positive surprise effect)
- In a second step, the **access to information** about the politician's corruption activities in pre- and post-election audits is taken into account. Interaction variables with the number of **radio stations** show that the clear downwards trend only holds for the case of municipalities with pre-election audits and the presence of a local radio station.
 - The **triple interaction pre-election audit x corrupt violations x radio** is highly significantly negative
 - For low corruption it is above the post-election average: The effect is therefore *relative* to prior beliefs. Good types are rewarded, bad types punished.
 - Controlling for the interaction or pre-election audit x corrupt violations x radio, the interaction pre-election audit x radio is actually significantly positive, i.e. controlling for corruption effects, **information increases re-election rates** for pre-election audits.
 - Also, controlling for the interaction pre-election audit x corrupt violations x radio, the interaction pre-election audit x corruption is not significant at all, i.e. **only if information about the corruption activities from pre-election audits is available does it have an effect on re-election rates**.

2.2.3 Duflo & Chattopadhyay 2004

- They make use of a natural experiment that 30% of village councils (Gram Panchayats, GP), randomly selected, have been required to have women as chiefs (Pradhan) since an amendment to the constitution in 1993 in India. They look at two districts, West Bengal and Rajasthan
- They answer **three questions**:
 - **Has women representation a policy impact**, i.e. does gender matter? Answer found: yes. **Theoretically**, it is **ambiguous** though. In a mean voter world, the politician should reflect the electorate's preference irrespective of their own identity. If politicians cannot commit to certain policies promised before elections (e.g. citizen-candidate model), identity representation plays a role.
 - If it does, are **policies** implemented under woman leadership more **in line with women's preferences**? Answer found: yes.
 - If so, does this come from the fact that women in power themselves have preferences more in line with the women they represent or the fact that they **respond more to issues raised by women** in the community? Answer: from the former effect only.
- Based on **citizen-candidate model** (Besley & Coate): policy implemented is $x \sim U[0,1]$ with women's (men's) preferences ω distributed over $U[0, W]$ ($U[M, 1]$). Cost of running for is δ_g , $g \in \{w, m\}$. Utility for individual i is given by $a - |x_j - \omega_i|$ for non-candidates and $-|x_j - \omega_i| - \delta_i$ for candidates. Successful candidates would implement $x_j = \omega_j$. In DC's refinement, it is $x_j = \alpha\omega_j + (1 - \alpha)\mu'$ where μ' captures the policy option by the local elite which is arguably male-dominated. This models has a number of predictions, among the important ones are:
 - **Under certain conditions, women may not even run**
 - **Reservation policy can alter the policy outcome**
- Issues raised by men and women are taken as a measurement of their preferences:

- Women raised the issue of drinking water more often in both districts and road improvement more often in West Bengal (less often in Rajasthan)
- Men raised the issue of irrigation and education more often in both districts
- This is ***perfectly in line with public good investment*** of GPs with reserved female chiefs relative to GPs without reserved chiefs; most effects are individually (for the respective good) significant and all are jointly significant.
- They also find that, controlling for the effect of a reserved seat and different preferences, that ***the extent to which the preferences are different is insignificant*** in West Bengal and sign. pos. in Rajasthan. Insignificance says that women do invest more in goods preferred by women because this is in line with their own preferences but, *beyond that*, are ***not more responsive to women's needs relative to men's needs***. Significance would suggest the opposite.

2.3 Empirical Strategy and Discussion

2.3.1 Besley & Burgess 2002

- Basic regression:

$$g_{it} = \alpha_i + \beta_t + \delta s_{it} + \gamma(s_{it})(z_{it}) + \phi z_{it} + u_{it}$$

where g is responsiveness (public food distribution or calamity relief expenditure), s is the need for state intervention, and z a variable of economic, political and media covariates (newspaper circulation, turnout, political competition, income, population, urban population share). While ϕ captures general *activism*, γ is supposed to capture *responsiveness* to specific shocks.

- Concerns:
 - **Newspaper** variable might be **endogenous**. Some underlying social development or education variable *not* in z may be driving both media coverage and g beyond what is captured in z . However, the fact that only **local language** newspaper circulation has a significant effect in the language differentiation would stand against that.
 - **Reverse causality** between g and (local) newspaper circulation
 - Potential **endogeneity of crop production**
- Strengths:
 - **IV** carried out to deal with above concerns as a **robustness** with newspaper ownership as instrument. However, any underlying variable may for endogeneity just as well drive g and the instrument, i.e. the exclusion restriction may not be fulfilled. This may particularly apply to the reverse causality issue. In terms of **private vs. state ownership of the newspapers** which certainly affects circulation and thus government responsiveness, the exclusion restriction arguably holds.
 - **External validity** (16 states, 34 years)

2.3.2 Finan & Ferraz 2008

- Basic regression:

$$E_{ms} = \alpha + \beta_0 C_{ms} + \beta_1 A_{ms} + \beta_2 M_{ms} + \beta_3 (A_{ms} \times M_{ms}) + \beta_4 (A_{ms} \times C_{ms})$$

$$+ \beta_5(M_{ms} \times C_{ms}) + \beta_6(A_{ms} \times C_{ms} \times M_{ms}) + X_{ms}\gamma + v_s + \epsilon_{ms}$$

Where E is electoral outcome, C is corruption, A is a pre-election audit dummy, M is the # of local radio stations, and X is a vector of covariates. Municipalities and states are indexed by m and s respectively. Interaction with corruption level ensures that the differential effect for *constant* corruption is measured which is important because the pre-election audit effect might depend on how much corruption there is *relative to priors*. Semi-parametric approach with own **dummy for each possible level of corruption** is carried out too. Interaction with media measure is relevant to see the effect of the accessibility to information about the incumbent after audit.

- Concerns:
 - The availability of **local radio stations** is not random and might be **endogenous** (for the same reasons as above in Burgess and Besley 2002, e.g. different education levels or population size)
 - Pre-election audits were different from post-election audits (e.g. **corrupted auditors** before elections). However, this is unlikely as auditors were regionally supervised, the corruption realisations are balanced and auditors are well-paid public employees selected after a public examination.
 - Corruption could just be a **proxy for political inability** because of which incumbents were not re-elected. Audits showing high corruption were interpreted to reveal political inability. However, given political ability influences the outcome of elections *generally* and the level of corruption is correlated with political ability, there should be some effect for post-election audit municipalities as well which we don't see. Also, there is no correlation between re-election and education level which is arguably correlated with political ability.
 - Revealed corruption levels might have **altered campaign strategies** by incumbents and opponents which might have been the channel of effect for re-election, rather than just the information about corruption. However, this would also suggest that the earlier pre-election audits took place (relative to the election), the stronger should be the results which is not the case (no difference). Also, this would suggest that the effect is stronger for higher political competition which is not observed either.
 - Mayors revealed to be corrupt might have got **less funding** for campaigns. However, including an interaction between pre-election audits and campaign contributions is not significant.
- Strengths:
 - The audit variable is **randomised**
 - Some **robustness against potentially endogenous radio variable** delivered: interaction terms with potential confounding factors included. Triple interaction effect remains significant and similar in size.
 - As another robustness checks, **different media sources** like newspapers and television are investigated. These might, however, suffer from the same endogeneity problems.

2.3.3 Duflo & Chattopadhyay 2004

- Because of randomisation, simple mean comparison suffices:

$$E[Y_{ij}|R_j = 1] - E[Y_{ij}|R_j = 0]$$

where Y is an outcome of interest (investment in a specific public good) and R indicates whether the chief position was reserved for a woman or not. NB: This comparison captures the ATE *in a world with reservation*. It does *not* capture the causal effect of having reserved positions for women relative to having no reservation system at all.

- To answer the question whether differences come from the fact that female leader preferences are more in line with women's preferences or, possibly beyond that, from the fact that female leaders are more responsiveness towards issues raised by their "peers", they run:

$$\begin{aligned} Y_{it} = & \beta_7 + \beta_8 R_j + \beta_9 (D_i \times R_j) + \beta_{10} (D_{ij} \times R_j) + \beta_{11} (S_{ij} \times R_j) \\ & + \beta_{12} S_{ij} + \beta_{13} D_{ij} + \sum_{l=1}^N \beta_l d_{il} + \epsilon_{ij} \end{aligned}$$

where d are good-specific dummies and D_i is the average difference between the share of requests about good i from women and from men (for goods more requested by women, it's positive; for goods more requested by men, it's negative – and this is captured by differential public good investment Y for the respective goods which means the coefficient is positive). D_{ij} is the difference between indicators for whether issue i was brought by women or men in village j , basically capturing how salient an issue is for women relative to men across all villages. S_{ij} is the sum of the indicators, i.e. how salient an issue is generally for men and women.

- Concerns:
 - **Other characteristics of leading women** than gender in reserved villages may be driving the effects, e.g. that they are "weaker" on social and economic characteristics (education, literacy, poverty, household assets, less political experience) than leading men in unreserved villages. However, controlling for these covariates in a robustness check, doesn't change the relevant coefficients.
 - The effects may be driven by the fact that **women may not be willing or able to run again for election** (the randomisation for reserved seats rotates every term). Controlling for this in a robustness checks, does not affect the relevant coefficients though.
- Strengths:
 - **Randomisation** means credible exogeneity of R
 - Especially the fact that **preferences on roads** are *different* for men and women in the two regions looked at (in West Bengal, it's more important to women while it is more important to men in Rajasthan) and this is reflected in differential investment going in opposite directions in the two regions as well is very convincing evidence for the theory (for all other issues, the relative preferences between women and men are the same in the two regions)

2.4 Policy Conclusions

- Beyond economic development, **improved accountability through media** plays a great role in **determining government action** – in line with World Bank's efforts to improve "governance".

- In particular, minorities who have **common interest** can profit from media to affect policy choices.
- Information on politicians' activities promotes accountability through enabling voters to **reward good types and punish bad types**. This channel, however, only works if the information is made publicly **accessible** through media coverage
- Even at very decentralised/low level of government, **politicians' identities** and requirements on them (such as gender reservation) has a **policy impact**. The candidate-citizen model seems to apply.
- Specifically, **reservations for women** which are increasingly implemented in numerous countries, seem to be a good idea to promote accountability, given the citizen-candidate model seems to apply.

2.5 Further papers

Banerjee et al. (2012). *Corruption*. WP.

- Overview / Lit Review / Framework / Future research outlook on corruption
- Empirical challenge: measurement, observability. Corruption inherently intransparent
- Theoretical challenge: good model / concept going beyond pure moral hazard aspect. They come up with different environments for corruption which lead to different forms of corruption.
 - The typical one is the nature of monitoring and punishment
 - Another one is intrinsic motivation (e.g. it is often observed that bureaucrats try to justify their taking money even when it's clear that it is a bribe.) This may differ across persons or countries
 - The nature of the economic decision
- they define corrupt behaviour as bureaucrat's behaviour in public office for personal gain involving the *explicit breaking of rules*. This makes measurement easier (no ethical judgement necessary)
- they see future research fruitful in
 - understanding links between corruption and competition. Sometimes, more competition may lead to more corruption
 - links with illegality and nontransparency: if rules keep being changed, is corruption actually increasing because citizens are so confused about the rules?
 - learning: for corruption, the rules of the game have to be acquired in some learning process as they are not public.
 - norms of corruption and signalling
 - social psychology of corruption (see above): corruption could potentially be much higher in many settings if people did not have the desire to legitimate.

3 Ethnicity and Development

3.1 Papers and General Introduction

(Main) papers discussed:

- i. Burgess, Jedwab, Miguel, Morjaria, Miquel 2013. *The Value of Democracy: Evidence from Road Building in Kenya*, Working Paper.

- Assess how ethnicity of the president affects asymmetric public infrastructure investment in times of autocracy vs. democracy
- ii. Hodler & Raschky 2014. *Regional Favoritism*, QJE.
 - Assess the effect of a country leader's birth region on the region's economic activity as measured by night-time light and its interaction with different variables
- iii. Easterly & Levine 1997. *Africa's Growth Tragedy: Policies and Ethnic Divisions*, QJE.
 - Cross-country study on how ethnic diversity explains differences in many economic factors driving growth across countries

General Theme:

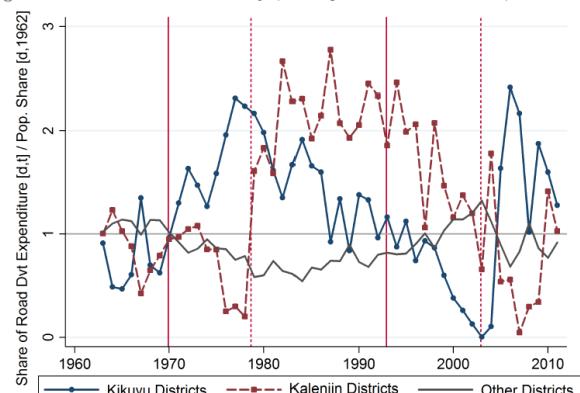
- Ethnical diversity may undermine development in several ways:
 - **Misallocation** of goods through ethnical favouritism
 - Suboptimal **social policies**
 - More independent **bribe takers**
 - **Common pool problem**: pool of rents is over-exhausted by rent takers from different ethnicities
- Importance of ethnic diversity in **economic development** and its pronouncement under **different political circumstances**
 - BJMM 2013: its effect on **misallocation** through **ethnic favouritism** under autocracy relative to **democracy**
 - HR 2014: its effect on **misallocation** through **regional favouritism** (not necessarily the same) and interaction with **political and economic factors** worldwide
 - EL 1997: its **effect on growth directly and indirectly** through important political and economic determinants of growth

3.2 Main Ideas and Results

3.2.1 Burgess et al. 2013

- Uses micro-data on road expenditure in post-Independence Kenya (1963-2011) to
 - Quantify the **extent of ethnic favouritism**
 - Determine whether **democracy vs. autocracy** affects this
- The date is convenient because there is a **high degree of ethnic segregation** and almost every district is **dominated** by one ethnic group
- Kenya is a very convenient country because, post-independence, it had **changes in the ethnicity of the president** and within the time in power of one president **changes in the political system**
- Findings: **Co-ethnicity** with the president has generally a highly significant effect on **road development expenditure** in the respective ethnic districts (in which the majority of people belongs to the respective ethnic group). When interacting with a **democracy** dummy, this effect is **completely offset**,

Figure 5: Road Investment in Kikuyu, Kalenjin and Other Districts, 1963-2011

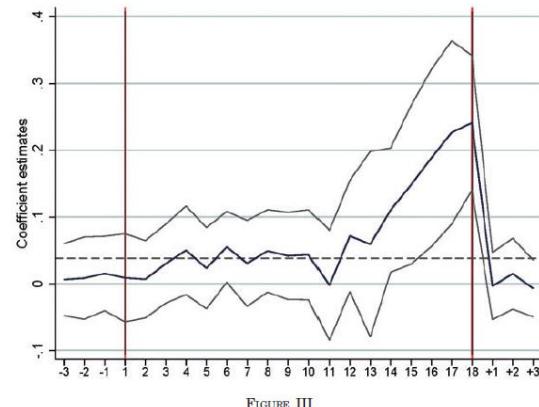


i.e. during democracy, there is no significance any more (this holds for all specifications). In line with figure (vertical red line = political system change, vertical dotted line = president's ethnicity change)

- Only **significant for president**, not for other cabinet members, except the **vice president** who is usually of a different ethnicity. He influenced the decision to some extent and ethnicities which were neither in common with the president nor vice-president really received relatively much less road building investment.
- They also **replicate Easterly-Levine's** growth accounting exercise, including the years 1980s to 2000s and a democracy dummy with interaction variable. Without the democracy dummy, the ethnicity effect disappears while, including the democracy dummy and interaction variable, ethnic diversity displays a sign. **negative effect on growth which is completely offset by democracy**. Serves as confirmation of the importance of the political system / institution.

3.2.2 Hodler & Raschky 2014

- Use **night-time light satellite** data as a proxy of GDP for almost 40 000 subnational regions in 126 countries to
 - Detect **regional favouritism**
 - Determine under which conditions (**political institutions, education standard, foreign aid flows, oil rents**) it is more or less pronounced
- Influence of leader on leader's birth region's night-time light intensity is highly significant under region and country-year FE. The effect **abruptly stops** when the leader is not in power any more.
- **Weaker institutions** (polity score), higher **education** and **national GDP weaken** the effect. This interaction is strongest for the polity score – a high polity score (democracy) basically completely offsets the regional favouritism effect relative to a low score (autocracy).
- Regional favouritism is stronger in countries where **ethnical diversity** is great but this effect vanishes once all interactions are taken into account. **Institutions** and **education** seem to be the **main determinants** of how pronounced regional favouritism is
 - Difficulty in interpretation given the result on ethnic diversity: regional favouritism might work along different lines than ethnic favouritism.
- **Aid flows increase** regional favouritism (again, this effect is decreased for better political institutions) while **oil rents have no significant effect** (neither for low nor high polity realisation)
- The use of different geographical areas as dependent suggests that regional favouritism is **very local**.

FIGURE III
The Dynamics of Regional Favoritism

3.2.3 Easterly & Levine 1997

- Explain why Africa has grown more slowly than other regions by ethnic diversity

- Explore **direct effect** of ethnic diversity (using mainly ethnolinguistic diversity as a measure, although robustness for other measures is provided)
- Explore **indirect effect** through the influence on schooling, political instability, financial development, government deficits, infrastructure, etc.
- They find a significant **negative direct effect** of ethnic diversity on growth when controlling for economic covariates (income, schooling, financial depth, black market premium, gvnt. deficit, telephones per worker, assassinations), decade dummies and continental dummies. The effect becomes modest and less significant when *all* covariates are included.
- They also find significant effects of ethnic diversity on the **economic covariates** except assassinations and gvnt. deficits which suggests that policies differ for different levels of ethnic diversity.
- Growth accounting part is bringing this together: Out of 3.4ppt growth differential between Africa and East Asia, the covariates and ethnic diversity (directly) together can account for 1.7ppt., while, tracing back the influence of ethnic diversity on the covariates, ethnic diversity directly and (mostly) indirectly accounts for 1.0ppt, i.e. **28% of the overall observed growth differential**.
 - This cross-country growth accounting exercise is “heroic” and **extremely crude**. It may give a sense that ethnicity plays a role, but no more than that.
 - These calculations are net of the convergence effect: At lower initial income levels, Africa should have grown 1.1ppt faster than East-Asia. 1.7ppt is net of this.
- In both a regression and an income accounting exercise, the authors also find a significant impact of ethnic diversity on the level of income
- Interpretation of the results (direct and indirect effect)
 - Ethnic diversity encourages the **adoption of growth-retarding policies** which foster **rent-seeking behaviour**
 - Ethnic diversity makes it more **difficult to form a consensus for growth promoting public goods**

3.3 Empirical Strategy and Discussion

3.3.1 Burgess et al. 2013

- Estimating Equation:

$$\text{road}_{dt} = \gamma_d + \alpha_t + \beta(\text{coethnic district}_{dt}) + \delta(\text{coethnic district}_{dt} \times \text{democracy}_t) + \theta(X_{d1963} \times [t - 1963]) + u_{dt}$$

where road is the share of road expenditure in d to the population in d relative to the share of road expenditure to the population in the whole country. X are covariates (baseline population, area, urbanisation rate, formal earnings, formal employment share, value of cash crops, location) which are interacted with **linear time trends** (allow for diff. effects over time)

- Concerns:
 - Quite difficult to make a case for **OMV wrt. co-ethnicity**: district fixed effects included and relevant covariates, even with linear time trend. Also, it is questiona-

- ble if the relationship with economic development as underlying variable is so different in democracy times.
- **Endogeneity of democracy.** There may have been structural changes that caused both change in the political system and weakening of ethnic favouritism. What led to democracy? However, strong **co-movement** between changes in political systems in Kenya and **sub-Saharan Africa** suggest exogeneity.
- Strengths:
 - Delivers the same results for **two measures** of the dependent: road building and road expenditure
 - Tiny changes in β and δ when different sets of covariates are used (**robustness to covariates specification**)
 - Distinguishing between five periods of leadership, you observe a significant difference in road building between co-ethnic and non-co-ethnic districts only under autocracy. Further robustness to results.

3.3.2 Hodler & Raschky 2014

- They run:

$$Light_{ict} = \alpha_i + \beta_{ct} + \gamma Leader_{ict-1} + \delta_j INT + \epsilon_{ict}$$

where α_i is a regional dummy, β_{ct} captures country-time fixed effects, and INT are all sorts of interactions looked at (my addition to the equation)
- Concerns:
 - **Endogeneity of leader**, i.e. some underlying OMV driving both GDP of a certain region and the probability of the country leader coming from this region. However, they provide robustness in so far as **future and past variables** on leaders do **not** have a **significant** effect (which they should in this case as an OMV to this end that varies over time – region FE included – would certainly not drive the abrupt sort of effect we observe, but reflect some more smooth development)
 - **Endogeneity of interaction variables**, especially polity and education
- Strengths:
 - Very **clear trend** observable / neat evidence
 - **External validity** as it covers 126 countries

3.3.3 Easterly & Levine 1997

- To assess the influence of covariates and ethnic diversity on growth rates, a simple regression with these regressors using Seemingly Unrelated Regressions, i.e. averaging over separate regressions for each time period.
- **Growth Accounting** for Africa-East Asia growth differential is done by multiplying the difference in explanatory variables between these regions with the regression coefficient from above.
- Concerns:
 - Massive **country heterogeneity** issue ignored: Only continent dummies are included. Beyond the few covariates, many omitted variables, heterogeneous across countries, could cause endogeneity: geography, more specific education measures (relating to languages), social policy, etc.
 - **Growth rates** and dummies **for decades** looked at in order to overcome business cycle problems and capture long-run growth. However, this is very crude.

- **Measurement error** for the ethnic measure. However, robustness is provided to the extent that other measures deliver similar results.
- Strengths:
 - **Big picture question.** This is an analysis of two continents which gives an idea that ethnicity seems to play a role for growth and development. However, specific figures should not be taken too seriously.

3.4 Policy Conclusions

- Ethnic diversity may have a considerable **impact** on a country's **growth and development in general**
- This impact works along different lines:
 - It might encourage the adoption of **growth-retarding policies** that foster **rent-seeking behaviour**
 - It may make it more difficult to find a **consensus for growth promoting public goods**
 - Much clearer evidence on **misallocation**: Favouritism of the ethnic group / region of the political leader that is purely economically speaking inefficient
- The extent of favouritism depends on political factors. **Political institutions** and **education** seem to play a key role. In particular, **democracy** – as opposed to autocracy – may completely offset favouritism.

4 Conflict and Development

4.1 Papers and General Introduction

(Main) papers discussed:

- i. Dube & Vargas 2013. *Commodity Price Shocks and Civil Conflict: Evidence from Colombia*, Review of Economic Studies.
 - Evaluates the influence of price shocks on labour and capital intensive goods
- ii. Yanagizawa-Drott 2014. *Propaganda and Conflict: Evidence from the Rwandan Genocide*, QJE.
 - Assesses the effect of information on the extent of conflict
- iii. Miguel, Satyanath, Sergenti 2004. *Economic Shocks and Civil Conflict: An Instrumental Variable Approach*, JPE.
 - Tries to identify economic causes of conflict

General Theme: Conflict seems to be more prevalent in developing countries and has **huge economic and political cost** (estimate suggests that one year in conflict reduces growth rate by 2.2%). Very recently, the **economic route to conflict** has been investigated and evidence became more reliable (not simple cross-country regressions):

- MSS 2004: **Economic shocks** can have an influence on conflict in general
- DV 2013: There are two effects for price shocks: the **opportunity cost effect** and the **rapacity effect**
- YD 2014 (not so much economic route): **Mass media** has a direct and indirect effect on conflict

4.2 Main Ideas and Results

4.2.1 Dube & Vargas 2013

- Using data on conflict events in over 950 **Colombian municipalities** between 1988 and 2005, they assess the effect of coffee and oil price shocks on conflict in Colombia. There are two theoretical channels:
 - **Opportunity cost effect:** positive price shocks increase wages for the respective good and thus decrease labour supplied to criminal / conflict activity (increased opportunity cost)
 - **Rapacity effect:** positive price shocks increase the return to predation of the respective good and promote conflict over the resource
- They find **reduced conflict** as a response to **increased coffee prices** (opportunity cost dominates) and **increased conflict** as a response to **increased oil prices** (rapacity effect dominates)
 - For coffee, the response is significant for all four measures of conflict looked at (guerrilla attacks, paramilitary attacks, clashes, casualties); for oil, the only significant coefficient is on paramilitary attacks. The authors argue only paramilitary attacks are affected because of the geographic concentration of oil production to areas where paramilitary groups originate from
 - These results are **robust** to estimations using **further (valuable) natural resources** like coal and gold and **further (labour-intensive) agricultural goods** such as bananas, sugar, African palm and tobacco
- Looking at **wages and hours** for the opportunity cost effect and **capital revenue, paramilitary and guerrilla political kidnappings** for the rapacity effect as dependent variables separately, they find that indeed **only the opportunity cost mechanism** is at play for **coffee** (sign. coefficients only for wages and hours) and **only the rapacity mechanism** is at play for **oil** (sign. coefficients for capital revenue and paramilitary kidnappings only)

4.2.2 Yanagizawa-Drott 2014

- Use nation-wide village level data on **persons prosecuted** for crimes during the Rwandan genocide to look at the direct and indirect effect of radio propaganda by radio **RTLM** on violence:
 - **Direct/fundamental effect:** propaganda **signals a high value of participation** in violence which may induce individuals to engage in violent activity. However, it is unclear whether such an effect is strong enough given potential intrinsic aversion against violence, risk and material opportunity cost
 - **Indirect/strategic effect:** people know that others listen to radio too which impacts on their propensity to participate ambiguously
 - **Strategic complements:** marginal benefit of participating increases in # of participants because of safety in numbers, desire to conform, etc.
 - **Strategic substitutes:** marginal benefit of participating decreases in # of participants because of free-rider incentives, etc.
 - These effects are **neutralised** if people have access to **independent information**

- The **direct effect** is found to be highly sign. **positive** for **militia** and **total violence** and sign. positive (10%) for **individual violence**.
- There are highly significant **positive spillover effects** for **militia violence within 10km**, no sign. **change** for **individual** or even **total violence** though – only the share militia/total violence increases significantly.
- **Counterfactual** analysis gives an insight into how massive the effects of radio propaganda were: **9.9% of total violence** was caused by RTLM, for militia violence excluding spillovers it was 7.7%.
- **Heterogeneous effects** analysed controlling for all sorts of interaction:
 - **Primary education** interaction yields that education reduces violence significantly although this is theoretically ambiguous: in school, people may be taught government ideology which would increase violence in education; or through education, people may become more critical independent thinkers which would decrease violence in education
 - **Size of the Tutsi minority** interaction yields that violence decreases in the size of the minority

4.2.3 Miguel et al. 2004

- Performs a **cross-country IV** analysis for **40 African countries** between 1981-1999
- Economic growth **instrumented by rainfall** has a significantly **negative impact** on civil conflict (defined as at least 25 battle-related deaths per year and per incompatibility)
 - 5ppt drop in GDP growth increases likelihood of civil war in the next period by over 12ppt or 50%
 - Interpretation: **Opportunity cost effect** – armed militia groups have an easier time recruiting combatants in times of negative growth
- **Interaction** specifications yield that impact of income shock is **not significantly different** for **richer or more democratic or more mountainous or more ethnically diverse countries**

4.3 Empirical Strategy and Discussion

4.3.1 Dube & Vargas 2013

- They run the following (2nd stage) regression:

$$y_{jrt} = \alpha_j + \beta_t + \delta_r t + Coca_{jrt}y + (Oil_{jr} \times OP_t)\lambda + (Cof_{jr} \times CP_t)\rho + X_{jrt}\phi + \epsilon_{jrt}$$

where y is the respective conflict outcome in municipality j , region r , Oil is the oil production level in 1988 (thus no endogeneity as this is in the first period), Cof is the coffee production level in 1997 (mid-sample period; because of endogeneity concern **instrumented by rainfall and temperature**), OP is the international oil price (exogenous to Colombia as it is a tiny oil producer globally) and CP is the coffee price **instrumented by the export volume** of the other three leading coffee exporting nations Brazil, Vietnam and Indonesia (Colombia is one itself which causes endogeneity issues and necessitates instruments). X are covariates including population, $\delta_r t$ are regional linear time trends and $Coca_{jrt}$ are linear time trends for coca and non-coca cultivating municipalities.

- λ and ρ capture the **DID** estimates for oil and coffee prices respectively (differential changes in coffee or oil municipalities)

- Regression for separate analysis of the two mechanisms just takes wages, hours, capital revenue, paramilitary political kidnapping and guerrilla political kidnapping as dependents and a different vector of covariates – the rest stays the same
- Concerns:
 - Reverse causality issue could persist for the instrument used if **exporting nations** adjust their exports according to Colombian coffee prices and thus react to differential violence increases in Colombian coffee regions (**violation of exclusion restriction**) To some extent, robustness is provided by the ability to explain international price dynamics and expansion decisions by some national and international events and policies unrelated to conflict in Colombia. This is not very rigorous however.
 - **Mean reversion:** We only have coffee and oil production for one point in time, although an average over several points in time would probably be better.
 - **Rainfall and temperature** as instruments have effects on other agricultural products through which conflict is influenced – potential **violation of the instrument's exclusion restriction**
 - **Local shocks** may impact on **national production** (in particular for oil) and what we pick up is an effect on that level (interpretation issue).
- Strengths:
 - **Instruments** to overcome major endogeneity issue of the coffee price (reverse causality argument would say that increased conflict could decrease coffee production leading to price increases in Colombia and globally, i.e. going against the hypothesis); however validity of instrument (s.above)?
 - **Separate analysis** for the two mechanisms neatly confirms the results
 - They provide **robustness** against a range of alternative explanations.
 - **Migration** (which could have driven the effect) does not significantly react to price changes.
 - **Government enforcement** increases in coffee areas when coffee prices decrease (which goes against hypothesis that there may be more violence because the government has less resources to spend on military inducing less enforcement and promoting armed group attacks)
 - No effect on **coca production** (which could have served as substitution for coffee production and involves more drug fights)
 - **Cross-sector spillovers** are not observed, i.e. decreased coffee wages led to people being recruited by groups that predate on oil rents. The effects are indeed separate.

4.3.2 Yanagizawa-Drott 2014

- To test the direct effect, they run:

$$\log(h_{vci}) = \beta_v r_{ci} + X'_{ci} \pi + \gamma_c + \epsilon_{ci}$$

Where h is the # of prosecuted persons for crime type v (individual or militia) in commune c and village i (communes > villages). r is the radio coverage, X are covariates. The covariates include latitude, longitude and second-order polynomials in the distance to the nearest transmitter, village mean altitude and variance of altitude within the village. Along with **commune FE**, this ensures that only a **very local variation in radio cover-**

age within communes is **exploited** – this variation would virtually only come from hills lying between transmitter and receiver which is arguably random

- The indirect effects are estimated using:

$$\log(h_{vci}) = \lambda_{vd}\bar{r}_{dci} + \bar{X}'_{dci}\phi_d + \gamma_c + \epsilon_{ci}$$

Where upper bar denotes *other* villages within distance d .

- Concerns:
 - **Simultaneity problem** in indirect effects estimation: conflict in other villages certainly plays a role and this biases all estimates in the above given reduced from. This is surprisingly not dealt with at all! We are only left with a comment that such peer effects may exist and this is “beyond the scope of this article”
 - **Measurement error** in dependent measuring prosecutions not actual participation in killings: if the error is correlated with radio coverage, e.g. negatively because more killings mean less witnesses, the true effect is underestimated.
- Strengths:
 - **Exogeneity check** for radio coverage after controlling for the bunch of variables given above: it is indeed not correlated with any village variables that would suggest endogeneity in the remaining variation of r
 - **Placebo test** using another (**non-propaganda**) **radio** channel is successful

4.3.3 Miguel et al. 2004

- They run the following IV-2SLS:

$$1^{\text{st}} \text{ stage: } growth_{it} = a_{1i} + X'_{it}b_1 + c_{1,0}\Delta R_{it} + c_{1,1}\Delta R_{i,t-1} + d_{1i}year_t + e_{1it}$$

$$2^{\text{nd}} \text{ stage: } conflict_{it} = \alpha_{2i} + X'_{it}\beta_2 + \gamma_{2,0}growth_{it} + \gamma_{2,1}growth_{i,t-1} + \delta_2year_t + \epsilon_{2it}$$

Where R is rainfall and the subscripts 1 and 2 just denote the equation number

- Concerns:
 - **Exclusion restriction of rainfall**: much rainfall may also directly decrease conflict because it has an effect on the infrastructure of a country (floods, etc.). Robustness against that: Road network data from World Bank used without significant effect; heat waves may raise temper. However, the lagged variable is significant in the case of the 25 deaths definition which doesn't match this sort of direct influence.
 - **Different interpretations** of the growth channel than **labour market effect** possible: e.g. negative growth promotes more inequality which translates into more resentment across social classes
 - **Cross-country endogeneity** problems with **interaction** variables, although this is only a minor aspect of the paper. Particularly weak institutions and conflict may have an obvious reverse causality problem
- **Data limitations**:
 - No within-year **monthly** data
 - Only **politically motivated conflicts** captured
- Strengths:
 - **High significance** and **robust estimators** in the **first stage** to different specifications; plus virtually zero estimate for the placebo lead rainfall regressor

- Measurement error for GDP addressed through **instrumental approach**
- For **alternative definition** of conflict as >1000 deaths, **still a significant coefficient** on current economic growth is found

4.4 Policy Conclusions

- There is an **economic route** to conflict
 - Generally, shocks to economic growth tend to have an impact on conflict
 - more specifically, price changes can influence the extent of conflict via the **opportunity cost effect** and the **rapacity effect** depending on whether the commodity is rather **low-value and labour-intensive** or **high-value and capital-intensive**
 - To counter the opportunity cost effect, **insurance schemes** or **price stabilisation schemes** might be a viable policy option
 - To address the rapacity effect, **less fiscal decentralisation** may be an option as local revenues invite (local) predation
- Information provision and media coverage is not always beneficial to society. **Propaganda may heavily exacerbate conflict**. Policy lessons could be to **restrict (particularly state-sponsored) mass media** in times of civil conflict or genocide.

5 Motivation of Public Servants

5.1 Papers and General Introduction

(Main) papers discussed:

- i. Dal Bó, Finan, Rossi 2013. *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*. Forthcoming: QJE
 - Assess how higher wages impact on the pool of applicants and acceptance decision for a public servant job
- ii. Ashraf, Bandiera, Lee 2014. *Do-gooders and Go-getters: Career Incentives, Selection, and Performance in Public Service Delivery*. Working Paper.
 - Evaluate how salience of career vs. social benefits of a job influences the type of applicant and performance on the job in public service
- iii. Khan, Khwaja, Olken 2014. *Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors*. NBER Working Paper.
 - Investigates the impact of incentive schemes on performance of tax collectors in Pakistan

General Theme:

- It is unclear how public servants who – as opposed to politicians – do not face elections are motivated to work effectively. (Monetary) incentives would work through two channels: selection of different types into public service jobs and increased effort on the job.
- For a public service job, not only qualification, but also pro-social motivation is important. Incentive schemes could potentially attract better qualified types but also crowd out pro-social motivation as the social motivation “tax” (difference between higher wage in the private sector and wage in the public sector) reduces. This might not hold true if qualification and pro-social motivation are in fact positively correlated

- DBFR 2013: higher wages attract more qualified but *not* less motivated applicant pool for public service jobs and help bridge less favourable job conditions
- ABL 2014: stressed career incentives attract more qualified but *not* less motivated applicant pool for public service jobs and improve performance on the job
- KKO 2014: incentive schemes improve performance of tax collectors and lower bribes in some cases while also increasing bribe payments in other cases

5.2 Main Ideas and Results

5.2.1 Dal Bó et al. 2013

- Using **randomised wage and job offers** in Mexico for 350 (real) positions as Community Development Agents (public servants) in the context of a Rural Development Programme launched in 2011, they answer three questions:
 - Do higher wages attract more qualified applicants? Yes, higher wages highly significantly **improve all job skills measures**, and significantly affect some **personality traits** in a favourable direction (increase conscientiousness, reduce neuroticism).
 - Do higher wages crowd out pro-social behaviour? No, in fact, higher wage offer induces most **public service motivation and prosocial behaviour measures** to be **significantly higher** (only exception: measure of charity done in the past year is significantly lower)
 - How do higher wages influence acceptance decisions, especially under worse job conditions? Generally, higher wages induce sign. **higher acceptance rates**. Worse **working conditions** (higher distance, higher drug-related deaths in the community, lower human development index in the community) all have (individually) significant neg. impact on acceptance decisions. In interaction with higher wage offers it is shown that higher wage offers reverses this effect, i.e. helps to “**bridge the recruitment gap**”
- Underlying theory: For market quality v , public service motivation π , wage offer w and a mean zero random outside utility shock ϵ , candidates will accept if $v + \epsilon < w + \pi$. Showing up for an interview incurs cost c .
 - From this, it can be shown that the **average quality of the applicants increases in w** ; and the **average π of the applicants**
 - **Decreases in w if v and π are independent.** The separating type \bar{v} who is just indifferent between applying and not increases in w
 - **Increases in w if v and π are sufficiently positively correlated**
 - The **labour supply elasticity** equals the elasticity of the applicant pool size + the elasticity of the convergence rate (the rate at which offers are converted into filled vacancies) – elasticities wrt. wage. From above, the elasticity of the applicant pool size is positive, but the **conversion rate elasticity is ambiguous**. Two effects play a role and the overall sign depends on which one dominates: Higher wages make infra-marginal types $v < \bar{v}$ more likely to accept but the added marginal types less likely to accept. The convergence rate elasticity is found to be positive, i.e. the **inframarginal effect dominates**.

5.2.2 Ashraf et al. 2014

- Using **randomised salience** of **career** (treatment) vs. **social incentives** (control) on different job advertisements in different communities for a Community Health Assistant (CHA) placement programme in Zambia in 2010, the authors answer two questions:
 - How do career incentives influence **selection** into public service?
 - Sign. pos. treatment effects in **skills** for applicants, but not so for selected trainees anymore
 - **No crowding-out of pro-social motivation.** In fact there are neither significant differences in pro-social motivation between treatment and control for applicants nor for selected trainees (only exception: belonging to a village community or self-help-group which is even sign. *higher* for career incentive selected trainees).
 - Sign. treatment effect in **career motivation** for both applicants and selected trainees
 - How does the differential self-selection into applying impact on **performance on the job?** For all three tasks (main: household visits, secondary: assisting at health posts, community mobilisation), the treatment effect is **highly sign. positive and substantial.**

5.2.3 Khan et al. 2014

- Using administrative data on (property) tax collection, property survey data on taxpayer satisfaction, assessment accuracy and bribes, and phone survey data on self-reported effort by tax inspectors, to investigate the effect of tax collector incentive scheme on
 - Tax revenues collected
 - Non-revenue outcomes such as accuracy and taxpayer satisfaction
 - Corruption / bribes
- The following incentive schemes were randomised among tax collectors:
 - Revenue treatment: bonus is purely based on higher tax revenues
 - Revenue Plus treatment: bonus payment based on tax revenues as well as taxpayer satisfaction and tax assessment accuracy
 - Flexible Bonus treatment: broad range of (partly subjective) criteria by some committee determined bonus payment
- Findings:
 - **Sign. pos. effect** of treatment on **tax revenue** which is **most pronounced for the revenue treatment.** The flexible bonus treatment hardly displays any significance. The effects are largely driven by changes in the tax base (as opposed to changes in the non-exemption or recovery rate; tax revenue = tax base * non-exemption rate * recovery rate where recovery rate = revenue/tax base after exemptions)
 - **No strong** (hardly any significant) effects of any treatment on **non-revenue outcomes** such as quality, satisfaction, accuracy, tax gap
 - There is a significant and substantial pos. **effect of reassessments on tax payment** and a **significant pos. treatment effect on bribe payments** and their frequency in general. The **interaction** of reassessment and treatment yields a **sign. pos. coefficient on tax payment** and a **sign. negative coefficient on the frequency of bribes** paid.

- Theory: In a bargaining setting between taxpayer and collector, the introduction of performance pay posits a **higher cost of not reporting** for the tax collector which translates into higher (and/or more frequent) bribes demanded. On average, this should lead to **some bargaining breaks down**, i.e. instead of bribes, taxes are now paid (especially large properties with large tax bases) and **increased bribe payments** for other properties.
 - Third **empirical finding** is perfectly **consistent** with this: the additional tax payment from reassessment is higher while bribe payments for reassessed properties is lower in the treatment group (bargaining break downs through increased cost). Generally, w/o considering reassessments, treatment increases bribe payments (higher cost for tax collectors translates into higher bribes for non-reassessed properties)
- Tax collectors seemed to focus on taxing a small group of largeholders more (instead of a large group of smallholders)
 - There is a **trade-off**: smallholders have more democratic influence per dollar, but largeholders are more influential absolutely speaking.

5.3 Empirical Strategy and Discussion

5.3.1 Dal Bó et al. 2013

- Given randomisation, the estimations are straightforward:

$$Y_{icr} = \beta_1 T_c + \zeta_r + \epsilon_{icr}$$

$$A_{icms} = \gamma_1 T_c + \gamma_2 (T_c \times W_m) + \gamma_3 W_m + X_i' \beta + \zeta_s + \epsilon_{icms}$$

where Y is the skill/motivation/behaviour measure of i in locality c and region r/s (locality $\ll r$), A is an acceptance dummy, T is the treatment dummy, W is a characteristic of the assigned municipality m (distance, drug-related deaths, HDI) and X are individual characteristics.

- Concerns:
 - **Measurement error** for personality traits, pro-social motivation which could be systematic – less pro-socially motivated people might be more inclined to lie on such tests. However, correlations in the data do not seem to suggest misreportings according to the authors
- Strengths:
 - **Randomisation** eliminates endogeneity of differential wage and job offers.

5.3.2 Ashraf et al. 2014

- For the differences in applicants and selected trainees with regards to ability, pro-social motivation and career motivation, simple differences in means are looked at
- For the performance on the job, the following regression is run:

$$v_{ih} = \alpha + \beta C_i + X_i \gamma + Z_h \delta + \epsilon_{ih}$$

where v are household visits (main task), C is the treatment dummy, X captures individual characteristics, Z are area characteristics.

- There are **two channels** that influence performance on the job: **Selection** into the job and **effort** on the job which may both depend on what agents know about the job. Dif-

ferential incentives advertised on job ads opens the selection channel while providing the same actual incentives to candidates once hired closes the effort channel. This provides identification of the selection effect provided that the following important **identifying assumption** holds: **Once real benefits are known and the same to all selected trainees, the salience treatment does not affect the applicants' utility any more.**

Could fail in two directions:

- If **true value** of the job is found to be **below the advertised value**, trainees might ex-post not meet the participation constraint anymore. Drop-out of the programme, however, is found to be extremely low.
- If **true value** of the job is found to be **above the advertised value**, trainees might put extra effort in on the job which confounds the mere selection effect to be analysed. To some extent, robustness is provided by looking at the effect in different time intervals after beginning the job. The effect is very similar in three semester instalments over 1.5 years which might not be expected if there is a positive "surprise" effect that would arguably diminish over time. Not very rigorous though.
- Concerns:
 - Systematic **measurement error** in the same way as for Dal Bó et al. 2013 on the application/ **social motivation questions**
 - Systematic **measurement error** in **performance on the job** by false reporting of household visits. However, some visits were audited which may reduce or eliminate the incentive to engage in such cheating.
 - Selection effect might be confounded with "**surprise" effect**, s. above.
- Strengths:
 - **Randomisation** eliminates endogeneity concerns of career types.
 - Experimental design generally enables **identification of selection effect** (although this might not be perfectly neat, s. above)

5.3.3 Khan et al. 2014

- For the effect on tax revenue, satisfaction, and accuracy, they estimate (1) and for the effect on bribes they estimate (2) and (3) :

$$(1) \ln Y_{cst} = \alpha_s + \beta Treatment_{cst} + \gamma \ln Y_{cs0} + \epsilon_{cst}$$

$$(2) Y_{ics} = \alpha_s + \beta Treatment_{cs} + \epsilon_{ics}$$

$$(3) Y_{ic} = \alpha_c + \beta_1 Treatment_c \times ReAssessed_{ic} + \beta_2 ReAssessed_{ic} + \epsilon_{ic}$$

where Y_{cs0} indicates baseline outcome for tax circle (unit) c in stratum s and i denotes properties. $ReAssessed$ is a dummy indicating whether the property value in the sample changed, i.e. was either reassessed or newly added to the tax roll.

- Concerns:
 - **Measurement error** in survey regarding **bribes**. This data was collected by indirect questions on what people thought a bribe would be on a property similar to theirs or what the typical frequency of bribe payments is. This might well be a systematic error, e.g. underreporting in case of high value properties when bribe demands increased (treatment).
 - Incentive schemes have **more than just price effect**: inspectors receive more information and monitoring, have higher income, worked more together with

their supervisors (behavioural change) and treatment and control group knew each other (spillovers?). Authors provide some arguments, however, that price change is major effect. Also, it would be a general effect and investigating these channels in more detail provides a more nuanced understanding.

- Strengths:
 - **Randomisation** of incentive schemes.

5.4 Policy Conclusions

- **Material benefits do not crowd out pro-social motivation** and behaviour in public service. In fact, they increase the qualification of the applicants while leaving the pro-social motivation unchanged.
- People **selecting** into public service on the basis of material benefits also **perform better on the job**.
- **Incentive schemes** on the job in public service also increase performance.
- Findings would make a case for **stressing material benefits** and **efficiency wages**.
- Introducing incentive schemes for tax collectors has a **bribe effect in two directions**: Some bribes will increase because of increased cost of non-(or under-)reporting but others will go to zero as actual reporting is now better for the taxpayer than paying the increased bribe.
 - Potential policy solution: have **performance pay** in general. In times when true information is important, however, accompany them by **external audits** or other verification mechanisms.

6 Taxation and Development

6.1 Papers and General Introduction

(Main) papers discussed:

- i. Pomeranz 2013. *No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax*. NBER Working Paper.
 - Assesses the role of third party information for tax enforcement by looking at the differential effect of audit announcements among transactions with and without paper trail
- ii. Best, Brockmeyer, Kleven, Spinnewijn, Waseem 2014. *Production vs. Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan*. Forthcoming: JPE.
 - Evaluate the impact of a Minimum Tax Scheme in Pakistan on tax evasion
- iii. Gordon & Li 2009. *Tax Structures in Developing Countries: Many puzzles and a Possible Explanation*. JPE.
 - Develop a theoretical model to explain tax policy of developing countries

General Theme:

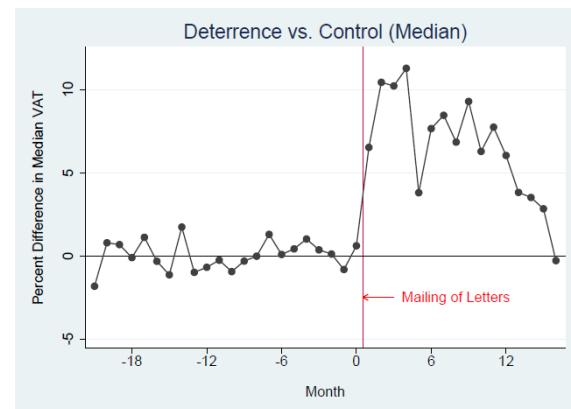
- Tax revenue is not only influenced by the **effort of tax collectors** but also by the **decision of the taxpayer** (to register on the extensive margin and to comply or evade on the intensive margin)
- There is a **stark differences** between developing and developed countries
 - in **tax revenue to GDP ratio**, and

- in the **share of “modern” taxes** (personal income tax, VAT, profit tax) used as opposed to inflation, capital tax, tariffs, revenue tax, etc.
- For **efficiency** reasons, tax literature suggests that “modern” taxes should be used. However, it usually **assumes perfect tax enforcement** and **availability of all tax instruments**.
- In light of these **assumptions not holding** in developing countries, the papers assess what sort of taxes could in fact be **optimal** in such countries:
 - P 2013: VAT has a **self-enforcement** mechanism that is effective in preventing evasion
 - Best et al. 2013: **Minimum tax schemes** or revenue tax may be welfare increasing when one faces the trade-off between production and revenue efficiency
 - GL 2009: If firms can evade taxes by not using the financial sectors, all sorts of **inefficient taxes** such as the above might in fact be **optimal**.

6.2 Main Ideas and Results

6.2.1 Pomeranz 2013

- Using **randomised letters** sent by the Chilean tax authority (deterrence letter announcing check for irregularities, tax morale letter affecting social norms and placebo letter just informing about website changes as well as a letter announcing audit), and tax data from the Chilean tax authority, she investigates:
 - Whether **deterrence** has an effect on reported VAT: Yes, the percentage of **VAT payments larger** than previous year payments, predicted payments, and zero (extensive margin effect) all increase significantly
 - Whether this effect is different for transactions with and without **paper trail**: Yes, relative to the previous year, **final sales** reported **increase significantly** while **input costs and intermediary sales reported do not change**
 - Interpretation/ explanation: firms naturally report input cost truly because it decreases their tax burden. They also seem to report intermediary sales truly because there exists a **paper trail** as the buying firm needs a receipt to declare the purchase itself as input cost. For final sales, however, none exists because final customers rarely need a receipt
 - Whether **audit announcements** have **spillover effects** in VAT payments along the value chain: Yes, it **increases VAT payments** relative to the previous year and predicted payments for the receiving **firm itself** and **suppliers** significantly, but there is **no significant change for clients**.

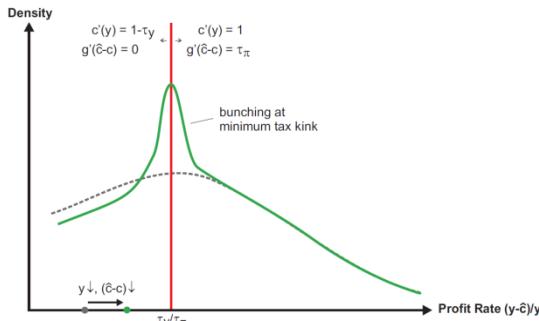


6.2.2 Best et al. 2014

- Use **minimum tax scheme** in Pakistan (you pay either revenue or profit tax depending on which liability is larger) to estimate tax evasion – excess bunching at the kink point

τ_y/τ_π (revenue tax rate over profit tax rate), to the left of which you would pay revenue tax and to the right of which you would pay profit tax, is probably due to tax evasion rather than real production decision.

- Basic **trade-off**: profit taxes ensure **production efficiency** while they make tax evasion much easier because profit is harder to observe than revenue. Revenue taxes ensure **revenue efficiency** while they distort production decisions.
- The **bunching response** depends on the **real output elasticity** (wrt. the marginal tax rate) and the **evasion response**. For such a large bunching response, the real output elasticity must be extremely (and completely incredibly) large around the kink.
- It can very neatly be seen on graphs how changes in the kink rate in some years induce a change in the bunching around it.
- Theory: When a minimum tax scheme is introduced, **movements towards the kink** from *both* sides are optimal:
 - From the right: reduction in declared profits reduce tax paid up until the kink.
 - From the left: initially, you paid profit tax. Now, your marginal return to real output decreases from 1 to $1 - \tau_y$ and your marginal return from evasion decreases from τ_π to 0. As a response, you will decrease output and increase profit which increases the profit rate.
- Under some assumptions, the **evasion reduction** by changing to a revenue tax is estimated to be **60-70%**
- **Corporate tax revenues increase by 74%** without decreased aggregate after-tax profits
 - Production efficiency losses are more than compensated by revenue efficiency gains



6.2.3 Gordon & Li 2009

- Observed **tax policy** in many developing countries is **very inconsistent** with optimal tax theory (latter forecasts e.g. no tariffs, no capital tax, uniform tax on consumption, preserve production efficiency).
- Key assumption in their setting: firms can avoid taxes by using **cash transactions** instead of the financial sector and thus leaving no paper trail.
- They will only use the financial sector if the value they receive from so doing (a_j) exceeds the tax (s_j) they have to pay when so doing:

$$\max_{\beta_j, L_j, K_j} \left((1 - \beta_j)p_j^*f_j + \beta_j p_j^* \frac{(1 + a_j)}{1 + s_j} f_j - rK_j - wL_j \right)$$

where p_j^* is the world market price, s_j the (revenue) tax rate, a_j the gains from using the financial sector as a mark-up on output in industry j (arguably heterogeneous across industries) and $\beta_j \in \{0,1\}$ essentially the decision variable whether to use it ($\beta_j = 1$) or not ($\beta_j = 0$).

- Given this setting, the following **tax policies** may be optimal:
 - Taxes that **favour labour-intensive** as opposed to capital-intensive sectors as the latter tend to benefit more from the financial sector. This could include capital tax, no or little personal income tax, differential excise tax according to sector
 - **Tariffs** to compensate tax differences across tradable industries
 - **Inflation** to indirectly tax the informal cash economy
 - **Restricting foreign entry** of firms which would ease tax evasion for domestic firms using foreign banks and partners
 - High level of **red tape** to indirectly tax the informal cash economy
- Generally, **restricted foreign entry** and the **threat of financial disintermediation** by domestic firms should lead to **low tax revenues relative to GDP**.
- All this can be observed in many developing countries and is certainly **inconsistent with optimal tax theory** (which however assumes tax enforcement is perfect) and also **hinders growth**.

6.3 Empirical Strategy and Discussion

6.3.1 Pomeranz 2013

- For the first and third question, (1) is used, and for the second question (2) is used:

$$(1) Q_\tau(VAT_{it} | Z_i, t) = \alpha_\tau + \beta_\tau(Z_i \times post) + \gamma_\tau Z_i + \delta_t$$

$$(2) LineItemInc_{itl} = \alpha + \beta(c_i \times post) + c_i + \delta_t + \phi(Z_i \times post \times l) + \eta(post \times l) + \iota_l + e_{itl}$$

where Q denotes quantiles (zero, median, previous year value, predicted values using control group outcomes, firm's pre-treatment VAT payments, size and final sales share), Z is the treatment variable, l denotes the line item and β and ϕ are the relevant DID coefficients
- Concerns:
 - **Interpretation:** first finding that intermediary sales are unaffected does not necessarily support a self-enforcement effect of a paper trail. Firms could simply not change anything here because they cannot hide anything anymore once they are audited (because of the paper trail) and want to signal good intentions on accurately reported final sales. However, this does not mean that they reported intermediary sales truly in the first place.
 - The fact that **intermediary sales** are non-significant for the deterrence letter but **supplier effects** are significant in the spillover effects analysis is somewhat **inconsistent**. If the interpretation is that intermediary sales are reported correctly (which are the sort of sales suppliers of a firm engage in), why does the spillover analysis suggest they change their behaviour?
 - **Sample selection** of firms which have both intermediary and final sales.
- Strengths:
 - **Randomisation** removes essential endogeneity.
 - **Testing for both suppliers and clients** (and, in a robustness check, controlling for differential characteristics) with the respective results in the **spillover investigation** very clearly identifies that the effects have to do with the VAT chain

6.3.2 Best et al. 2014

- For the graphical analysis, they just plot the density of reported profit rates.
- Evasion is estimated by comparing the observed density to a counterfactual density which is computed by fitting a flexible polynomial to the observed density excluding observations in the range around the kink that is (visibly) affected by bunching:

$$d_j = \sum_{i=0}^q \beta_i (\pi_j)^i + \sum_{i=\pi_L}^{\pi_U} \gamma_i \times 1[\pi_j = i] + v_j$$

where $[\pi_L, \pi_U]$ is the excluded range. Excess bunching can be calculated by subtracting the observed d_j from the estimated \hat{d}_j over this range.

- Concerns:
 - Some **strong assumptions** have to be made to estimate **specific numbers** (GE effects excluded, specific output elasticity assumed, functional form assumptions for production and evasion cost function). So, specific numbers can arguably not be taken too seriously here. However, it is a nice graphical analysis
- Strengths:
 - Very obvious tax evasion going on seen in **graphical analysis**.

6.3.3 Gordon & Li 2009

n.a. (theory paper)

6.4 Policy Conclusions

- In developing countries where **tax enforcement is not perfect**, deviations from the theoretically optimal tax policy might in fact be *optimal*
- In particular this regards:
 - **Relatively higher tax on capital** than wage
 - **Inflation tax**
 - Using taxes and mechanisms that have **some inherent enforcement mechanism** even if this compromises on tax efficiency (since efficient taxes might be evaded on if they cannot be enforced)
- A **minimum tax scheme** or payment of revenue tax may be a welfare improvement when the trade-off between revenue efficiency and production efficiency is faced
- **VAT** has an inherent self-enforcement mechanism through the paper trail it leaves which turns out to affect evasion behaviour substantially

7 Firms, Markets and Entrepreneurship

7.1 Papers and General Introduction

(Main) papers discussed:

- i. **Atkin, Faber & Gonzales-Navarro 2015. Retail Globalisation and Household Welfare: Evidence from Mexico. Working Paper**
 - Assess the effect of entry of global foreign retail (Walmart) in Mexico on consumer welfare and competition in the retail market
- ii. Hsieh & Klenow 2014. *The Life Cycle of Plants in India and Mexico*. QJE

- Examine the different life-cycle properties of firms in India and Mexico compared to the US and assess their impact on aggregate manufacturing productivity
- iii. Bandiera et al. 2017. *Labor Markets and Poverty in Village Economies (earlier draft: Can Basic Entrepreneurship Transform the Economic Lives of the Poor?)* QJE.
 - Large scale RCT to examine how a transfer of capital and skills impacts the occupational choice for the poorest of the poor

General Theme:

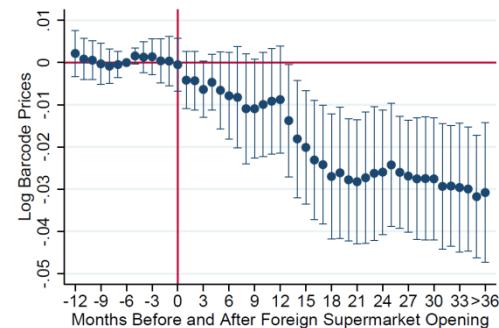
- How does the structure of the market affect developmental outcomes? Alternate between showing how entry + liberalization can increase welfare, but also that the ultra-poor face binding constraints: S-shaped development, poverty traps, etc.
- **What are the constraints that trap individuals or firms in low-productivity states?**
 - HK 2014 suggest there are a number of size-dependent barriers that prevent firms from growing in size and productivity as they age in India and Mexico
 - taxes more strictly enforced on, and greater adjustment costs for larger (registered) establishment
 - credit constraints which prevent firms from investing in firm-specific intangible capital (management) and also causes an inefficient allocation of capital across firms
- high trade and transport costs make it more difficult for firms to grow their market
- Bandiera et al. 2013 shows that the ultra-poor face capital / skills constraints that result in underemployment. Why so many poor trapped in low skilled occupations?
 - **Low skills**
 - **Limited capital, credit constraint**
 - **Limited opportunities on demand side**
- **What is the effect of global foreign retail entry on development?** Politically salient topic: Walmart not yet allowed in India. Retail is also a little studied but 15-20% of total employment, 10-15% GDP and **50% of household expenditures** in many developing countries – lowering consumer prices could have massive effect.
- **The idea of management as a proximate cause:**
 - Even within **very narrow sectors**, there are (large) **differences in TFP**. This may come from management differences.
 - Management can be seen as
 - **Technology:** $\frac{\partial F}{\partial M} > 0 \quad \forall M$
 - **Design:** $\exists \bar{M}(r, w) : \frac{\partial F}{\partial \bar{M}} = 0$ where r and w are the prices of capital and labour
 - **Difficulties** in research on the impact of management
 - **Data:** Mostly in anecdotes, quantitative data difficult to get
 - **Endogeneity:** For low wages, management would be substituted by labour

7.2 Main Ideas and Results

7.2.1 Atkin et al. 2015

- They investigate three questions:

- What is the effect of **FDI** on average **household welfare** in municipalities of entry? **+7.5% of average household income**
- What are the **underlying channels**? Direct effect, pro-competitive effect
- To what extent do gains **differ** across the **pre-existing household income** distribution? Regressive welfare gains: effect is twice as large for richest as compared to poorest income group.
- Walmart entry in Mexico analysed. Generally, huge growth of retail sector in Mexico in last since 2000: 365 stores in 2001 to 1335 stores in 2014
- They decompose the welfare effects into two main effects. By looking at the compensating variation, they separate these two effects
 - **Cost of living effect:** fall in consumer prices (holding income constant). Change in P for a given bundle => this is why **barcode data** is essential: **allow for changes in composition** & quality; also provides more **reliable** measure of prices and can distinguish prices charged by Walmart and those by tradition retailers.
 - **Direct effect:** foreign retailers charge lower prices and offer new products (accounts for 80% of COL reduction)
 - **Pro-competitive** has an **intensive** (see figure) and an **extensive margin** effect. Incumbents must lower their prices or exit (20% – reflects smaller market share of incumbents?)
 - **NB:** while intensive margin pro-competitive effect is observable, extensive margin pro-competitive effect and direct effect from more varieties and amenities are not. Make functional forms assumptions to estimate demand at virtual price: CES
 - **Income effects:** effect on household nominal income through changed employment (esp. in the retail sector), etc. Use household surveys and construct quarterly time-series of incomes, wages, occupation, etc.
 - Retail business profits for domestic retailers. Fall in profit or exit => large adverse effect
 - Retail labour income: small positive
 - Effect of retail entry on incomes from other sectors: small positive
- **Non-pecuniary welfare effects:**
 - customers have to use their car to get to large store => **pollution** externalities
 - **changes in product composition** / cheaper junk food => obesity in Mexico (although requires another distortion: time inconsistency, lack of information, etc.)
- Given strong positive effects, **why might voters in a country be against entry of Walmart**: negative effect on traditional retailer profits: so **if the median voter is a small retailer** or employed in traditional retail (likely if large chunk of population engaged in this sector) could explain reluctance. NB: because of scale effects, unlikely that all staff from exiting small retail will find jobs at Walmart.



7.2.2 Bandiera et al. 2013

- Conduct a large **RCT** to assess whether ***in situ* occupational choice** can be engineered when the **asset and skills constraints are relaxed simultaneously**. Question: why don't poor women specialize in wage labour or self-employment? Do they have **different preferences** from the slightly wealthier women (norms, etc.) or are they subject to **binding asset/skill constraints**? Test: will the poor choose to liquidate or work with the assets provided?
- Also look at **spillover** effects in the community.
- **Background:** Bangladesh – landless labourers facing irregular employment in agriculture, also work as domestic servants (mostly women) – the poorest of the poor: no skills, no capital, very vulnerable
- **Program design** – collaborate with BRAC
 - Survey: large labour survey: cross-section 23,026 households from 1409 communities. 4 waves: baseline, T + 2yrs, T+4yrs, T+6yrs (last one treated only). All beneficiaries surveyed, plus sample of households from other wealth classes => allow for benchmarking and spillovers effects. Attrition over the 4 years is 13% both in T and C communities.
 - Eligibility: 6,000 beneficiaries are selected by same procedure in both treatment and control communities (some poverty line + identified by the communities as their poorest members)
 - Random allocation of basic entrepreneurship: select from menu: livestock, small crafts, small retail (=> almost all choose livestock (cow, goats)) + receive asset-specific training over 1 year. Randomisation at the branch rather than community level to minimise contamination
 - **Asset transfer** amounting to **USD 560 PPP** (massive!)
- **Channels:** income effect would push consumption, substitution effect increase work hours
- **Results: Generate sustainable change in occupation towards self-employment although still perform wage-labour. Large increase in work on the intensive margin. Greater earnings.**
 - Extensive margin & occupation choice: After 4 yrs, share of women in wage labour falls by 17%. **Share of women in self-employment (husbandry) increases by 15%** and those in both occupations increases by 8%. Casual wage labour is characterised in this setting by uncertain and low earnings, for unskilled labour only and spot contracts.
 - Intensive margin: Net annual **labour supply increases by 218 hours** (+22%)(decrease of hrs in wage labour more than compensated by increase in self-employment hours) => income seasonality reduced
 - 15% in labour productivity, **37% increase in earnings** (8% increase in expenditure) – lifts 11% of women above poverty line. Results more pronounced after 4 years than 2 => **sustainable** path out of poverty.
 - Even sustainable gains after 7 years.
 - Both mean and dispersion of the treated are increased; the ultra-poor overtake the non-eligible very poor, and close 40% of the gap with middle households.
- **Interpretation:** the poor do face underemployment constraints that can be alleviated by skill/capital transfer. No crowding out of other village activities.

- Alternative explanation: was the constraint really binding, or is there some **learning** about opportunities going on too? Like in Bryan et al Seasonal Migration paper.
- **Three factors** are **critical** in this programme and likely decisive for the positive impacts observed
 - **Capital and skill** were provided together. Other literature suggests that they are **complementary**
 - The sheer **magnitude** of capital and skill transfers provided is huge. This sets the programme apart from standard microfinance programmes which tend to be not so successful.
 - **Outside options** for women under study were rubbish: insecure wage labour. Remember: ITT on eligible, selected very vulnerable women.
- **Cost-benefit analysis:**
 - Costs TK 20,700 per household (including value of the asset + training) and yields TK 1,737 per year. Would need to compare with return on cash transfer of same amount (zero if spent, more if invested in individual “best activity”).
 - **Benefit-cost-ratio: 3.21, IRR: 0.22**
 - Compare programme to **counterfactual cash transfer** and assume people would save the money and profit from interest. Then programme makes them better off
 - However, they could also use the money to make an **investment of their choice**; they argue the menu from which they could choose in the programme is quite long.

7.2.3 Hsieh & Klenow 2014

- Document different patterns in **life-cycle of firms in India, Mexico and US** Relate this to aggregate manufacturing productivity growth. Stylised facts: in US, 40-yr old firms are 7 times larger than young plants, whereas in India they are only 40% larger.
- The patterns are consistent with firms in Mexico and India having more **difficulty to accumulate of firm-specific intangible capital** (managerial or organisational investments that are necessary for the firm to grow beyond a certain size). Reasons for why firms in developing countries invest less in this type of capital? See part 7.1.

7.3 Empirical Strategy and Discussion

7.3.1 Atkin et al. 2015

$$\ln p_{gsbmt} = \sum_{\tau=-12}^{36} \beta_\tau \mathbb{1}(MonthsSinceEntry_{mt} = |\tau|) + \delta_{gsbm} + \eta_t + \epsilon_{gsbmt}$$

where $\ln p_{gsbmt}$ is the log price of a barcode-product b in product-group g , individual store s , in municipality m and month t . $I(MonthsSinceEntry_{mt} = \tau)$ is an indicator function, and $MonthsSinceEntry_{mt}$ counts the months since foreign entry for each municipality m at a given point in time t (with negative values counting months before entry, positive values counting the months after entry, so that $MonthsSinceEntry_{mt} = 0$ in the month that a foreign store enters a municipality for the first

- Income effect similarly estimated.
- Strengths:
 - **Barcode data!**

- Look at **long-run** effects. 3 months after entry and even beyond (>36 months is last data point)
- Concerns:
 - Challenge is **endogeneity in foreign stores entry**. Deal with this by event study design. Walmart eventually wanted to go into every municipality but went into different ones at different points in time depending on when they got the permit. If this is only decisive for when they enter, there shouldn't be a pre-entry price trend. Tested by looking at single months (non-parametrically) before and after entry. Works very neatly (see figure).
 - **Functional form assumptions**

7.3.2 Bandiera et al. 2013

- Basic regression:

$$y_{idt} = \alpha + \beta_1 W_1 T_{id} + \beta_2 W_2 T_{id} + \gamma T_i + \delta_1 W_1 + \delta_2 W_2 + \eta_d + \epsilon_{idt}$$

District d, individual i. Time t=0,1,2: baseline, midline T+2, endline T+4. W are survey waves.

β_t measures impact of treatment (more precisely effect of "intent to treat". 87% of selected beneficiaries opt in)

- **Randomisation time of roll-out** at BRAC branch level (40 BRAC branches in 1409 communities): 20 get treatment in 2007, 20 in 2011 (control)
- Communities and eligible households within a community are *selected* according to the same procedure for treatment and control branches
 - As it is the same procedure, the randomisation of timing provides causal effect on *eligible* households (ITT; because of selection, these households are not representative households of society or even within their communities)
- No differential attrition between control and treatment wrt. baseline occupational choice
- **Two important assumptions:**
 - **Common trend**
 - **No contamination** between control and treatment
- Strengths:
 - large **RCT**
 - **long-run** follow-up surveys to check persistency
- Concerns:
 - Perhaps only **external validity**: there are some places perhaps where differences in preferences / norms heavily shape occupational choice independently of constraints: e.g caste system
 - Also would be good to know effect later than T+4 and see if spillovers

7.3.3 Hsieh & Klenow 2014

- If Robin asks about this, I'll eat my hat.

7.4 Policy Conclusions

- **Entry of global retailers** into local markets have **positive** and **negative** effects. While the debate is often focused on negative effects for local incumbent retailers, Atkin et al.

2015 show that **overall income gains are positive** (coming from a vast majority benefitting from higher competition and lower prices as well as more variety)

- It is hard to understand why some really profitable investments are not undertaken by the very poor and why they do not seem to engage in profitable self-employment or use microloans to do so. Bandiera et al. 2013 actually demonstrate a successful programme in Bangladesh. The key differences to existing mostly unsuccessful analysed programmes are:
 - **Capital and skills are both provided** – they seem to be **complementary**
 - **Magnitude** of capital and skill transfer provided is very large
- While rural-urban gaps and shifts from agriculture in (urban) manufacturing are often seen as important occupational changes, **in situ occupational changes** from insecure wage labour to more secure and profitable self-employment are important upgrades on the occupational ladder too.

7.5 Further papers

De Mel et al. (2008). Returns to Capital in Microenterprises: Evidence from a Field Experiment. QJE.

- DMW **randomly give grants to Sri Lankan microenterprises** (capital other than land and buildings < 1000 USD)
- **Four treatments: USD100 cash, USD100 in-kind, USD200 cash, USD200 in-kind.** These are large treatments: 3 (6 for USD200 treatment) months of median profits or 55% (110%) of median initial invested capital. In-kind equipment was selected by the business owner
 - First stage impact on capital stock is (pointwise) a lot higher for cash treatments than in-kind treatments, but the difference is not statistically significant.
- They estimate treatment effects, that is, how the capital injection changed profits and also – more structurally – estimate the marginal return to capital directly using an IV approach (s. below)
- DMW find that both treatment effects (per unit of capital) and the **return to capital estimates range between 4.6-5.3% per month** (55-63% pa.; note that the paper's abstract has a typo here). This is substantially higher than market interest rates. Such high findings at very low capital stock **speak against the idea of poverty traps** because of non-convex production sets, ie. very low returns at low capital (however, selection of those who are actually in business observed)
 - this effect is even larger than what some non-experimental studies in the area give. They believe that therefore, attenuation bias through measurement error play a major role in non-experimental studies.
- **heterogeneity:** returns strongly **vary with entrepreneurial ability and wealth** (60% of women and 20% of men have returns < market interest rate of 12-18% pa. in spite of substantial findings), **but not with measures of risk aversion or uncertainty**; returns are much **higher for men than for women** (not significantly different from zero for the latter)
- **Results more consistent with missing credit markets than missing insurance markets:** Missing credit markets and perfect insurance markets mean higher marginal returns for firms with higher entrepreneurial ability, and lower for households with more workers and liquid assets, but risk shouldn't make a difference. Perfect credit markets

and missing insurance markets mean that the difference between marginal return and interest rate should be explained by the level of business risk, but they don't find significant heterogeneity along the lines of risk.

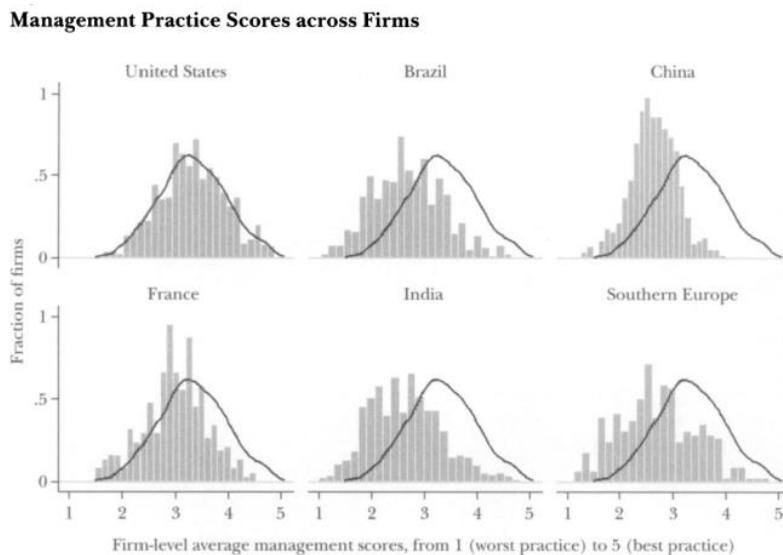
- In a 2009 follow-up paper on the gender question, they look at potential channels for gender differences. They consider household wealth, liquidity, entrepreneurial ability and risk attitude: even including controls for them, the differences remain.
- They also **directly estimate returns to capital** in a specification where **profit is regressed on capital and time and individual fixed effects**. They use both linear and log measures for both profits and capital. The **treatment is used as an IV** for capital. In the absence of heterogeneous effects, the IV estimate is a measure of the returns to capital. **In the presence of heterogeneous effect, it's a LATE:** weighted average of marginal returns where the weights are how much the capital stock responds to treatment – if the degree of the response and the marginal return are positively correlated, ie. higher return entrepreneurs invest more of the treatment in their business, then the LATE is larger than the average treatment effect.
 - However, entrepreneurs equate their marginal returns from business activity and household activity (their opportunity cost of investing cash in the business as opposed to having it in the household). So, higher marginal return entrepreneurs also have higher marginal returns in the household and would only invest more in the business if their household returns drop at a *higher rate* than the business returns once they have more cash. The authors test this using some measures of ability, risk aversion, etc. and don't find that this holds.
- QNS:
 - They use a direct question on profits and the difference between sales and cost as measures for profit. The latter is substantially lower and the correlation between the two is only .33. Do the results hold up for both profit measures?
 - Is Digit Span a good measure for entrepreneurial ability?
 - How do they get their return on capital measures? (Section V, Table IV):
 - For real profits (1), they claim 5.85% return. In a lin-lin specification, the coeff. 5.85 means a 1 LKR increase in capital leads to a 5.85 LKR increase in profits, ie. 485% return. In a lin-log specification, the coeff. 5.85 means a 1% increase in K leads to a 0.0585 LKR increase in profits. Or a 1 LKR increase in K ($=1/K \times 100\%$ increase in K) leads to a $5.85/K$ LKR increase in profits. Given the mean K, this is tiny. They probably measure the profits in 1/100s to get the direct return on capital as a number. Then a 5.85 unit increase in 1/100 profits means a 0.0585 increase in profits which corresponds to a 5.85% return (note that their LHS variable is *profits*, not *sales*, otw. A coefficient of 1.0585 would give a 5.85% return)
 - For log real profits (2), they claim a 5.51% return based on mean $K = 26500$ LKR, mean profit = 3850 LKR. This should be due to a log-log specification. In a log-log specification, the coefficient is the elasticity, that is, a 1% change in K increases profit by 0.379%. In units at the mean, this translates in the following way: A 1 LKR change in mean K (that is, $\frac{100}{K}\%$) translates into a $0.00379 \bar{Y} * \frac{100}{K} = 0.379 * \frac{\bar{Y}}{K} = 0.379 * \frac{3850}{26500} = 0.0551$ LKR change in profits, corresponding to a 5.51% return.

McKenzie and Woodruff (2008). *Symposium on Access to Finance: Experimental Evidence on Returns to Capital and Access to Finance in Mexico*. World Bank Economic Review

- They randomise cash and in-kind grants for micro-entrepreneurs (less than 1000 USD capital stock) in Mexico and find returns as high as 20-33% per month
 - Two treatments: Cash grant and in-kind grant, both worth 140 USD.
 - Panel of 207 microenterprises in the city of Leon over five quarters
 - This is larger (at the upper end of) than a previous non-experimental study found
- In interaction with firms reporting their degree of financial constraints, they find that the return is much higher (70-79%) for firms reporting to be financially constrained.
- They estimate both ITT effects and IV-LATE effects. For both, they also run RE and FE specifications to take potential firm-specific elements in the error term into account (there are multiple observations for one firm) and to increase precision.
 - Note that using a 2SLS procedure, the estimator is essentially ITT / uptake rate, that is the TOT. Their returns estimate is based on the IV effect.
 - They don't seem to report how many people did not comply. (?) The difference between ITT and TOT is rather small, but not zero (contrary to DML where both the pure treatment estimation and the IV give very similar results)
- Attrition is non-negligible and smaller in treatment firms. It seems like some selection is going on in that firms that were assigned to treatment had to stay in the survey and be able to be located
- They pool both treatments, but test for different effects and cannot reject. However, huge volatility in results anyway.

Bloom and van Reenen (2010). *Why do Management Practices Differ across Firms and Countries?* J Econ Persp

- While productivity differences on the country level are studied, productivity differences on the firm level that may be very sizeable were less studied. Economists did not bother too much about management because, firstly, most studies have been done rather as case studies than systematically across firms and countries and, secondly, it seems that if management mattered, it should be pretty doable to employ good managers and overcome the transaction and agency costs
- Based on **telephone interview**, Bloom & Van Reenen (2010) score firms from many countries on 18 basic **management practices** along three main lines:
 - **Monitoring**: know what happens and using this information
 - **Targets**: setting targets, taking action
 - **Incentives**: rewards for good performance, consequences for bad performance
- Average overall score is highest in the US and China and Brazil are considerably **left shifted** from the US distribution (explanations for TFP differences?)



- Management score is significantly positively **correlated with sales, number of employees and profitability**
- Difference in management scores across countries is mostly due to the **size of the left tail** of badly managed firms
- **Strong product market competition** seems to boost average management practices
- **Multinationals are generally managed well** in every country while **family run firms** and **government-run firms** are generally quite **badly managed**
- Firms with **intensive use of human capital** (more educated workers) tend to have **better average management** practices
- Countries with **lighter labour market regulations** tend to have **better average management** practices
- Because it's only macro correlations, no causal inference is really possible. There may be **two explanations**:
 - **Management does matter** and developing countries need better management
 - This is due to **misallocation** driving both business outcomes and management
- It turns out that the effects are quite sizeable. A combination of (i) **imperfectly competitive markets**, (ii) **family ownership of firms**, (iii) **regulations restricting management practices** and (iv) **informational barriers** (both pure information about the own need/ lack of good management and slow 'technology' adoption) seem to allow bad management to persist.
- There are essentially **two issues** with the finding of a huge **left tail** if management matters: **(a) why do these firms not adopt good management practices?, (b) given they don't, why don't they get pushed out of the market?** What are the **market failure** reasons behind these?
 - (iv) could be a market failure as explanation and (ii) another explanation for (a).
 - (ii) could be a reason or catalyst for (iv)
 - (i) could be a market failure and (iii) a policy reason explaining (b)
 - Note that (ii) may be a rational decision. If rule of law is uncertain and outside managers cannot be trusted, the negative effect from productivity may be offset by the positive effect of trust/ less uncertainty around conflicts of interest.
- If management makes the difference (and further micro evidence suggest so, s. below), **policy conclusions** could be:

- **Foster market competition** which leads to better managed firms gaining shares and prospering and badly managed firms being forced to exit
- **Get rid of taxes and distortive policies that favour family run businesses** which tend to be managed worse (this is because of misallocation of talent: not talent but family members are managers)

Bloom et al. (2012) explore how **social capital**, proxied by trust, impacts on **management/organisation of firms** and thus on productivity

- Basic story: CEOs hire n managers to solve problems. They share of problems they solve is z_M and trust λ determines how many they solve correctly, namely λz_M . Management skill α is the cost of acquiring information and determines the cost αz_M to train managers to solve z_M . Trust and training cost therefore determine the optimal level of decentralisation z_M^* and the optimal firm size n^* through:
- **Trust is like a reverse cost:** When it is high, you can rely on public officials, don't have to spend resources protecting your property, capital markets work better as borrowing and lending is easier, etc.
- **Decentralisation, firm size and productivity increase** in higher trust - theoretically and found empirically across 4,000 multinational firms with HQs in Europe, Asia and the US
- Relation to **developing countries**: there may be **reverse causality** issues in determining whether poverty induces low trust. You can argue that poor countries have weaker institutions which may induce low trust.

Bloom et al. (2013). Does Management Matter? Evidence from India. QJE

- In order to identify more causal influence of management practices on productivity, Bloom et al. (2013) randomise **consulting treatment** for 11 larger (100-1000 employees) **textile firms in India** with a control group of 6 larger textile firms
 - small sample size! But they do some technical stuff about it: they use some statistical procedure and they also use asymptotics in T rather than N since they have a huge amount of within-firm across-time/weekly data.
 - Essentially the idea is similar to taking within-firm averages over time. An average within a firm would give you a very precise estimate. So, you don't need that many different firms any more.
 - Treatment: 4 months free high quality consulting
- They estimate ITT effects
- There are **three phases**:
 - **Diagnostic phase:** detailed recording of management practices prior to the intervention and presentation of results
 - **Implementation phase:** implementation of consulting services over four months
 - **Measurement phase:** collection of performance and management data and light consulting advice, requiring 1.5 days per month
- There are **four groups**:

- **Treatment plants:** 14 plants among the 11 treatment firms that were subject to all phases
- **Control plants:** 6 plants among 6 control firms (1 each) that were subject to the diagnostic and measurement phase
- **(Treatment and Control) Non-experimental plants:** Plants among the treatment and control firms where only measures on management score were taken on a bi-monthly basis (not the fully fledged programme of any of the phases)
- The ‘first stage’ shows that the consulting did seem to improve management practices, but also demonstrates an effect of the diagnostic phase. Note that this means – if anything – that the treatment effect is a lower bound:

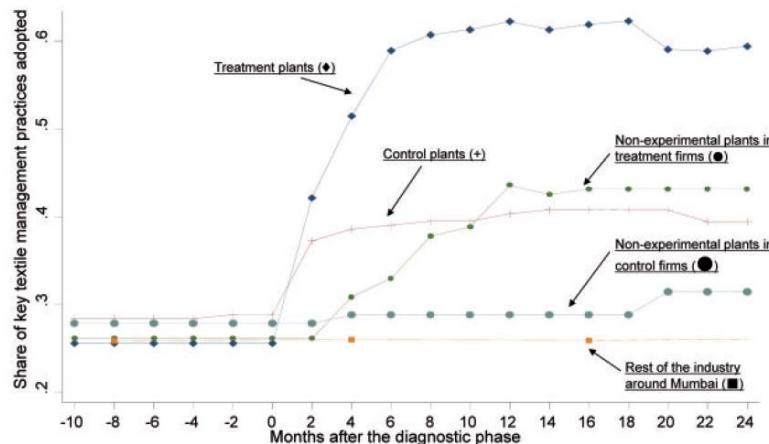


FIGURE V
The Adoption of Key Textile Management Practices over Time

- The impacts on **quality defects, inventory** and **TFPR** are significant (**50% decrease, 40% decrease, increase**) and they do not find spillover effects
- Costs of USD 250,000 increase profit by USD 325,000 → **profitable**
- **Productivity increased by 17%** in treatment relative to control firms
- **Reasons for non-adoption** seem to lie in **information barriers** and the **choice of male family members only**. Little **competitive pressure** through high tariffs and **lack of external finance** may have been reasons why badly managed firms were not pushed out of the market
 - s. also above discussion on the Bloom-vanReenen-2010 paper what underlying reasons for non-adoption w/o consequences are
- 17/66 firms complied with the offered consulting services, but the 17 firms were not significantly different from the other firms in some observables, inter alia the management score, employee numbers, total assets etc.
 - Reason for international consultancy: local consultants who see the firms' accounts could bribe them afterwards b/c many firms are evading taxes in this setting. This could also (still) be a reason why take-up is so low
- The study and results contribute to a number of aspects / different literatures:
 - The role of management as discussed above
 - Impact of training on business performance (note that this is a very expensive and comprehensive consulting service provided to big plants; so the results may not be directly comparable to the studies looking at short training for micro-entrepreneurs)

- Management literature debate between 'best practices view' and 'contingency view': there seems to be a certain set of management practices that does make a difference for all firms and is not firm-specific; i.e. this is evidence for the former

De Mel et al. (2014). *Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka.* JDE

- De Mel et al. (2014) conducted an experiment with Sri Lankan **women** operating a **subsistence business** or thinking about starting one.
 - Treatment 1: **business training only** (ILO programme)
 - Treatment 2: **business training + cash**
- Finding:
 - Both treatment (mostly not significantly differently) **pos. sign.** affected all sorts of **management practices**
 - The impact on profit is insignificant for treatment 1 and significant for treatment 2, but only in the short-run (no more significance after two years) → **no impact of business training**
 - There is **some impact** of training on the women thinking of **starting a business: quicker entry induced and positive profit impacts**
- **Concerns:**
 - **Low power**
 - **Measurement** of key outcomes, they are self-reported.

8 Infrastructure and Trade

8.1 Papers and General Introduction

(Main) papers discussed:

- i. **Jensen 2007. *The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector.* QJE**
 - Impact of mobile phones adoption on market performance of fisheries sector in Kerala
- ii. **Donaldson 2012. *Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.* Forthcoming: AER.**
 - Assess impact of building India's railroad network on trade costs and quantity, price dispersion & incomes using colonial archives data and guided by a GE trade model
 - See also EC535 slides on this paper, quite comprehensive.
- iii. **Dinkelman 2011. *The Effects of Rural Electrification on Employment: New Evidence from South Africa.* AER**
 - Assess the effect of mass roll-out of electricity in rural South Africa on employment growth

General Theme:

- **Impact of improvements of infrastructure on trade, welfare and growth**
 - Infrastructure projects tend to **improve the efficiency of markets** in ways which are often **complementary**:
 - A reduction in **transport costs** is more useful, the more producers know where they need to go sell
 - Better **information** / reduction in **search costs** will have greater effect on trade if agents can take advantage of this by travelling
 - Improving effectiveness of output markets has effect on **consumer & producer surplus**
 - Reduction in **barriers to trade** (tariffs) often seen as key in development because of competition. Potential adverse effects: incentives for innovation of domestic firms ; infant industries
 - Facilitating **internal trade** through the reduction of transport costs can be just as important: (1) most trade occurs within countries, allow regions to read benefits from trade (2) **Better intra national trade spreads out the benefits of international trade** by allowing remote regions to have access to international markets⁴.
- **Impact of infrastructure is difficult to measure** because of **endogenous project placement** (correlated to unobserved changes in area's growth, income, etc.) well as **scale** and **non-replicability** of roll-out of new infrastructure. Attempts to address this:
 - Donaldson 2012: must **argue for exogeneity** of placement with respect to outcomes (placement responded to military rather than economic reasons; test using "placebo" railroad projects that were never completed)
 - Dinkelman 2011: **instrument** for electricity roll-out using land gradient which is unrelated to employment, and construct panel and estimate within variations in employment, **taking into account district fixed effects**
 - Jensen 2007 **very lucky**: exogeneity of timing of cell phone coverage across the fishing villages, clear identification

8.2 Main Ideas and Results

8.2.1 Jensen 2007

- Using **weekly survey data** from three regions in India in which mobile phones were sequentially introduced by a private mobile phone seller, Jensen looks at the effect of such information technology infrastructure on market performance and welfare, in particular:
 - **Location** of markets fishermen targeted
 - **Price dispersion**
 - **Waste** of supply
 - **Violations of the Law of One Price**

⁴ See Epifani & Gancia (2011) on complementarity of trade liberalization with internal policy reforms (including internal transport or industrial policy). The standard argument for trade liberalization is that it increases welfare through competition. However, it can also result in welfare reductions by increasing the *dispersion* in markups, when there is heterogeneity on who benefits the most from trade (because of intersectoral misallocation)

- **Profits and consumer surplus**
- Theory: In an imperfect information setting, the **First welfare theorem** and the **law of one price do not necessarily hold**. In particular, in the setting of local fishermen catching fish in their catching zone and deciding on the location of the market where they want to sell it, there is a BNE (Jensen shows), in which
 - There is a threshold $x(\tau)$, with τ being transportation cost, of fish caught above which fishermen sell at the **nonlocal market** (because if you catch much fish, it's likely that there is a lot of fish and your competitors catch a lot too which pushes prices down on the local market)
 - **Price dispersion** between the two markets you can sell at **exceeds transportation cost**
 - There is a threshold for transportation cost τ above which fishermen **always sell at their local market**
- Also, as soon as a search technology is introduced with which fishermen can learn about the catch for a certain cost ψ , there is a BNE in which
 - There is a threshold of fish caught $x(\psi)$ above which fishermen will **search and switch markets** when the other zone has a low fish realisation.
 - **A reduction in ψ reduces price dispersion**
- For a **downward sloping demand curve, better information increases welfare**. If you have a low realisation in one zone and a high realisation of fish in the other, fishermen from the high realisation zone will come to the market in the low realisation zone. This increases welfare in the low zone market (A+B) by more than it decreases welfare in the high zone market (E+F)

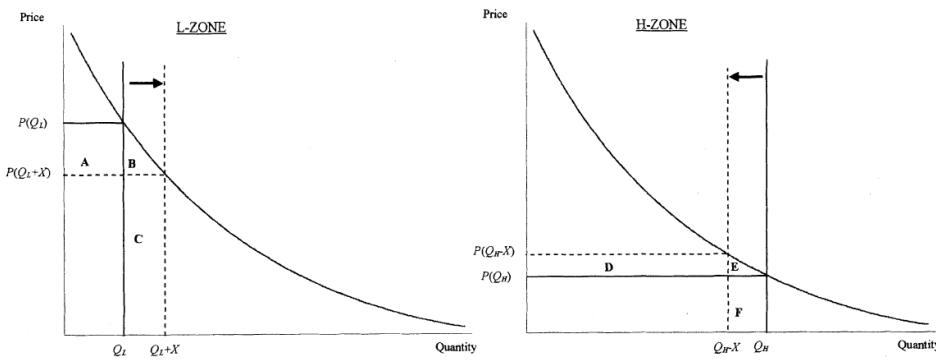
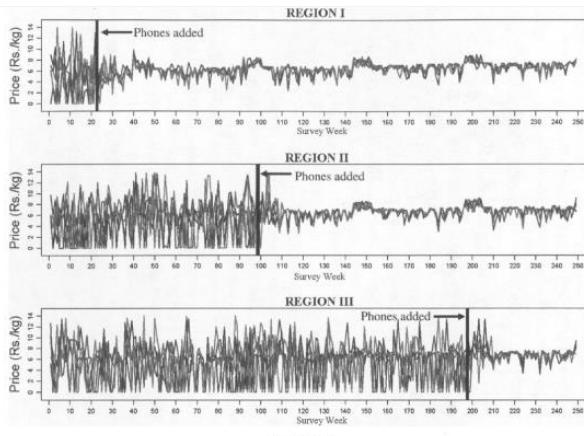
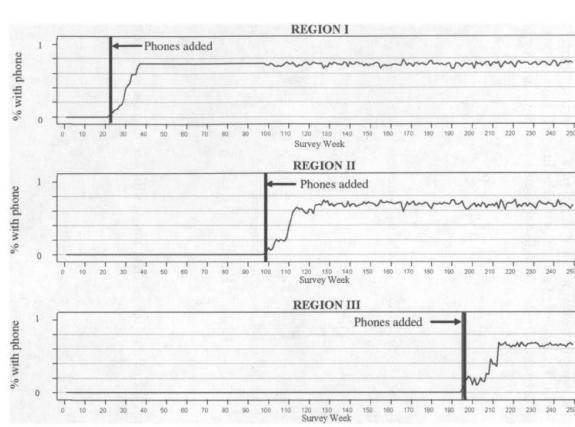


FIGURE I
Changes in Welfare Associated with Arbitrage

- The **first stage** is very convincing: People very quickly adopt mobile phones when they are offered [left]
- **Findings (extremely neat!):** The introduction of phones in the regions
 - **decreases the share of fishermen selling at local markets** on average from 100% to 58-70% (abruptly)
 - **decreases price dispersion** substantially and significantly (around a fifth) [right graph, Table III]
 - **decreases waste** from strictly positive levels to virtually zero [Table III]
 - **decreases the violations of the Law of One Price** from 54-60% to 3-8%. [Table VII]
 - **Increases both profits and consumer surplus** significantly.

TABLE III
PRICE DISPERSION AND WASTE IN KERALA SARDINE MARKETS

	Period 0 (pre-phone)	Period 1 (region I adds phones)	Period 2 (region II adds phones)	Period 3 (region III adds phones)
Max-min spread (Rs/kg)				
Region I	7.60 (0.50)	1.86 (0.22)	1.32 (0.10)	1.22 (0.44)
Region II	8.19 (0.44)	7.30 (0.29)	1.79 (0.19)	1.57 (0.16)
Region III	8.24 (0.47)	7.27 (0.27)	7.60 (0.25)	2.56 (0.34)
Coefficient of variation (percent)				
Region I	.68 (0.07)	.14 (0.01)	.08 (0.01)	.07 (0.01)
Region II	.62 (0.04)	.55 (0.04)	.12 (0.01)	.08 (0.01)
Region III	.69 (0.09)	.57 (0.04)	.54 (0.03)	.14 (0.02)
Waste (percent)				
Region I	0.08 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Region II	0.05 (0.01)	0.04 (0.01)	0.00 (0.00)	0.00 (0.00)
Region III	0.07 (0.01)	0.06 (0.01)	0.06 (0.01)	0.00 (0.00)

TABLE VII
VIOLATIONS OF THE LAW OF ONE PRICE

	Period 0 (pre-phone)	Period 1 (region I has phones)	Period 2 (region II has phones)	Period 3 (region III has phones)
Overall				
Region I		0.54	0.03	0.04
Region II		0.57	0.55	0.06
Region III		0.60	0.58	0.58
With time + depreciation				
Region I		0.50	0.01	0.02
Region II		0.53	0.52	0.03
Region III		0.57	0.55	0.54
All markets combined				
Without time + depreciation		0.47	0.35	0.20
With time + depreciation		0.44	0.31	0.16
				0.03

8.2.2 Donaldson 2012

- He comes up with a **trade model** based on Eaton-Kortum that delivers four key predictions Donaldson tests using data from the railway construction in India by the British (1861-1930)
- **(1) Inter-district price differences measure trade cost:** looking at the price for salt (which was only retrieved in some regions) in different regions he finds that the lowest-cost effective distance to the source (region) has a highly significant positive influence on the price. The roadway is about eight times (!) as expensive as the railroad, the river about four times
 - Without explicitly modelling a market for transportation, he just models transportation cost as a wedge between prices: $\ln p_d^o - \ln p_o^o = \ln T_{od}^o$ (origin o , destination d , for a good that is only produced in location o – the superscript that usually denotes commodity k)
- **(2) Bilateral trade flows take the “gravity form”, that is, they decrease in rising trade cost:** looking at expenditures on goods X , he finds that the effective distance between origin and destination along the lowest-cost route has a highly significant negative influence.
 - He needs the trade cost from (1) to estimate the gravity equation:

$$\ln X_{od}^k = \ln \lambda_k + \ln A_o^k - \theta_k \ln r_o - \theta_k \ln T_{od}^k + \theta_k \ln p_d^k + \ln X_d^k$$

where X_{od}^k is expenditure in region d on commodity k from region o , X_d^k is total expenditure on k in d , λ and A are parameters in the distribution that models the production process, r is the land rental rate as a cost factor, p is the price.

- Note that estimation of θ_k , the elasticity of expenditure with respect to transport cost is straightforward to estimate since all other regressors are some fixed effects.

Table 3: Railroads and Trade Flows (Step 2)

Dependent variable: Log value of exports	(1)	(2)
Log effective distance between origin and destination along lowest-cost route	-1.141 (0.203)***	-1.194 (0.446)***
(Log effective distance between origin and destination along lowest-cost route) x (Weight per unit value of commodity in 1880)	-0.052 (0.041)	
(Log effective distance between origin and destination along lowest-cost route) x (High-value railroad freight class of commodity in 1880)	0.035 (0.053)	
Observations	1,315,079	1,315,079
R-squared	0.949	0.955

Notes: Regressions estimating equation (14) using data on 17 commodities and 45 trade blocks annually from 1880 to 1920. Regressions include origin and destination fixed effects, separately for each commodity and year. 'Effective distance between origin and destination along lowest-cost route' measures the railroad-equivalent kilometres (due to the normalized railroad freight rate to 1) between the centroid of the origin and destination trade blocks in question, along the lowest-cost route given relative freight rates for each mode of transport (as estimated in Table 2). 'Weight per unit value in 1880' is the weight (in maunds) per rupee, as measured by 1880 prices. 'Railroad freight class in 1880' is an indicator variable for all commodities that were classified in the higher (more expensive) freight class in 1880; salt was in the omitted category (low-value commodities). Heteroskedasticity robust standard errors adjusted for clustering at the exporting block level are reported in parentheses for columns 1 and 2 respectively. *** indicates statistically significantly different from zero at the 1% level; ** indicates 5% level; and * indicates 10% level.

- **(3) Railroads raise real income levels:** he finds an increase of **16% (!)**. This comes from the estimation of the welfare function that can be retrieved as:

$$\ln W_o = - \sum_k \mu_k \ln \lambda_k + \sum_k \frac{\mu_k}{\theta_k} \ln A_o^k - \sum_k \frac{\mu_k}{\theta_k} \ln \pi_{oo}^k$$

where the first two terms come from exogenous parameters (μ_k is essentially the preference for k in the utility function of consumers) and $\pi_{oo}^k = \frac{X_{oo}^k}{X_o^k}$ is the total expenditure in o on commodity k produced in the own region as a share of the total expenditure in o on commodity k .

- This is the **main result of the paper**

Table 4: Railroads and Real Income Levels (Step 3)

Dependent variable: log real agricultural income per acre	(1)	(2)	(3)	(4)
Railroad in district	0.164 (0.056)***	0.170 (0.095)*	0.188 (0.095)**	0.182 (0.073)**
Unbuilt railroad in district, abandoned after proposal stage	0.008 (0.020)			
Unbuilt railroad in district, abandoned after reconnaissance stage	-0.004 (0.050)			
Unbuilt railroad in district, abandoned after survey stage	0.012 (0.037)			
Unbuilt railroad in district, abandoned after sanction stage	0.008 (0.075)			
(Unbuilt railroad in district, included in Lawrence Plan 1869-1873) x (post-1869 indicator)	0.013 (0.057)			
(Unbuilt railroad in district, included in Lawrence Plan 1874-1878) x (post-1874 indicator)	-0.051 (0.067)			
(Unbuilt railroad in district, included in Lawrence Plan 1879-1883) x (post-1879 indicator)	0.005 (0.054)			
(Unbuilt railroad in district, included in Lawrence Plan 1884-1888) x (post-1884 indicator)	0.073 (0.098)			
(Unbuilt railroad in district, included in Lawrence Plan 1889-1893) x (post-1889 indicator)	-0.096 (0.088)			
(Unbuilt railroad in district, included in Lawrence Plan 1894-1898) x (post-1894 indicator)	0.044 (0.066)			
(Unbuilt railroad in district, included in Kennedy plan, high-priority) x (year-1848)	0.001 (0.025)			
(Unbuilt railroad in district, included in Kennedy plan, low-priority) x (year-1848)	-0.003 (0.029)			
Observations	14,111	14,111	14,111	14,111
R-squared	0.744	0.766	0.768	0.764

Notes: OLS Regressions estimating equation (16) using real income constructed from crop-level data on 17 principal agricultural crops (listed in Appendix A), from 235 districts in India, annually from 1870 to 1930. All regressions include district fixed effects and year fixed effects. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Unbuilt railroad in district, abandoned after X stage' is a dummy variable whose value is one if a line that was abandoned after 'X' stage penetrates a district, in all years after the line was first mentioned as reaching stage 'X' in official documents. Stages 'X' are: 'proposal', where the line was mentioned in official documents; 'reconnaissance', where the line route was explored by surveyors in rough detail; 'survey', where the exact route of the line and nature of all engineering works were decided on after detailed survey; and 'sanction', where the surveyed line was given official permission to be built. 'Lawrence 1868 plan' was a proposal for significant railroad expansion by India's Governor General that was not implemented; the plan detailed proposed dates of construction (in 5-year segments) over the next 30 years, which are used in the construction of this variable. 'Kennedy plan' was an early construction-cost minimizing routes plan drawn up by India's chief engineer in 1848 (divided into high- and low-priorities), which was rejected in favor of Dalhousie's direct routes plan. Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1% level; ** indicates 5% level; and * indicates 10% level.

- **(4) There exists a sufficient statistic for the welfare gains from railways:** this follows from the theory. The impact of railroad on welfare is captured by the impact of one endogenous variable, namely the share of that district's expenditure that it sources from itself (see (3)). Regressing income on railroads *and* this trade share, the coefficient of railroad becomes indeed insignificant. Only regressing on trade share, the coefficient is indeed very close to -1 , as it should be (see (3)).
 - **This is ultimately the test how much the model can explain.** And indeed it seems it can explain almost everything. The whole 16% increase in income comes from the trade cost reduction as modelled.

Table 5: A Sufficient Statistic for Railroad Impact (Step 4)

Dependent variable:	Log real agricultural income per acre, corrected for rainfall	(1)	(2)
Railroad in district		0.169 (0.051)***	0.023 (0.054)
"Trade share", as computed in model			-0.936 (0.131)***
Observations		14,111	14,111
R-squared		0.610	0.634

Notes: OLS Regressions estimating equation (18) using real income constructed from crop-level data on 17 principal agricultural crops (listed in Appendix A), from 235 districts in India, annually from 1870 to 1930. Dependent variable is log real income, corrected for crop-specific rainfall of each of 17 crops, weighted across crops as in equation (18). Regressions include district fixed effects and year fixed effects. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Trade share' is the share of a district's expenditure that it buys from itself; this variable is computed in the equilibrium of the model, where the model parameters are set to those estimated in Steps 1 and 2, and the exogenous variables (the transportation network, rainfall, and district land sizes) are as observed. Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1% level; ** indicates 5% level; and * indicates 10% level.

8.2.3 Dinkelman 2011

- Evaluates **impact of electrification** "on the ability of the poor to use their labour resources for market production", that is, community employment (especially female), wages and migration in KwaZulu-Natal, South Africa.
 - Finding: women's labour supply increases** significantly by around **9ppt** while men labour supply is not changed.
- Background:** After apartheid, National Electric Program: government-financed set up of grids in previously excluded areas. Quick roll-out into rural areas. Target poor African households rather than industrial users.
- Channels involved in electrification**
 - labour supply** (a priori ambiguously affected, but positive effects more likely):
 - (1) Housework becomes more productive.** The substitution effect could imply increase in home production. However, given that there are only a limited number of tasks (cooking, etc.) to perform at home, it is likely that (2) the dominant effect is the increase in effective time of day, for women mostly ("income effect" in terms of **greater time endowment**) which can be used to participate in paid labour.
 - (3) there could also be more work opportunities** outside the house: activities that become possible with electricity: corner shop with a fridge, selling hairdryers. NB: intensive (more hours) & extensive margin (more jobs) involves.
 - Effects do **not** come from some **industrialisation** effect which would have increased **labour demand** as no employment spillovers are observed
 - There is **migration** into electrified areas. However, excluding in-migrants, the effects on female employment growth are still significantly positive.
- FE effects on wages** are **negative** for men and positive for women, but **not significant** for either!

8.3 Empirical Strategy and Discussion

8.3.1 Jensen 2007

- Using the quasi-experimental setting, he runs:

$$Y_{rt} = \alpha + \sum_{r=1}^{II} \beta_r Region_r + \sum_{p=1}^3 \beta_p Period_p + \sum_{r=1}^{II} \sum_{p=1}^3 \beta_{rp} Region_r \times Period_p + \gamma Z_{rt} + \epsilon_{rt}$$

where Z is a vector of controls such as sea conditions, fuel price etc. This specification allows for different “treatment” effects in the different regions

- Concerns:
 - **Common trend** across regions has to hold for verification of causal effect. Look at the figures including pre-“treatment” trends. Extremely convincing.
 - **Non-random timing** of mobile phone introduction. However, again, no pre-trends visible that would suggest so.
 - **Migration** flows to regions where mobile phones were already introduced. However, figures on fishermen show that virtually all of them stayed in their catching zone before phones were introduced in their region.
- Strengths:
 - **Quasi-experimental setting**.
 - **Weekly survey data**.
 - **Extremely neat identification**.

8.3.2 Donaldson 2012

- He carries out four different regressions for the four different propositions explained above
- Concerns:
 - **Prop. 3.** There might be **endogeneity**. Migration flows, institutional changes and many other variables could be OMVs driving both railroad access and income. Donaldson provides some robustness by placebo tests: unbuilt railroad projects that were planned until some point in the final stage do not have a significant impact
 - **Old data?**
- Strengths:
 - **Placebo** test: planned, but unbuilt railroads have no significant effect, see Table 4 above.
 - **Big picture question**
 - **External validity**, not least because of the model.

8.3.3 Dinkelman 2011

- For the IV she runs:

$$\Delta y_{jdt} = \alpha_1 + \alpha_2 \Delta T_{jdt} + X_{jdt} \beta + \lambda_d + (\delta_j + \epsilon_{jdt})$$

$$\Delta T_{jdt} = \pi_0 + \pi_1 Z_j + X_{jdt} \pi_2 + \gamma_d + \tau_{jdt}$$

Where α_1, δ_j and λ_d stem from linear general, community j and district d specific time trends, X are covariates, T is electrification as instrumented by land gradient Z .

- **Identification assumption: land gradient not correlated with change in employment**, once electric grid and baseline community characteristics taken into account. Use placebo experiment to argue for the identification.
- Strengths:

- **Second regression (FE)**
- Many **robustness** checks, s.below.
- Investigation of **different channels** (labour supply, industrialisation, migration)
- Concerns:
 - For the FE regression, concerns about **strict exogeneity violations**.
 - **Exclusion restriction?** Land gradient might be correlated with better employment especially given agriculture plays a great role. She provides two robustness checks:
 - Inclusion of **placebos** of time periods electrified pre-1996, the starting period of her data are insignificant
 - The **instrument is not correlated with growth of major potential employers**, schools and white Indian adults (women would majorly work as teachers or domestic servants). This essentially provides robustness against a reverse causality argument saying labour demand shocks are concurrent with electrification
 - **Measurement:** land gradient might be picking up wealth differences between communities. **Robustness:** placebo technology changes like water nearby or flushed toilets are not influenced by electrification which would be the case if we're just picking up wealth differences.

8.4 Policy Conclusions

- **Infrastructure facilitates functioning markets** and **trade** and thus impacts on **welfare** and **growth**. It is of **overarching importance** for development efforts.
 - Jensen provides very neat evidence for how **ICT infrastructure** provides for the essential perfect information assumption and thus makes **markets function better** which in turn positively impacts on welfare.
- Infrastructure can take **many different forms**:
 - **Physical** infrastructure: roads, railway, etc.
 - **ICT** infrastructure: technology, computers, mobile phones, etc.
 - **Energy** supply/ electrification
- While most effects are obviously positive, some argue that more trade may do good to some people, but harm to others.
 - Donaldson makes a strong case for the **overall income increase** through trade – which is substantial (16%!).
- Dinkelman concludes that it is important to **consider all sorts of economic outcomes** electrification can have: income, health, education, etc.
- She also stresses that it is important to consider the **specific context** in which it works (in her case, post-Apartheid Africa)

9 Energy and Development

9.1 Papers and General Introduction

(Main) papers discussed:

- i. **Allcott, Collard-Wexler, O'Connell 2015. How Do Electricity Shortages Affect Industry? Evidence from India.** Working Paper

- Evaluate the effect of electricity shortages in India on firm profits, revenues and productivity, and material inputs
- ii. Lipscomb, Mobarak, Barham 2013. *Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil*. AEJ: Applied Economics
 - Assess the impact of electrification on local development in Brazil
- iii. McRae 2015. *Infrastructure Quality and the Subsidy Trap*. AER.
 - Explains the persistence of low-quality electricity infrastructure as a result of a subsidy trap and finds empirical confirmation in Colombia

General Theme:

- Energy is an **important constraint to growth**: it is key for most production activities
- **Energy consumption differential** between developing countries and US much larger than GDP p.c. differential
- The **causes** for energy under-supply and its **consequences** are important to consider for economic growth and development:
 - MR 2015: **subsidies** may be a reason for a **persistent low-quality and low-supply equilibrium (subsidy trap)**
 - ACO 2015: energy **shortages** are a **substantial drag on revenue** because they work like a high input tax. Contract solutions may help
 - LMB 2013: **electrification** has large **effects on development** in the **long run**

9.2 Main Ideas and Results

9.2.1 Allcott et al. 2015

- With extensively digitised data from India's Central Electricity Authority (CEA), they look at the effect of electricity shortages as instrumented by supply shifts on firm revenues, material inputs, productivity and profits in India.
- **Two reasons** for shortages:
 - McRae's (2015) **subsidy trap**
 - **Supply not keeping up with rapid demand growth**
- An electricity outage impacts on firms via two mechanisms:
 - **Input tax effect**:
 - For firms **with generators**: this is a **finite positive tax** which induces them to reduce material inputs. As not all of them are fully flexible, firms experience reductions in revenue productivity too.
 - For firms **without generators**: this is an **infinite tax** so they shut down temporarily.
 - **Input variation effect**:
 - Relative to using the same weighted (outage and non-outage periods) average of materials and electricity in *all* periods, firms have to vary materials and electricity depending on whether they are in a outage or non-outage period → **productivity loss**
- Findings:
 - Shortages do **not affect electricity intensity** of an industry, i.e. they cannot produce at lower levels of electricity intensity in the short run.
 - They significantly **increase the fuel share** as input cost in revenue (because of generators)

- **Materials** are **significantly reduced** while workers are insignificantly changed (not flexible enough).
- **Revenue is significantly reduced** (8.6%) and **TFPR decreases insignificantly**: TFPR unchanged as material input reduction is possible and accounts for revenue reduction
- Carrying out a **simulation** using a **production function model**, they find *no* significant difference in predictions on self-generation share increase, materials reduction, labour reduction, revenue loss and TFPR loss than the IV estimates
 - While **self-generating firms** experience a **producer surplus loss of 8.0%** of which 7.7% come from increased production cost (input tax effect) and only 0.3% from actual variable profit loss, **plants that don't use generators** experience a **10% producer surplus** loss, all of which naturally comes from variable profit loss as they simply have to shut down
 - The small insignificant changes in revenue productivity for self-generating firms in the short run show that **energy is not able to explain the productivity differences** between developed and developing countries in the short run
- A counterfactual analysis reveals that **variable profit losses could be substantially reduced with interruptible contracts** (especially for those firms not using generators: reduction from 10% to 0.7%; for generator-using firms from 0.3% to 0.1%)
- As generators have substantial economies of scale, shortages also **distort the distribution of plant sizes**: small firms are more severely affected by shortages.

9.2.2 Lipscomb et al. 2013

- Using data on electricity infrastructure from 1960-2000 in Brazil and **modelling of hydropower dam placement** (as an instrument), they look at the effect of electrification on development
- **Rate of return** based on the increase of land values *after* 10-12% payment to electrification companies are **184%** (61m USD increase in land value for a county electrified with cost of 33m USD for provision)
- They find **large and highly significant effects on HDI and housing values** which exceed OLS estimation findings (clearly suffering from endogeneity such as reverse causality, ...). Two potential reasons:
 - **Productivity** increases
 - **In-migration** of most productive workers
- They find that in-migration is unlikely to account for the large magnitude
- Underlying mechanisms: highly significant and mostly substantial effects on
 - **Employment** (pos.)
 - **Income** (pos.)
 - **Education** (pos.)
 - **Poverty** (neg.)

9.2.3 McRae 2015

- **Subsidy trap** works along the following lines:
 - **Households** with informal connections get **low-quality electricity provision** and **do not pay** for that
 - **Utility firms** can cover their provision cost through **government subsidies**

- The **government provides subsidies** in order to **retain political support** from the poor households and prevent (violent) unrest (often observed)
 - **Utility firms do not upgrade** their network as the **capital cost of doing so exceeds the extra profit** they would make
- Therefore, government subsidy policy perpetuates the existence of low-quality equilibrium and **disincentivises investment** in electricity infrastructure
 - **Empirically**, he finds **confirmation** that extra profit from an upgrade does not exceed the capital cost of the upgrade
- **Upgrading** connections involves
 - **Installation of meters** to bill households according to their true usage (non-zero marginal price)
 - **Improved quality and reliability** through higher technical standard
 - **Individual connection** to the network so that disconnection as a punishment for non-payment is possible and incentivises payment
- Using data from metered households with similar characteristics to households in informal settings from Colombia and a structural model, he **estimates electricity demand**. **Upgrade** will induce **two changes**:
 - **More demand because of higher quality**
 - **Less demand because of positive (non-zero) marginal price**
 - Empirically, demand slightly decreases, i.e. the **second effect dominates**
- The government could pursue **four policies** to circumvent the problem:
 - **Reduce transfers before the upgrade**
 - **Provide additional transfers after the upgrade**
 - **Increase household consumption** after the upgrade
 - **Subsidise cost of the upgrade**
- It is found that, under an optimal policy mix, every county could be **upgraded at a cost of 34% of the current programme**; if there are constraints that firms cannot be made worse off, it is 23% of current programme cost.

9.3 Empirical Strategy and Discussion

9.3.1 Allcott et al. 2015

- They construct an **instrument for shortages** using two supply shifts:
 - **Hydroelectric generation** which depends on rainfall (it turns out that nearly all the identification comes from this part)
 - **Capacity addition** through new plants conditional on state trends
- They estimate:

$$Y_{ijst} = \rho S_{st} + \beta R_{st} + \nu W_{st} + \Lambda_{st} + \psi_s t + \theta_t + \mu_{jt} + \phi_i + \epsilon_{ijst}$$

Where Y is the outcome of interest in plant i , industry j , state s and time t , S is the instrumented shortage, R are rainfall controls (as rainfall might have a direct effect other than via hydroelectric generation supply shifts), W are cooling-degree controls, Λ account for state splitting, and some fixed effects with and without time trends.

- Concerns:
 - For **capacity addition** to be **exogenous** they assume that it is stochastic in which a new plant comes online because of the **unpredictable time-to-build**. Is

it really so unpredictable? Probably not quite. However, the first part of the instrument is much more important for identification anyway (apparently).

- Strengths:
 - Inclusion of **rainfall controls**
 - **Counterfactual analysis**

9.3.2 Lipscomb et al. 2013

- They construct a **model forecasting hydropower dam placement** to develop a hypothetical electrical grid over Brazil and use it as an **instrument** in an IV regression:

$$\text{1st stage: } Y_{ct} = \alpha_c^1 + \gamma_t^1 + \beta E_{c,(t-1)} + \epsilon_{ct}$$

$$\text{2nd stage: } E_{c,(t-1)} = \alpha_c^2 + \gamma_t^1 + \theta Z_{c,(t-1)} + \eta_{ct}$$

Where Y is a development outcome in county c , time t . E is the proportion of grid points in the county electrified and Z the respective hypothetical proportion (instrument). The hypothetical grid is a function of geographic characteristics determining suitability and the construction budget in each decade.

- Concerns:
 - **Exclusion restriction:** If people move along the same lines as are hydropower dams built because of the geographic characteristics (water-rich area, etc.) and this influences development, the exclusion restriction would be violated. They provide robustness against that: **Rank order correlation** between suitability for hydropower dams and suitability for movement according to GDP and population density is very small.
 - **Lagged development** could be a driver of the hypothetical placement and serially correlated. They show that **this is not the case**.
- Strengths:
 - **External validity**
 - A number of **robustness checks** (alternative specification, additional controls, etc.) run

9.3.3 McRae 2015

- He comes up with a **structural model of electricity demand** through household appliances needing electricity and depending on the probability of outage
- To figure out demand, regress annual electricity consumption on appliance holdings, household characteristics and dummies for regions
- **Demand model** is used to **predict consumption** before and after the hypothetical upgrade for the policy analysis.
- The rest is just a simple accounting exercise computing profits before and after the change depending on the change in subsidies and revenue. Capital cost estimation is based on figures from the national Energy ministry.
- Concerns:
 - **Unmetered households** for which demand is calculated based on the structural model may be **different** in decisive characteristics **from metered households** which are the basis for finding an electricity demand function
 - Rather **rough calculations** and assumptions on some figures

- Strengths:
 - **Policy counterfactual analysis**
 - **Strong theoretical foundation**

9.4 Policy Conclusions

- The **importance of energy provision** for development and growth is underlined by both
 - The **large long run effects** of electrification observed in Brasil (184% return!)
 - The **effect of shortages on producer surplus** in India of almost -10% which comes from profit loss directly if firms cannot even substitute to using (more costly) generators
- **Productivity** is however **not significantly touched** in the short run if plants can substitute to using generators
- Still, there is a **huge energy consumption differential** between developed and developing countries (larger than GDP p.c. differential). This may be due to a **subsidy trap**, i.e. consumption subsidies disincentivising investment in better energy provision
- **Capital investment subsidies or conditional subsidies** may be an answer to incentivise investment in energy upgrades
- Technocratic solutions such as **interruptible contracts** may be a technocratic solution to reduce producer surplus in the face of energy shortages.

10 Environment and Development

10.1 Papers and General Introduction

(Main) papers discussed:

- iv. Burgess, Hansen, Olken, Potapov, Sieber 2012. *The Political Economy of Deforestation in the Tropics*. QJE.
 - Assess how the number of political jurisdictions at the local level impacts on deforestation
- v. Duflo, Greenstone, Pande, Ryan 2013. *Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India*. QJE.
 - Investigate pollution reporting and consequences of incentivising auditors on it in India
- vi. Davis 2008. *The Effect of Driving Restrictions on Air Quality in Mexico City*. JPE.
 - Assesses the effect of a driving restriction programme in Mexico on air quality.

General Theme:

- Pollution and other environmental issues are problems bound to many market failures which is why people or even states will not *privately* deal with them to the appropriate extent:
 - Global externalities without any global social planner
 - Intergenerational inconsistency
 - Long-run consequences that are not immediately observable
 - Clean air and water are common resources (free-rider problem)

- Environmental regulations or policies cannot ignore basic incentive structures of the people they affect
 - BHOPS 2012: increased number of political jurisdictions with control over deforestation will – according to Cournot competition – increase deforestation
 - DGPR 2013: auditing of environmental standards is corrupted because of the auditors' conflict of interest and third-party auditing both induces truth-telling and lower pollution
 - D 2008: Mexico's air pollution restriction HNC failed because people simply bought more cars instead of substituting to public transport

10.2 Main Ideas and Results

10.2.1 Burgess et al. 2012

- Use satellite data and official data on deforestation and oil and gas rents from the government after the implementation of a decentralisation law in 2001 which gave local governments responsibilities for deforestation, oil and gas revenues and led to substantial district splitting over years, they investigate
 - The effect of district splitting on the (legal and illegal) quantity of deforestation and the price of deforestation
 - The influence of oil and gas rents in this context
- Using satellite data, a main finding is that legal and illegal deforestation increases significantly (in a province) as the number of districts increases in the short-run and the medium run (sum of three lag period coefficients)
 - Some zones are conservation zones in which deforestation is entirely illegal. Increases here can be said to be illegal with certainty
- Using official data (which may not capture illegal quantities of deforestation but certainly the market price), they find that the price of wood decreases as the number of districts increases in the short and in the medium run
 - Price elasticity is estimated to be around -2.3
- They find that deforestation decreases significantly in higher oil and gas rents in the short run, but this effect doesn't persist in the long run
 - They provide some suggestive evidence that medium-run effects do not exist because the political equilibrium changes in that new politicians who are more interested in rent-seeking come into power to extract rents from both oil/gas and deforestation once rents from the former increase.
- Theory: Cournot competition between districts. Firms are willing to pay $b = p(Q) - c$ for a unit permit or bribe and districts d solve:

$$\max_{q_d} b(q_d)q_d - \pi(q_d, \bar{q})r_d = p(Q)q_d - cq_d - \pi(q_d, \bar{q})r_d$$

which yields from the FOC:

$$\frac{p - c}{p} = \frac{1}{n\epsilon} + \frac{\pi\left(\frac{Q}{n}, \bar{q}\right)r_d}{p}$$

where π is the probability of detection when going beyond the legal quota \bar{q} and $\epsilon = \frac{dq}{dp} \frac{p}{q}$ is the demand elasticity, Q is the total production sum and r is the politician's discounted

sum of future rents. Cournot competition predicts higher quantities and lower prices for increased competition n and substitution between logging and other form of rents r

- The findings are perfectly consistent with the theory.

10.2.2 Duflo et al. 2013

- Using data from audit reports and backchecks in India for an RCT with treatment consisting of random auditor allocation, flat auditor payment from central pool rather than the plant, random backcheck verification for some auditors and second-year performance pay on audit report accuracy, they investigate environmental audits and find:
 - Current audit-reporting is corrupted; 29% of audits falsely report below the regulatory standard
 - Treatment caused auditors to be more truthful – 50% to 70% higher readings
 - Treatment reduced emissions; average pollution decreased by 0.21 standard deviations
- Some tentative numbers on costs and benefits suggest there is a net benefit in providing these incentives, i.e. third-party auditing.

10.2.3 Davis 2008

- *Hoy No Circula (HNC)* programme in Mexico foresees that car owners with a certain last digits on number plates must not use their car on certain weekdays. It's well enforced so compliance is great
- Using hourly data on air quality, he finds
 - That there is *no* overall improvement in air quality → programme failure; with no improvement in air quality and high social cost from more car purchases (indirect measure of what people are willing to pay to circumvent the driving restriction), this programme is difficult to justify on a cost-benefit basis
 - There is some substitution to increased weekend and late night air pollution which is when the regulation doesn't apply
- Further results:
 - The number of cars used *increases* and there are more high-emissions vehicles used
 - Gasoline sales do not change significantly
 - There is no substitution effect towards public transportation or use of taxis

10.3 Empirical Strategy and Discussion

10.3.1 Burgess et al. 2012

- For the quantity ((1)), price ((2))and oil and gas rent ((3))) analysis, they use the following FE Poisson Quasi-MLE specifications ((1) and (3)) and FE ((2)) specifications:

$$(1) E(deforest_{pit}) = \mu_{pi} \exp(\beta NumDistrictsInProv_{pit} + \eta_{it})$$

$$(2) \log y_{wipt} = \beta NumDistrictsInProv_{pit} + \mu_{wpi} + \eta_{wit} + \epsilon_{wipt}$$

$$(3) E(deforest_{dit}) = \mu_{di} \exp(\beta PCOilGas_{dit} + \gamma Numdistricts_{dit} + \eta_{it})$$

where i denotes island, p denotes province (< island), d denotes district (< province) and w denotes wood type

- Concerns:
 - Critical assumption of the model and thus its consistency with the empirics: firms need to bribe only one district and not several. This might not hold true
 - The regressions have no controls beyond some fixed effects. In (1), differential economic growth in districts or population development may drive both the number of districts in a province and the level of deforestation.
 - Timing of splitting might be correlated with underlying trend in logging. However, robustness regressions with various types of trends do not change the coefficients.
- Strengths:
 - Satellite date reveals illegal logging as well and gives perfectly reliable data
 - Poisson Count model provides good specification for many zeros as dependent variable and allows for arbitrary distribution

10.3.2 Duflo et al. 2013

- For the first two findings, the authors estimated:

$$1\{\text{Compliant}\}_{ij} = \beta_1 1\{\text{AuditReport}\} + \beta_2 1\{\text{AuditReport}\} \times T_j + \beta_3 T_j + \alpha_r + \epsilon_{ij}$$

Where the indicator on *Compliant* is 1 if the reading is between 75% and 100% of the regulatory standard in the backcheck and the indicator on *AuditReport* is 1 if it is so in the audit report (in a second regression, the indicator indicates readings simply below the regulatory standard). *T* is the treatment variable. This is run without the interaction effect and treatment effect for the first finding.

- For the third finding the authors just regressed pollution result on the treatment dummy and a regional FE
- Concerns:
 - Since treated plants react by reducing emissions, auditors reporting lower pollution levels in treatment are not only driven by more truthful behaviour but also merely by the fact that pollution is lower in treatment. As a robustness, the author analyse the difference between backcheck readings and audit report readings and still find a significant reduction in this difference for treatment.
- Strengths:
 - Randomisation makes identification very reliable.

10.3.3 Davis 2008

- The basic estimating equation is:

$$y_t = \gamma_0 + \gamma_1 1(HNC_t) + \gamma_2 x_t + u_t$$

Where variables are in logs and *x* is a vector of covariates including month of the year, day of the week, hour of the day and interactions between weekends and hour of the day as well as weather variables

- To account for endogeneity, a regression discontinuity design is employed with a highly flexible polynomial time trend - observations in a very narrow time interval around the HNC implementation are looked at where other unobserved factors are arguably constant.
- Concerns:

- There may be changes in monitoring stations' reportings after the HNC implementation which are correlated with the pollution level. That would pose an OMV problem. However, some of the alternative specifications take this into account and display no significance.
- Strengths:
 - A number of alternative specifications is run
 - Variance matrix accounting for serial correlation is estimated. For little time intervals there is probably a high degree of serial correlation
 - A great number of alternative specifications is estimated with no significant results.

10.4 Policy Conclusions

- Incentive structures have to be taken into account when designing effective environmental policy
 - A higher degree of decentralisation of control over natural resources can actually decrease the degree of conservation
 - Auditing on pollution standards is not truthful when auditors are paid by the firms they are supposed to audit because they find themselves in a fundamental conflict of interest. This improves if adequate incentives are introduced
 - Behavioural responses to circumvent regulations can seriously undermine the benefits of a mere regulation
- Specifically, the evidence on deforestation suggests that it can be decreased by
 - Increasing the probability of detection through providing more resources to the enforcement of the quotas
 - Providing alternative rents to local politicians such as funding from environmental funds
- Third-party auditing of environmental regulations (and generally third-party auditing) provides incentives for truth-telling and the benefits of that may exceed the cost of third-party auditing.
- Before providing driving restrictions to decrease air pollution, the substitution responses by people should be carefully thought out to prevent ineffectiveness

11 Child Labour

11.1 Child Labour Model (Basu & Van, 1998)

- Two hypotheses
 - **Luxury Axiom:** families send their children to work if they need to because consumption drops very low
 - **Substitution Axiom:** firms can substitute between adult and child labour
- There are N identical households with one adult and one child
- For $c \geq s$, $(c, 0) \succ (c + \delta, 1)$ and for $c < s$, $(c + \delta, 1) \succ (c, 0)$ where $e = 1$ means child labour. The household's budget constraint is $2c \leq ew_C + w_A$.
- Firms produce $y_i = f(A_i + \gamma C_i)$ where A and C are adult and child labour input respectively
- There are **multiple equilibria**:

- For $2s < w_A < w_C/\gamma$ **only adults work at high wage**
- For $2s > w_A > w_C/\gamma$ **adults and children work at low wage**
- There are multiple equilibria because high adult supply not only means lower wage (market mechanism) but can also lead to w_A dropping below $2s$ which introduces child labour (**pecuniary externality**)
- A ban on children can **lead to the first equilibrium** here

11.2 Empirical Verification and Work

- Doran (2013) verifies the **substitution axiom** using a **Mexican programme** which increased **returns to education** and thus reduces child labour supply.
 - Adult and child labour being substitutes means reduced child labour induces increased demand for adult labour. No matter whether adult labour supply shifted or stayed constant, such a shift induces *both* wages and employment to increase
 - He finds increases in both using a DID
- Edmonds and Schady (2012) use a **cash transfer programme** in Ethiopia to verify the **luxury axiom**. If it holds, a cash transfer should reduce child labour.
 - They empirically (IV) find this.
 - They find that it does not completely offset child labour which is reasonable because the transfer is far below the child wage.
- Bharadwaj et al. (2013) extend the model to the case of **two sectors** and show that a ban may decrease wages for children and **increase child labour**
 - They can empirically (DID) verify this in India and also find a **negative effect on household welfare**

12 Growth Theory / Development and Growth Accounting

12.1 The Development Accounting Exercise

- Development Accounting compares **differences in income/GDP** between developed and developing countries to **differences in observable components** of physical and human capital
- From $Y = F(A, K, Lh) = AK^\alpha(Lh)^{1-\alpha}$ with h being human capital, you get the following “success” rate in per worker terms:

$$\text{success} = \frac{\text{var}[\log(k^\alpha h^{1-\alpha})]}{\text{var}[\log(y)]}$$

- This is typically around **40%** (Caselli, 2005 [“Accounting for cross-country income differences”]). Taking not only schooling, but also experience into account when evaluating h raises it to around 60% (Lagakos et al., 2014)
 - The measure of 40% is highly **sensitive to the specification of α** . It increases a lot as α increases.
- In order to do development accounting, you **need to find out α , k and h** (and y).
 - Assume factors are paid their marginal products: $r = MPK = \frac{\alpha Y}{K}$, $w = MPL = \frac{(1-\alpha)Y}{L}$

- **α is easily retrievable** as $1 - \alpha = \frac{wL}{Y}$, a **country's labour share**. Turns out to be around 2/3 pretty much all over the world.
- k can be retrieved from developing a measure of K and dividing by L . The **inventory method** gives $K_t = I_t + (1 - \delta)K_{t-1}$ with $K_0 = \frac{I_0}{g+\delta}$, g is the average geometric growth rate for investment, δ is depreciation (expression for K_0 from the Solow Model steady state solution). This is quite **crude** however.
- Better is to **micro-estimate MPK** to find $K = \frac{\alpha Y}{MPK}$ from there.
- **h is found in micro-estimates** using wages and schooling data, typically running a regression of \ln wage on schooling where the coefficient reflects the marginal return⁵. The underlying idea is that \ln wage and $\ln h$ are proportional and h is specified as $h = \exp(\phi S)$ where s is yrs. of schooling.

$$w = MPL = \frac{(1 - \alpha)Y}{L} \rightarrow \ln w \propto \ln h = \phi S$$

- You can also account for schooling quality using measures of test scores and their effect on wages (instead of yrs. of schooling). In Caselli's (2005) paper, this does *not* make an appreciable difference. More recent evidence suggest there is a 10% gain in the accounting evidence using this approach.

Lagakos et al. (2014). *Life-Cycle Wage Growth Across Growth*. JPE forthcoming. [formerly: *Experience Matters: Human Capital and Development Accounting*.]

- Lagakos et al. (2014) provide an estimate of **human capital** including **not only schooling but also experience** in a cross-country regression. They find that the effect of experience on wages, the **experience-wage profile**, is **lower in developing** than in developed countries. If this is incorporated in **development accounting**, around **60%** of income differences can be explained.
 - They have $h_{ict} = \exp(g(s_{ict}) + f(x_{ict}))$ for human capital where x_{ict} is experience for individual i in cohort c and time t . Wage is $w_{ict} = w_{ct}h_{ict} \exp(\epsilon_{ict})$ where the skill price $w_{ct} = \bar{w} \exp(\gamma_t + \psi_c)$.
 - **Concerns** with this regression
 - Basic assumption that **wage equals MPL** is very unrealistic
 - **Cross-country heterogeneity**: difference in contract structures, composition private/public sector, repeat years of schooling, other country differences
 - **Sectoral composition** may differ across countries
 - **Spillover effects** might systematically differ
 - They also run a **regression for US immigrants**, so the country is the same
 - **Concerns**:
 - **Selection into US migration** which may be systematically different for developed and developing countries

⁵ If ways are found to deal with the obvious endogeneity problem. Note, however, that there is an argument (Card, 1999, I think...) that in many contexts the upward bias from omitted ability and the downward bias from measurement error cancel out pretty much in a simple OLS regression.

- Immigrants from different countries have **systematically different professions** in the US with different experience-wage profiles

12.2 MPK and Return on Financial Capital

- In simple regression of profits on firm capital (recall from Euler's Theorem, $Y = F_K K + F_L L$, underlying skill/entrepreneurial ability is an OMV causing **endogeneity** (influences both capital and profit)
- De Mel et al. (2008) conducted an experiment providing **cash grants** as (random) treatments for **small businesses in Sri Lanka**
 - Two concerns:
 - Stable Unit Treatment Value Assumption (**SUTVA**): is the control group unaffected by treatment?
 - **GE effects**: are there **spillover effects**. These would be important to see the effect on the aggregate economy
- They make **three investigations**:
 - (General) treatment effect: **return on financial capital** extremely high (55-63% real annual return); this works via many routes not only increase in capital.
 - They use the treatment as instrument in an **IV regression** of profit on capital, time dummies and individual fixed effects to **isolate the marginal return to capital**. Since treatment not only affects capital, but for instance labour too, they adjust profits to this end.
 - Finding: **5% monthly** (!) return to capital
 - To account for **GE effects**, they use **GPS data** to run a regression of profits on treatment and N_{it} which is the number of firms within 500m and 1km
 - Finding: **negative spillover effects** in 500m environment
- Karlan et al. (2014) provide **Ghanaian tailors** with financial capital in an experiment and observe **increased capital** but **decreases profit** in the **short run** and leads to **disinvestment in the long run**
 - **External validity** is questionable because of specific context
- On the macro level, Caselli & Feyrer (2007) look at **adjustments to MPK** calculation taking the difference between **natural** and **physical capital** into account
 - α is the share going to *all* sorts of capital / all non-labour inputs. Natural capital (land, timber, etc.) is part of that but we are only interested in physical reproducible capital, the payments to which are less than α . If this is **differential between developed and developing countries** (likely), MPK (only physical capital) is asymmetrically wrongly assessed
 - Without accounting for this, **MPK** in developing countries is more than twice as large as in developed countries. Taking this into account, it very much **equalises** which would be an explanation of why capital doesn't flow massively to developing countries.

12.3 Growth

- **Kaldor facts:**
 - Labour productivity has grown at a sustained rate
 - Capital per worker has grown at a sustained rate

- The real interest rate or return on capital has been stable
- The ratio of capital to output has been stable
- Capital and labour have captured stable shares of national income
- Among the fast growing economies of the world, there is an appreciable variation in the rate of growth of the order of 2-5 percent

12.4 Growth Accounting

- Idea: **Decompose economic growth** into what is due to changes in capital and human capital over time as opposed to changes in TFP
 - This is the same accounting exercise **within a country in the time series** that development accounting does for the **cross-section between countries**
- From $Y = F(A, K, H)$, you have $dY = F_K dK + F_H dH + F_A dA$ and growth can be decomposed as:

$$\frac{dY}{Y} = F_K \frac{K}{Y} \frac{dK}{K} + F_H \frac{H}{Y} \frac{dH}{H} + F_A \frac{A}{Y} \frac{dA}{A} = \alpha \frac{dK}{K} + (1 - \alpha) \frac{dH}{H} + \Delta TFP$$

- Change in TFP is change in GDP that changes in capital and human capital cannot account for. Applying this to East Asian countries, Young (1995) finds that **TFP growth is actually far lower in Hong Kong, (especially) Singapore, South Korea and Taiwan** than other authors estimated.

13 Misallocation

13.1 The Idea of Misallocation

- There are two types of efficiency:
 - **technical/productive efficiency**: Operating at the best technical standard
 - **allocative efficiency**: Allocating factors of production according to their marginal product efficiently within firms, across firms, across sectors, spatially, ...

Misallocation is about allocative inefficiency
- TFP is a function of both technical and allocative efficiency. Research on misallocation tries to find out what part misallocation has to play for TFP
- We have misallocation on the **intensive margin** if marginal products of inputs are not equated across firms and all sorts of other dimensions, e.g. specifically for capital and firms i , if $p_i f'_i(k) \neq p_j f'_j(k)$ for some firms i, j .
- We have misallocation on the **extensive margins** if some firms fail to operate although they have a higher marginal product of capital at $k = 0$ than another firm, e.g. this is certainly fulfilled if Inada conditions hold.
- *Example:* For the simple production of lattes and steel $Y = X_s^\alpha X_l^{1-\alpha}$, and production of X from capital $X_i = AK_i$ with total capital $\bar{K} = K_s + K_l$, a share β is allocated to produce steel and $1 - \beta$ to produce lattes.
 - This yields: $Y = A\beta^\alpha(1 - \beta)^{1-\alpha}\bar{K}$ where A is productive efficiency and $\beta^\alpha(1 - \beta)^{1-\alpha}$ captures allocative efficiency in the sense that it is maximised for the efficient solution
 - The (allocatively) efficient share β given to X_s is $\beta = \alpha$ from $\max Y = AK_s^\alpha K_l^{1-\alpha}$ s.t. $\bar{K} = K_s + K_l$. This gives $Y = A\alpha^\alpha(1 - \alpha)^{1-\alpha}\bar{K}$

- Now, if the budget constraint becomes $\bar{K} = (1 + \tau)K_S + K_L$ because of a distortion or tax on buying steel, the optimal solution from the modified problem becomes $K_S = \frac{\alpha}{1+\tau} \bar{K}$, $K_L = (1 - \alpha)\bar{K}$. This gives $Y = A(\frac{\alpha}{1+\tau})^\alpha (1 - \alpha)^{1-\alpha} \bar{K}$ which is obviously less than before.
- This **distortion** can is a **way to capture/measure misallocation**
- **Restuccia and Rogerson (2008)** [“*Misallocation and productivity*”] provide a little summary and distinguish between the direct and indirect approach:
 - Direct approach: pick one specific mechanism of misallocation (e.g. factor affected by a certain tax), obtain a measure of it and assess to what extent the specific aspect generates misallocation and aggregate TFP
 - Indirect approach: tries to identify the net effect of the composition of various factors underlying misallocation. Essentially write down a model with all sort of wedges on the FOCs.

13.2 Macro Evidence and General Framework

Hsieh & Klenow (2009). Misallocation and Manufacturing TFP in China and India. QJE.

- Hsieh & Klenow (2009) develop a model with distortions capturing the degree of misallocation *across* firms (others have looked *within* firms) and use it to quantify misallocation in China and India relative to the US
- They essentially use a Melitz (2003) model w/o international trade and exploit the key result that *revenue* productivity (physical productivity x firm's output price) should be equated across firms in the absence of distortions.
- Key assumptions:
 - Intermediate inputs are **imperfectly substitutable**
 - Firms are in **monopolistic competition** within a sector according to **Dixit-Stiglitz**:

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

- Production functions are **Cobb-Douglas**
- TFP of a firm A_{si} and revenue productivity $TFPR_{si}$ ($= P_{si}A_{si}$) are **jointly log-normal**
- A_{si} is **exogenous**, i.e. no consideration of the extensive margin
- Consider firms i in a sector s with production function $Y_{si} = A_{si}K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}$, maximising

$$\pi_{si} = (1 - \tau_{Y,si})P_{si}Y_{si} - wL_{si} - (1 + \tau_{K,si})K_{si}$$

i.e. facing a **firm-specific output** and **capital distortion** that is to measure misallocation

- The aim is to compare TFP_s with a hypothetical TFP_s^* , TFP without misallocation, to compare the degree of misallocation
 - In a sector, $Y_s = TFP_s K_s^\alpha L_s^{1-\alpha}$
- Problem: A lot of **important variables are unknown** when firm level data on revenues and inputs is only available, namely Y_s , P_s and A_{si}
- Using the monopolistic competition set up, it can be shown that:

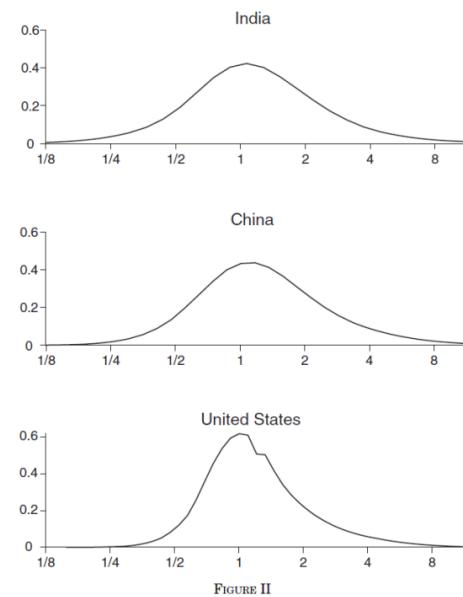
$$TFP_S^* = \left(\sum_i A_{si}^{\sigma-1} \right)^{\frac{1}{\sigma-1}} ; TFP_S = \left(\sum_i \left(A_{si} \frac{\overline{TFPR}_S}{TFPR_{si}} \right)^{\sigma-1} \right)^{\frac{1}{\sigma-1}}$$

$$\log(TFP_S) = \frac{1}{\sigma-1} \log \left(\sum_i A_{si}^{\sigma-1} \right) - \frac{\sigma}{2} \text{var}[\log(TFPR_{si})]$$

$$A_{si} = \kappa_s \frac{(P_{si} Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_i^\alpha L_i^{1-\alpha}}$$

where κ is just a scaling factor and need not be known for the comparison since it is the same for every firm and $\overline{TFPR}_S = P_S TFP_S$, the (average) revenue productivity

- While TFP may differ across firms, **firm revenue productivity $TFPR_i$ should be equated** (in which case TFP_S just becomes TFP_S^*) for allocative efficiency.
- Hsieh & Klenow (2009) find that so doing would **much more heavily increase TFP in China and India (both around 100%) than in the US (around 35%)**. The fact that misallocation may happen to a much greater extent in developing than in developed countries may explain why TFP differences are so large.
 - The more meaningful comparison would be to look at the gains if China's and India's TFPRs would be equated to the same extent as in the US (taking the **US as the benchmark** with the lowest possible amount of distortions). This gives gains of **30% in China and 60% in India in 2005**
- There are many more firms who have really low TFP in India and China (**larger left tail**)
- Hsieh & Klenow (2009) argue that this misallocation may in large parts be due to **policies and institutions** such as state-owned enterprises in China and entry restrictions in India
 - They show that state and foreign companies perform worse in terms of TFPR than the average in China (in particular the former: 34% lower!) and state or joint public/private firms perform worse than the average in India
- Ziebarth (2013) performs the same exercise for **19th century US** and finds great **similarities to today's China and India**. As such institutions and policies have not played a major role in 19th century US, he argues that different reasons may mainly be underlying misallocation such as
 - **Differences in transportation networks**
 - **Lack of competitive regulation**
- Further reasons for misallocation could be:
 - **Firm-specific mark-ups**
 - **Financial frictions**

FIGURE II
Distribution of TFPR

Fernandes & Pakes (2008). Factor utilization in Indian Manufacturing: A look at the World Bank investment climate surveys data. NBER WP.

- FP are trying to measure the extent of underutilisation of labour and overutilization of capital in India (factor misallocation), motivated by the fact that the growth rate of value added in the Indian manufacturing sector is half that of China and Vietnam while the growth rate of services is comparable with China and even higher than in many other countries. Why does Indian manufacturing perform so poorly?
- Some authors see reasons in labour market frictions (see, e.g., Besley-Burgess-2004)
- They set up a structural model to estimate the distribution of firm productivity and then find out whether factor employment (both labour and capital) in India is “abnormally” high or low conditional on (i) firm productivity, (ii) the amount of other factors used in production, and (iii) factor cost.
- Basic model (in logs!) is a sales-generating function (i.e. linking *sales* and input rather than output and input):

$$y_{jt} = \beta_l l_{jt} + \beta_k k_{jt} + \beta_z z_{jt} + \omega_{jt} + \eta_{jt}$$

Where (y, l, k, z, ω, η) are value added (sales minus material inputs), labour, capital, temporary labour, productivity (unobserved) and an error

- Endogeneity concerns they address using some IO techniques:
 - Firms with low productivity draw may shut down, so only the truncated distribution of ω can be found
 - Firms know productivity when choosing inputs and productivity shocks are correlated over time; therefore input choices are clearly correlated with the unobserved expected shocks
- Main results:
 - Massive under-utilisation of labour, defined as the ratio of optimal labour to actual labour used. Average ratio is 5.8 in 2001 and 3.4 in 2004
 - Over-utilisation of capital, though to a lesser extent. The average ratio is 0.75

Figure 2. Underutilization of Labor by Size Group

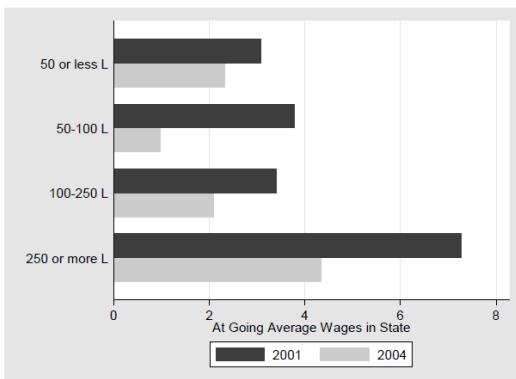
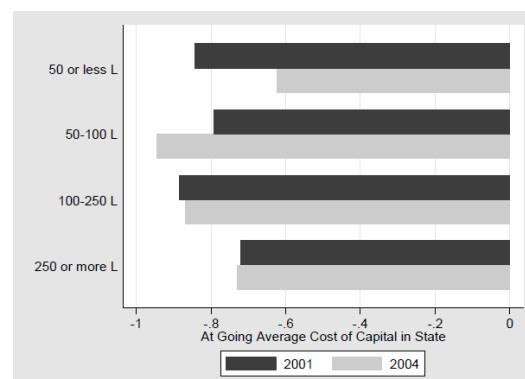


Figure 4. Overutilization of Capital by Size Group



- over-utilisation of capital could of course be a response to under-utilisation of labour
- Concerns:
 - Strong assumptions in the sales-generating function. Unclear whether this is robust to a non-Cobb Douglas function.

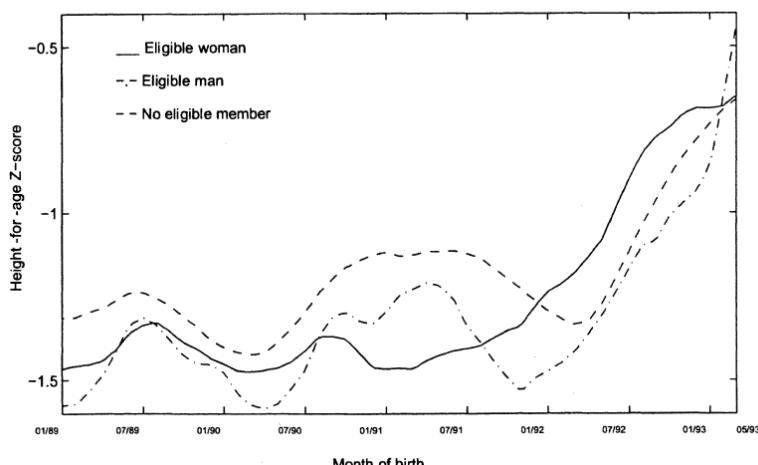
- What labour frictions could generate a shadow wage up to 10x the market wage in some states (alternative calculation: what wage would you need to have the actual amount of labour be optimal)

13.3 Intrahousehold Misallocation

Duflo (2003). Grandmothers and Granddaughters: Old-Age Pensions and IntraHousehold Allocation in South Africa. WB Econ Rev

- This is essentially a test of the unitary household, that is, the idea that – within a household – no matter who earns what amount of income, decisions are taken as a unit and each member's consumption responds equally to changes in each member's income (income pooling $\frac{\partial c_i}{\partial y_j}$ is the same for all i, j members of the household)
- Context: Post-apartheid, in the early 1990s, the pension scheme was expanded for the black population (significant income shock: twice their median wage) where women became eligible from the age of 60 and men from the age of 65.
- Duflo evaluates the effect on the health of children within the household of eligible people using two methods:
 - (i) RDD: Compare households with women between 55 and 60 to those with eligible women older than 60 and, equivalently, those with men between 60 to 65 to those with eligible men older than 65. The main result uses the cutoffs as an IV for receiving a pension (fuzzy RDD).
 - Result: If women are eligible, the weight-for-age score increases by 1.19 standard deviations for girls and not for boys. Eligible men don't show any effect for either gender
 - Concern: First, households with older grandparents are systematically different. Second, endogenous change in the household structure as a response to the pension.
 - (ii) DID-like... apparently, height-for-age scores are very quickly responsive to changes in the environment for *young* children. Suppose there are systematic differences in the health of children between eligible and non-eligible children, in particular, health might be better for non-eligible households (more privileged). Then the difference between eligible and non-eligible households should

FIGURE 1. Height for Age of Children Living with Eligible Women, Eligible Men, No Eligible Member



be smaller for households with younger children since they were well nourished as a response to the pension scheme and resulting income increase for a *larger fraction of their life*. Looking at height-for-age scores of children in households that are non-eligible, have an eligible man, and have an eligible woman *by birth year of the child* clearly shows the desired pattern: Non-eligible households score better for older

- children and those with eligible women score better for younger children (i.e. the trend is even reversed)
- Concern: Focus in identification very much on whether eligible households are comparable to non-eligible households. However, one of the distinctions drawn in the paper is the differential effect between an eligible *grandmother* vs. an eligible *grandfather*. Not clear whether households with eligible older women or men are systematically different.

Duflo & Udry (2004). *Intrahousehold resource allocation in Côte d'Ivoire: Social norms, separate accounts and consumption choices.* WP.

- Another test of the unitary or at least collective model. The idea of the collective model (Browning-Chiappori 1998) is that members of a household engage in some bargaining, e.g. Nash bargaining, within the household, using pareto weights θ_i . That is, the outcome of the household is found by

$$\max_{c_h} \theta_m u_m(c_m, c_f, c_h) + (1 - \theta_m) u_f(c_m, c_f, c_h), \quad s.t. \quad p \cdot c \leq \sum_i y_i$$

Where c_h is 'household consumption' and everything else follows the obvious notation

- The key result is that – even without the (strict) unitary approach – this problem leads to pareto efficiency and that, conditional on total expenditure, any change in income shares go through changes in the bargaining weights. However, foreseeable shocks should not influence the pareto weights, but be insured away. Within the household, there is perfect insurance. Non-persistent observable shocks are fully insured away and should not induce any differences in resource allocation.
- DU test this idea with households in Côte d'Ivoire where men and women traditionally grow different kinds of crops that react differently to various levels of rainfall. In particular, they distinguish between:
 - Crops grown by females
 - Crops grown by males
 - Yam (grown by males), the income from which – according to strict social norms – should be used for the household good
- They set up the household problem which is essentially separable so that households solve

$$\max_{c_i} E u_f(c_f) + \lambda E u_m(c_m), \quad s.t. \quad x \equiv p_m c_m + p_f c_f \leq \pi_f(r) + \pi_m(r)$$

- Separability here means that the optimal profit functions as an outcome of a maximisation over labour input, where r are rainfall shocks (not permanent income changes) can be found first, and then the household problems wrt. consumption can be solved. This gives consumption $c_i = c_i(\lambda, p, x)$ where rainfall in period j affects consumption of good k by individual i only through its influence on expenditure x . In particular, consumption is not affected via λ since risk is a non-persistent shock. (assume the effect on prices is zero here, but there is a discussion in the paper if that's not the case).

$$\frac{\partial c_i^k}{\partial r_j} = \frac{\partial c_i^k}{\partial x} \frac{\partial x}{\partial r_j} \rightarrow \frac{\frac{\partial c_i^k}{\partial r_1}}{\frac{\partial x}{\partial r_1}} = \frac{\frac{\partial c_i^k}{\partial r_2}}{\frac{\partial x}{\partial r_2}}$$

- This is the testable result: If rainfall in period 1 and 2 affects total expenditure in the same way, it should also affect consumption of good k by i to exactly the same extent. This is clearly not the case:

Total expenditure (1)	Food consumption (2)	Adult goods (3)	Clothing (4)	Prestige goods (5)	Education (6)	Staples (7)	Meat (8)	Vegetables (9)	Processed foods (10)	Purchased foods (11)	Food consumed at home (12)	
PANEL A												
OLS coefficients:												
Predicted change in male non-yam income	0.126 (0.049)	0.062 (0.054)	0.870 (0.425)	-0.164 (0.334)	0.683 (0.209)	-0.101 (0.128)	0.113 (0.072)	0.002 (0.126)	0.345 (0.210)	0.004 (0.139)	-0.029 (0.078)	0.098 (0.119)
Predicted change in yam income	0.207 (0.037)	0.227 (0.041)	-0.473 (0.320)	0.296 (0.252)	-0.272 (0.158)	0.320 (0.108)	0.345 (0.054)	0.135 (0.096)	0.023 (0.159)	0.122 (0.105)	0.087 (0.059)	0.444 (0.090)
Predicted change in female income	0.309 (0.056)	0.235 (0.061)	1.537 (0.490)	0.535 (0.382)	0.993 (0.239)	-0.098 (0.159)	0.193 (0.082)	0.492 (0.144)	0.995 (0.239)	0.474 (0.159)	0.412 (0.089)	0.313 (0.136)

- The independent variables are essentially the rainfall variations (they construct three particular dimensions of rainfall variation related to different sources of income) and the first column are then effectively estimates of the numerator above $\partial x / \partial r$ and the other columns estimators of the denominators $\partial c / \partial r$. Clearly, it matters considerably what sort of rainfall shock households experience for the consumption of goods. In particular:
 - Male income goes to tobacco and alcohol, but yam income does not
 - Female income goes to clothing, prestige goods and food
 - Yam income goes to education and food
- The authors argue that this is a rejection of the collective model, but would be in line with a limited commitment model within the household along the lines of Ligon et al. (2002)

Udry (1996). *Gender, Agricultural Production, and the Theory of the Household.* JPE.

- Udry (1996) looks at 150 households in **farming activity in Burkina Faso** over 4 years to analyse misallocation of resources between men and women within a household. He estimates:

$$Q_{htci} = X_{htci}\beta + \gamma G_{htci} + \lambda_{htc} + \epsilon_{htci}$$

Where Q is yield on the plot for crop c , household h , individual i in time t , X are plot characteristics and G is a gender dummy.

- Rationale: Within households, different factor inputs and outputs for a plot are only a function of the plot characteristics. This arises as a result of the production maximisation (for any crop k):

$$\max_{\{N_F^i, N_M^i\}_i} \sum_{i \in P^k} G^k(N_F^i, N_M^i, A^i), \quad s.t. \quad \sum_i N_F^i = N_{Fk}, \quad \sum_i N_M^i = N_{Mk}$$

where P^k is the set of plots i planted to crop k , N s are labour input, A_i the area of the plot capturing plot characteristics and the non-negativity constraints are omitted. Since this gives a constant $\partial G / \partial N_g$ for both men and women across i [in the interior solution], then, for any two plots i, j s.t. $A_i = A_j$, it follows that:

$$G^k(N_F^i, N_M^i, A^i) = G^k(N_F^j, N_M^j, A^j)$$

- Therefore, if controls in X are sufficient, gender should not make a difference, leading to the test on the above estimation equation $\gamma = 0$
- Finding: **women** have a yield that is almost **30% less** (highly sign.)
 - Assuming the same concave production function of capital for men and women , this can have two explanations:
 - Their technical efficiency is much worse. They are just much less able in farming than men
 - There is misallocation
- In further analyses, he finds:
 - Estimates of the **production function** give that it is indeed **not significantly different** between men and women except for a few aspects and that it displays decreasing marginal returns to scale (concavity)
 - **Women** get significantly **less fertiliser**
 - **Distance from home cannot explain difference** – estimating by the same distance still yields the same gender differences
- These findings indicate strongly that the lower yield is due to **misallocation**. Udry estimates the **potential production gain** through efficient allocation to be **6%**.

13.4 Firm Misallocation

- De Mel et al. (2008) who look at **return** to capital among **microentrepreneurs in Sri Lanka** not only find extremely high returns but also large **heterogeneity**:
 - **Women** have on average **zero return**
 - **Wealthy families** have very **low return**
 - Better **educated** people have **very high return**
 - Those with large **digit span** (proxy for intelligence) have **very high return**
- This is also very **suggestive of misallocation**: money/cash constraints seem to determine who has capital rather than efficiency.
 - NB: You can only conclude misallocation if you know the production functions are the same otherwise or are at least not able to explain these differences.
- Banerjee & Duflo (2005) have **similar findings**:
 - Variability in returns
 - Very high returns to some

Restuccia & Santaella-Llopis (2017). Land Misallocation and Productivity. NBER WP.

- RS use detailed micro-level data from Malawi (12000 HHs) and a simple model to estimate the degree of land misallocation with respect to farm productivity.
- They find that land size is hardly related to farm productivity (misallocation) and fixing this would imply the 3.6-fold output relative to the actual one
- They see a potential reason in the missing market for land in the context in Malawi by comparing farmers using marketed land with those not using marketed land – if one were to reduce the dispersion in marginal products among farmers with no marketed

land to the same level as among farms using marketed land, output would increase 2.6-fold

- Output is characterised by farmer productivity heterogeneity:

$$y_i = s_i \zeta_i k_i^{\theta_k} (q_i l_i)^{\theta_l}$$

- $(y_i, s_i, \zeta_i, k_i, q_i, l_i)$ are output, farmer productivity, rainfall shock, capital, land quality and land respectively. Measurements:
 - output, capital, land and rainfall shocks are directly taken from the data
 - The θ s are taken from US data: $\theta_k = 0.36, \theta_l = 0.18$ (capital and land shares)
 - Land quality is the predicted land productivity from a regression on various measures of land quality (elevation, slope, erosion, soil quality, etc)
 - In this way, farmer productivity s_i can be backed out
- A distribution of farmer productivity (in logs) very close to a normal can be found
- They report a number of main results (Table 5):
 - (a) The ratio of efficient output to actual output
 - (b) As Fig. 3a,c shows, there are two types of misallocation: (i) dispersion of factor inputs *within* the same productivity level and (ii) the same level of factor input across different productivity levels. Part (b) reports misallocation after partially (95%, 90%, 80%) or fully (0%) removing the within-productivity misallocation
 - (c) Part c looks at the potential output gain if redistribution within specific geographic areas is possible (e.g. for regions, efficient output in the numerator is the output if redistribution within regions is implemented)

Figure 3: Land Size, Capital, MPL and MPK: Actual and Efficient Allocations

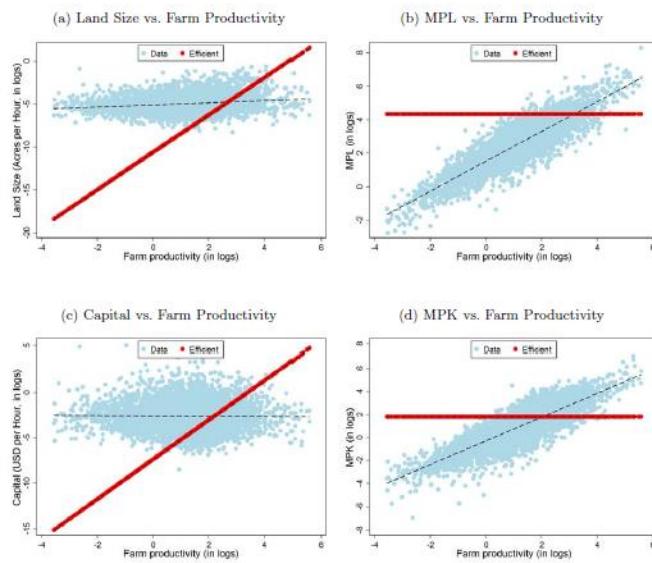


Table 5: Agricultural Output Gain (Y^e/Y^a)

(a) Main Results					
	Bootstrap Simulations				
Nationwide	Full Sample	Median	5th pct.	95th pct.	
	3.59	3.55	4.07	3.11	

(b) Within Productivity- s_i Variation					
	Benchmark	95%	90%	80%	0%
Output Gain	3.59	3.40	3.25	3.00	2.48

(c) By Geographical Areas and Institutions					
	Average	Median	Max	Min	
Geographic Areas:					
Regions	3.40	3.42	3.63	2.87	
Districts	2.93	2.80	7.10	2.07	
Enum. Areas	1.65	1.60	11.00	1.04	
Institutions:					
Traditional Authority	3.11	3.48	5.58	1.39	
Language	3.36	3.88	5.95	1.16	

Table 6: Land Markets—Output Gain for Farms with Marketed vs. Non-marketed Land

	By Marketed Land Share			By Marketed Land Type			
	No (%)	Yes (> 0%)	All (100%)	Rented Informal	Rented Formal	Purchased Unitit.	Titled
Output Gain	4.15	1.97	1.57	1.72	1.73	5.13	1.39
Observations	5,962	1,189	746	215	682	126	97
Sample (%)	83.4	16.6	10.4	3.0	9.5	1.8	1.3

Notes: The output gain is calculated as the ratio of efficient to actual output separately for subsamples of farm households defined by the share of different types of marketed land used. The share of marketed land is defined from the household-farm level information on how land was obtained, see Section 2. Each column refers to a particular subsample. The first column reports the output gain for the subsample of household farms that do not operate any marketed land. The second column refers to the subsample of household farms operating a strictly positive amount of marketed land, either purchased or rented-in. The third column refers to the subsample of household farms for which all their operated land is marketed land. The last four columns disaggregate the results by the main types of marketed land: (1) rented informally, i.e. land borrowed for free or moved in without permission; (2) rented formally, i.e. leaseholds, short-term rentals or farming as a tenant; (3) purchased without a title; and (4) purchased with a title. There is only 1% of households with marketed land whose type is missing in the Malawi ISA data.

13.5 Spatial Misallocation

Young (2013). Inequality, the Urban-Rural Gap, and Migration. QJE.

- **Rural-urban gaps** with regard to TFP, consumption and wages are commonly observed. This could be spatial misallocation but also reflect **selection of different skill levels**. Young (2013) argues so. The model is:

$$Q_i = A_i S_i^{\alpha_i} U S_i^{1-\alpha_i}$$

Where Q is production, S skilled labour and US unskilled labour, and $i \in \{U, R\}$ indicates urban and rural.

- The crucial assumption that drives the whole model is that $\alpha_U > \alpha_R$, i.e. that **urban production is skilled labour intensive**.
- $P(E)$ is the probability that someone is skilled depending on education E and the return to education is $R_E = \ln(w_s)P'(E)$ (from log utility). $P(E, i)$ is the probability of being skilled conditional on both education and location, interpretable as the probability of meeting someone with education E and in location i who is skilled. It can be shown that $P(E, U) > P(E, R)$ (comes essentially from $\alpha_U > \alpha_R$)
 - This is the selection story: Conditional on education, you are more likely to be skilled if you live in an urban area because skilled people have self-selected into living in urban areas where the demand for skilled jobs is higher.
 - Education is therefore not enough to control for skill level. The model goes one step further here.
- The UR Gap (in education equivalents⁶) is:

$$UR_{Gap} = \frac{\ln(w_s) [P(\mu_E, U) - P(\mu_E, R)]}{R_E}$$

⁶ That is, measured in years of education – essentially that's why the denominator is the returns to education. Young measures everything as education equivalents in the paper since average consumption data is quite noisy.

- The UR Gap can be shown to be (using some algebra and a 1st order Taylor approximation):

$$UR_{Gap} = \left[\frac{P(\bar{\mu})(1 - P(\bar{\mu}))}{P'(\bar{\mu})\bar{\Pi}(1 - \bar{\Pi})} (\Pi_S^U - \bar{\Pi}) \right] - \left[\frac{P(\bar{\mu})(1 - P(\bar{\mu}))}{P'(\bar{\mu})\bar{\Pi}(1 - \bar{\Pi})} (\Pi_{US}^U - \bar{\Pi}) \right] + 0 \times (\mu_E - \bar{\mu})$$

Where μ_E is the average education, $\bar{\mu}$ is the education level around which the Taylor approximation is computed, Π_j^i is the share of j ($\in \{S, US\}$) labour in i ($\in \{U, R\}$), and $\bar{\Pi} = \Pi_S^U = \Pi_{US}^U$ is the share around which the Taylor approximation is computed.

- This can be estimated as (c denotes the country):

$$UR_{Gap}(c) = \alpha + \beta_1(\Pi_S^U(c) - \bar{\Pi}) + \beta_2(\Pi_{US}^U(c) - \bar{\Pi}) + \beta_3(\mu_E(c) - \bar{\mu}) + \epsilon(c)$$

- Π_{16}^U , the share of those with 16 years of education in a city, is taken as a proxy for Π_S^U as $P(16) = 1$ according to Young and similarly Π_0^U is taken as a proxy for Π_{US}^U .
- Young finds indeed that he cannot reject $H_0: \beta_1 = -\beta_2, \alpha = \beta_3 = 0$ which confirms the rural-urban gap argument.
- Lessons:
 - The **share of skilled and unskilled workers in cities determines how large the rural-urban gap is** (high for high share of skilled workers in cities and low share of unskilled workers in cities)
 - This can **rationalise spatial variation in TFP**, wages and consumption **without misallocation**
- Movement cost** may be another explanation of such spatial variation (see Bryan et al., 2014)
- More general results:
 - The UR gap can explain much more of the variation in consumption than educational variation.
 - There is a similar share of people moving from rural to urban areas as moving from urban to rural areas – this is suggestive of sorting / selection.
 - The UR gap is *uncorrelated* with urbanisation rates or GDP. This is in contrast to Caselli (2005) who finds that the ratio of productivity in non-agriculture to productivity in agriculture decreases as GDP increases. Young argues that Caselli's results are largely a rich country phenomenon (where rich countries are absent from Young's sample). Young argues that – among poor economies – the relative productivity difference may reflect sorting based on skill in line with his model, that is, there is *no* misallocation.

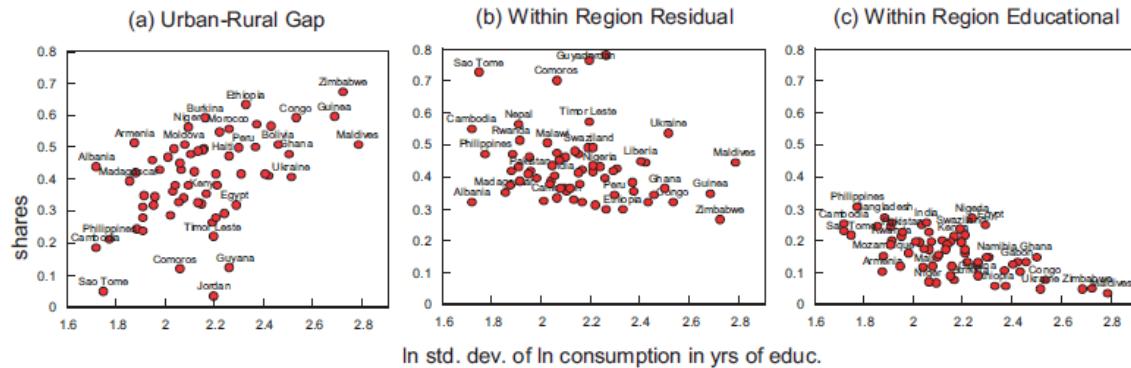


FIGURE V
Shares of Variance of Consumption (65 countries)

Bryan et al. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. Ecta.

- BCM analyse the **lack of seasonal migration** in **Rangpur, Bangladesh** during an annual **predictable and regional famine** (called Monga). It is puzzling that people would not (i) save (intertemporally arbitrage) or (ii) seasonally move for such famines (spatially arbitrage)
- Unlike Young (2013) and Lagakos-Waugh (2013), BCM consider moving cost.
- They conduct an experiment with some control village and **three treatments**: providing **cash, credit and information**.
 - 100 villages in 2 districts covered and most vulnerable households identified (<1/2 acre of land) → 1900 eligible households before the 2008 famine (Monga)
 - Cash and credit treatment also received information and the loan/credit amount of USD 8.50 was designed to cover transportation cost
 - Implicit limited liability on the loan means this may also serve as some kind of insurance (relevant if people don't invest in migrating because they may fail and incur huge utility losses relative to their consumption close to subsistence levels – Banerjee's "vulnerability view")
- Cash and credit (denote 'incentive') induced an additional 22% of households to migrate, information had no significant effect (denote info and control 'control'). Even one and three years later, there is a significant and considerable effect of the incentive treatments
- Three main results come out of the study:
 - (i) Migration is very profitable and under-utilised.
 - The **short-run and long-run migration** rate is significantly **higher** for the **cash and credit** treatment.
 - Migration has a **significant** and substantial effect on **consumption**. Assuming gains are realised over two months (duration of famine) and no disutility from separation, they estimate the **gross return to be 273%** (!)
 - (ii) Risk, income close to subsistence and learning about returns are key elements in explaining why there is such low utilisation

- (iii) There is a missing element in the model since the above components cannot fully explain the migration choices. **Risk aversion** at reasonable levels **cannot explain** the effects.
- This all is evidence for **misallocation** prior to the intervention.
- **External validity** is questionable here because of the very specific setting of the region affected by the Monga famine.
- BCM set up a model to try and rationalise their findings: There are good and bad migrants. Good migrants benefit from migration, bad migrants don't. Before having migrated, households don't know which type they are. But households can migrate and find out.
- Households engage in buffer stock saving (after passing a subsistence threshold in consumption, they start saving to smooth consumption in the future). In addition to buffering, households will at some level of 'cash on hand' x start to save up for migration.
 - Depending on their initial wealth, they may be stuck in a poverty trap and never (in a reasonable lifetime) be able to save up enough for migration.
- The incentive can clearly get them to migrate by providing a cash on hand shock.
- Calibration and simulation gives that
 - The findings are only rationalisable at plausible levels of risk aversion if households don't save up and are not forward looking (don't understand the ongoing value of migration)
 - If these elements of savings and forward looking are incorporated, unreasonably high levels of risk aversion would need to be assumed to get the results.

Bazzi et al. (2004). Skill transferability, migration and development: Evidence from population resettlement in Indonesia. WP.

See also referee report in EC428-MT

- Bazzi et al. (2014) use a **grand population resettlement programme in Indonesia** as quasi-experiment to assess **skill transferability** which, if not perfect, may be another reason for spatial variability.
 - People ("transmigrants") could choose to participate in the programme and be relocated.
 - For different reasons is the relocation process credibly exogenous
 - They find that **one s.d. increase in agroclimatic similarity leads to 20% productivity increase** and a **concave function** of skill transferability
- **Concerns** with their identification strategy:
 - **Participation is voluntary:** Those who chose to move were likely poorer and have fewer skills. This changes at least the **levels effect** of skill transferability, the intercept would be lower. This already is compromises on **external validity**. It might even change the **marginal effect** if the adaptability of these people is also systematically different.
 - Individual FE are included, but there may be **time-varying individual characteristics**. Those with low agroclimatic similarity might have made an extra effort or be frustrated and make very little effort to adapt; at least in the short run.
 - This is only a **rural-to-rural migration** programme and looks at agriculture only which again has impact on the external validity

13.6 Sectoral Misallocation

- Strange set of facts:
 - Lots of people in developing countries **work in agriculture**
 - **Agricultural productivity** is really **low** in developing countries relative to developed countries
 - Agricultural productivity is fairly **low relative to productivity in manufacturing**, developing countries (productivity in manufacturing is actually not that bad in developing as compared to developed countries).
- This hints at **misallocation** going on as well.

Lagakos & Waugh (2013). Selection, agriculture, and cross-country productivity differences. AER.

- LW try to explain the gap in productivity between agriculture and non-agriculture by selection (similar idea as Young (2013) essentially, just explicitly with sectors instead of urban-rural distinction). The basic idea is simple and hinges on two key assumptions:
 - A subsistence consumption requirement of the agricultural good \bar{a}
 - Non-homothetic preferences
- Suppose utility and production (in an agric. and non-agric. sector) take the following form:

$$U^i = \log(c_a^i - \bar{a}) + \eta \log(c_n^i)$$

$$Y_a = AL_a, \quad Y_n = AL_n$$

- Note that economy-wide TFP is the *same* in both sectors, but let workers have idiosyncratic skill types for both sectors (z_a^i, z_n^i). If A is low in one country, by the subsistence requirement, many people have to work in agriculture. Therefore, the average individual productivity in this sector \bar{z}_a is low since people select into sectors based on their idiosyncratic productivities and the few with highest z_n^i will be in the non-agricultural sector while many will be in the agricultural sector. This also means wages in the agricultural sector will be low relative to the non-agric. sector
 - This can essentially explain the main finding that poor countries (with low A) have (i) more people working in agriculture and (ii) a much higher ratio of non-agricultural to agric. productivity than richer countries
- They do quantitative theory. Based on some distributional assumptions on (z_a^i, z_n^i) , they match moments in US data to find the parameters. For cross-country comparison, they take $A = 1$ for the US and lower A so long as to arrive at another country's GDP p.c. They can then look at what the model would predict for these countries. Main results include:
 - 90th to 10th percentile (GDP p.c. distribution) comparison of productivity difference in agric. vs. non-agric.
 - Wage gap w_a/w_n
 - Share of employment in agric.

TABLE 2—90–10 PRODUCTIVITY DIFFERENCES, DATA AND BENCHMARK MODEL

	Agriculture	Aggregate	Non-agriculture	Ag/non-agriculture ratio
Data	45	22	4	10.7
Model	29	22	13	2.2
Without selection	19	19	19	1.0

Notes: The aggregate productivity difference is the ratio of GDP per worker between the ninetieth and tenth percentile countries. Sector productivity differences are the ratio of sector output per worker in the ninetieth and tenth percentile countries. The ag/non-agriculture ratios are the agriculture productivity differences divided by the non-agriculture productivity differences.

Source: Authors' calculations and Caselli (2005).

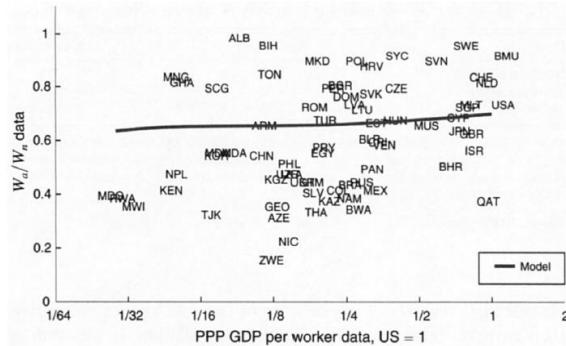


FIGURE 1. AVERAGE WAGE IN AGRICULTURE RELATIVE TO NON-AGRICULTURE, DATA, AND MODEL

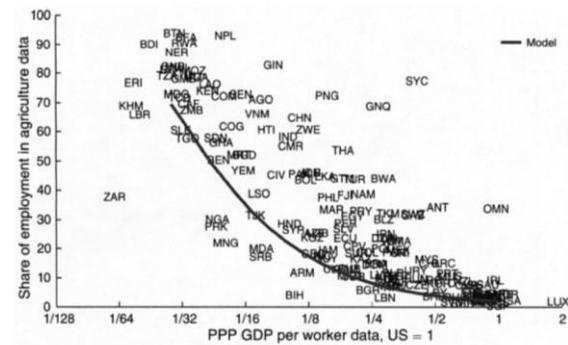


FIGURE 2. SHARE OF EMPLOYMENT IN AGRICULTURE, DATA, AND MODEL

14 Management

Integrated in the Section on “Firms, Markets and Entrepreneurship”

15 Human Capital / Education

15.1 General Summary / Remarks

- Classical approach following **Mincer**: $h = e^{\phi s}$ and $\ln w = \ln \beta_0 + \phi s$ where w is the wage and s is schooling
- In regressions of wage on schooling, however, we cannot capture the productivity increase through (mere) schooling by ϕ as this captures the **OMV ability** too.
- Different market failures may justify intervening in the market for (private) schooling to the extent that schooling may even be entirely publicly provided:
 - Parents choose for their children and do not perfectly reflect their children's preferences
 - Externalities in production with human capital as an input
 - Lack of credit market access
 - Incorrectly held beliefs about the marginal benefit of education
 - Difficulty of assessing educational quality
 - Peer effects as externality

- Local (natural) monopolies: there is a fixed cost of setting up a school and people usually want to be able to walk to school. Within a school's reach therefore, there is no incentive to set up another school
- Many the studied interventions are partial equilibrium analyses: what happens to education / returns if programme X is implemented? The results may vastly differ from the introduction of a nation-wide policy where GE effects play a role.

15.2 Returns to Education

Duflo (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. AER

- Duflo (2001) uses **INPRES**, an Indonesian government programme to build more than 60,000 primary schools between 1973 and 1979 as a **natural experiment**.
- Schools were not built in random places, but regions with **low enrolment rates targeted**
- Duflo uses essentially a DID design with two sources of variation:
 - School age in 1974
 - Programme intensity in different regions
- Regression of outcome (schooling, wage) on birth cohort and region fixed effect as well as an **interaction** between the **intensity in a region** and the **age of people in 1974** as coefficients in a regression with years of schooling as dependent variable. She thus uses the fact that only people aged 2-6 in 1974 benefitted from the programme as opposed to (slightly) older children in the same region. The coefficient is shown in the graph.
- Identifying assumption: Parallel trend between primary school aged and (slightly) older cohorts. Placebos run, s. also convincing graph for non-primary school aged children
- Finding: the coefficient for the interaction is **significant for all ages <10** in the schooling regression, 1 school per 1000 increases education by 0.12 to 0.19 years. The **increase in wages is 3-5.4%**
- Duflo also estimates the **returns to education** directly, instrumenting for schooling using (Birth dummies x programme intensity in region of birth) as IV. This gives returns estimates on the order of **6.8-10.6%**.
- For all estimates, the results are quite heterogeneous:

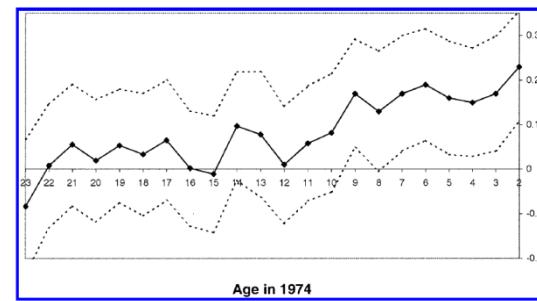


FIGURE 1. COEFFICIENTS OF THE INTERACTIONS AGE IN 1974 * PROGRAM INTENSITY IN THE REGION OF BIRTH IN THE EDUCATION EQUATION

TABLE 6—PROGRAM EFFECT AND RETURNS TO EDUCATION BY CATEGORIES OF REGION OF BIRTH

	Whole sample	Characteristics of region of birth						
		Density ^a		1976 Poverty ^b		Preprogram education ^c		
		<Median (1)	>Median (2)	High (4)	Low (5)	<Median (6)	>Median (7)	
<i>Panel A: Effect of the Program on Education</i>								
Dependent variable: Years of education.								
Sample: individuals ages 2 to 6 or 12 to 17 in 1974								
Interaction (2–6 in 1974)*program intensity in region of		0.15 (0.026)	0.19 (0.035)	-0.014 (0.048)	0.13 (0.058)	0.083 (0.035)	0.14 (0.040)	
<i>Panel B: Effect of the Program on Wages</i>								
Dependent variable: log(hourly wage). Sample:								
individuals ages 2 to 6 or 12 to 17 in 1974 (wage earners)								
Interaction (2–6 in 1974)*program intensity in region of		0.017 (0.0074)	0.032 (0.011)	-0.00084 (0.012)	0.051 (0.017)	-0.00083 (0.0094)	0.028 (0.013)	
<i>Panel C: Returns to Education</i>								
Dependent variable: log(hourly wage). Sample:								
wage earners								
Years of education		0.078 (0.00062)	0.11 (0.026)	No First stage	0.10 (0.028)	No First stage	0.12 (0.032)	
		[0.9]	[0.86]		[0.88]		[0.72]	
							[0.83]	

- External validity issues:
 - Emphasis on education at the time provided convenient context – higher returns than usual
 - GE effects since large programme: individuals' returns lower in later environment where many more people are educated in general than previously.
 - 2SLS as weighted ATE for compliers issue
 - Selection of compliers: on the one hand, affected children probably those who were previously prevented from going to school by lacking infrastructure and their poverty; on the other hand, those who have high expected returns choose to attend school once new schools are built
- Cost-benefit analysis somewhat heroic (extensions up to 2050 necessary including GDP growth assumptions, strong assumptions), but give large returns (internal rate of return between 8.8 and 12%) and show that it takes a very long time until positive returns are generated (takes a generation for school effects to show in wages, so this is not surprising)

Duflo (2004). *The medium run effects of educational expansion: evidence from a large school construction program in Indonesia. JDE*

Follow-up on Duflo (2001) paper

- Model and estimate GE/ macro effects of the INPRES programme.
- One of the key investigations is to use the INPRES programme as an instrument for education *average* increases and estimate the effect of that on wages.
- She uses labor force surveys from 1986 to 1999. The INPRES programme was carried out between 1973-79
- Essentially, Duflo employs a DID approach using the following variations:
 - Variation in regions by programme intensity

- Variation in cohorts: cohort '20-40' are those born *after* 1962 who were affected by INPRES and in working age during 1986-99. The 'older cohort' are those born *before* 1962 and *under* the age of 60 (in 1999) who were not affected by INPRES but still working age during 1986-99.
- First stage shows a positive effect on the share of primary school graduates in the 20-40 cohort relative to the older cohort – the share is increasing over time.
- Duflo uses the regional variation in programme intensity as an instrument for regional variation in education averages and estimates the effect of a human capital accumulation rate change on wage rate changes, holding skill constant. She essentially runs a differenced version of the following estimation equation:

$$\overline{\ln(w_{ijt})} = \overline{S_{jto}} b_{jt} + S_{jt} \alpha_U \epsilon_{jt} + \mu_t + v_j + \overline{v_{ijt}}$$

Where i, j, t denote individuals, districts, time, S is the share of the younger cohort or older cohort (with index o); programme intensity is used as an instrument for S_{jt} , the share of primary school graduates in the 20-40 cohort. (more details on the quite technical estimation strategy in the paper)

- She finds that higher human capital leads to lower wage (however, insignificant!) and higher formal employment. Consistent with a dual-economy model which an increase in productivity of the labour force is entirely absorbed by the formal sector.
- Capital accumulation in this sector is not fully responsive to the human capital increases. Why did the capital stock not adjust (to the extent it should have)?
 - Myopic investment behaviour, very large adjustment cost, financing constraints proposed as reasons in conclusion
- Relevance: this shows that an increased rate of human capital accumulation does not necessarily mean more growth.

2SLS estimates of the impact of average education on individual wages		
	Independent variable: % of primary school graduates in the 20–40 sample	
	Sample: rural and urban areas	Sample: rural areas only
	(1)	(2)
<i>Panel A: years 1986–1999</i>		in
Log (wage)	−0.204 (0.443)	−0.834 (0.701)
Log (wage) residual	−0.292 (0.355)	−0.633 (0.431)
Skill premium	−0.434 (0.916)	−0.982 (1.408)
Formal employment	0.441 (0.159)	0.454 (0.203)
Formal employment among educated workers	0.432 (0.197)	0.501 (0.259)
Formal employment among uneducated workers	0.379 (0.203)	0.409 (0.232)

Wantchekon et al. (2015). *Education and Human Capital Externalities: Evidence from Colonial Benin*. QJE

- Collected data from four villages where first schools in Benin were built in the early 20th century by missionaries/ colonialists. This was before any other colonial institutions were set up in Benin. They take villages within a 7-20km radius as controls arguing that the choice of the specific village to set a school up was as good as random relative to these controls. There was no information to select on between such villages.
 - Villages within <7km not considered because of spillovers / might have attracted students as well; villages >20km not considered because they might be too different geographically / ethnically. No significant differences between treatment and control villages detected on observables.

- Three groups created: individuals living near the school and enrolling (TG1), those living close but not enrolling (TG2) and controls (C).
- Children to attend school chosen at a time when there was hardly any information on the effects; if anything, anecdotal evidence suggests negative selection which would go against results.
- Data from school, church and family archives and random sampling within the villages
- They essentially try to answer three different questions with regard to education:
 - (i) what is the effect of education on my own living outcomes? [obviously this is a second-stage question to the trivial one whether building a school has a first-stage effect on education – the answer is yes]
 - (ii) what is the (externality) effect of education on non-educated within my village (where a school was built and some are educated)?
 - (iii) what is the intra-family (externality?) effect of educated family members in the higher generation (parents, uncles, etc)?
 - (iv) is there catching-up in the sense that people in the *second* generation with *non-educated* parents within a treatment village (where *other* parents are educated) are more likely to go to school?
- Results:
 - (i) Large positive effect on own measures of living standard (prob. of being a farmer, having access to water and electricity, having a means of transportation and an index are outcomes), measures of my social network and measures of my political participation
 - (ii) hardly any externality effects on others, except for speaking the French language and social networks, as well as being a candidate in an election (the latter may clearly be a matter of free-riding on another (educated) person from the village already running)

FIRST-GENERATION LIVING STANDARDS EFFECTS					FIRST-GENERATION SOCIAL NETWORKS EFFECTS				
	(1) Farmer	(2) Water	(3) Electricity	(4) Means of transportation	(5) Living standards	(1) French language	(2) White friends	(3) Social networks scale	(4) Social networks scale
Individual-level treatment	-0.641*** (0.095)	0.112*** (0.041)	0.077*** (0.012)	0.294*** (0.025)	0.872*** (0.171)	0.870*** (0.033)	0.373*** (0.016)	2.010*** (0.217)	1.999*** (0.228)
Village-level treatment	-0.060 (0.116)	0.055 (0.048)	0.018* (0.010)	-0.012 (0.018)	-0.004 (0.164)	0.072*** (0.024)	0.049 (0.039)	0.100*** (0.038)	
Observations	291	406	406	388	379	Distance from school			-1.102*** (0.382)

FIRST-GENERATION POLITICAL PARTICIPATION EFFECTS			
	(1) Campaign for party	(2) Member of party	(3) Candidate in election
Individual-level treatment	0.339*** (0.053)	0.317*** (0.047)	0.117*** (0.036)
Village-level treatment	0.045 (0.046)	0.057 (0.061)	-0.021*** (0.007)
Observations	365	362	373

- (iii) large intra-family effect of educated parents on my own education (and subsequent effect on living standard outcomes)
- (iv) large catching-up effect of educated parents of *others* on the education of people with *non-educated* parents within a treatment village (and subsequent effect on living standard outcomes)

SECOND-GENERATION EDUCATION EFFECTS				
	(1) Education	(2) Primary or more	(3) Secondary or more	(4) University
Individual-level treatment	0.374*** (0.087)	0.144*** (0.042)	0.162*** (0.038)	0.067*** (0.022)
Village-level treatment	0.566*** (0.065)	0.345*** (0.035)	0.163*** (0.029)	0.058*** (0.015)
Observations	1,898	1,898	1,898	1,898
L=D F-stat	2.104	9.335	0.000067	0.079
L=D p-value	.148	.002	.993	.779

SECOND-GENERATION LIVING STANDARDS EFFECTS						
	(1) Farmer	(2) Water	(3) Electricity	(4) TV	(5) Phone	(6) Means of transport
Individual-level treatment	-0.061* (0.034)	0.118*** (0.040)	0.142*** (0.050)	0.194*** (0.047)	0.211*** (0.050)	0.135*** (0.038)
Village-level treatment	-0.299*** (0.035)	0.054 (0.033)	0.426*** (0.040)	0.312*** (0.034)	0.228*** (0.035)	0.011 (0.033)
Observations	1,791	1,924	1,924	1,924	1,924	1,894
L=D F-stat	16.169	0.990	13.039	2.829	0.053	4.283
L=D p-value	.000	.321	.000	.094	.818	.039
						.182

Jensen (2010). *The (perceived) returns to education and the demand for schooling.* QJE.

- Motivational idea: students, in particular secondary school students in less developed countries, or their parents take education decisions based on perceived – not actual – returns which may be very inaccurate
- Jensen first reports evidence on the fact that perceived return < actual return in the Dominican Republic (unlike for college students in the US according to a number of studies)
 - Return to secondary schooling (relative to only primary schooling) is factually 1300 DOP and estimated by students to be 329
 - Returns to tertiary schooling (relative to secondary schooling) are factually 5200 DOP and estimated by students to be 1282 DOP.
- Then students are at randomly selected schools in the Dominican Republic in the eighth grade (last grade of compulsory schooling) were provided with information on the returns to schooling
- Main results:
 - First-stage increase in beliefs about return to secondary schooling. They perform a DID looking at the difference between T and C prior to and after the intervention (obviously, randomisation means that pre-intervention difference should be close to zero and it is). The secondary schooling return belief increases by 366 DOP in T relative to C
 - Second stage: increase by 0.20 more years of schooling in treatment relative to control group over four years. The prob. of returning to school next year increases by 4% and the prob. of completing secondary school does not show a significant effect (point estimate: 0.02)

	Panel B. Schooling			Round 3			
	Control	Treatment	Difference	Control	Treatment	Difference	
Returned to school?	0.55 (0.02)	0.59 (0.02)	0.042 [*] (0.025)		0.30 (0.02)	0.32 (0.02)	0.020 (0.024)
Completed secondary school?				9.75 (0.070)	9.93 (0.073)	0.18 [*] (0.098)	
Years of schooling completed							
Number of observations	1,118	1,123	2,241	1,033	1,041	2,074	

- heterogeneous effects: increase by 0.33 years for least poor and no effect for poorest students, consistent with the idea that poverty and credit constraints limit schooling demand

15.3 Private Schooling

Muralidharan & Sundararaman (2015). *The aggregate effect of school choice: Evidence from a two-stage experiment in India.* QJE

- Two-stage randomisation of school vouchers: Village-level and – within treatment villages – student-level. The voucher programme admitted access to private schools. The randomisation structure is used to assess both the individual effect of attending a private school and the aggregate effect of students in private schools (including spillovers)
- Setting in Andhra Pradesh, India. There was *voluntary* participation by private schools, but they were not allowed to selectively reject or accept voucher-winning students
 - Still, only low quality private schools would participate?
- The design of their experiment allows them to look at different margins. In particular, they have four different groups in both treatment and control villages
 - Non-applicants in public schools (in T and C villages)
 - Applicants in public schools who are not awarded a voucher (in T villages and all applicants in C villages)
 - Applicants in public schools who are awarded a voucher (obviously only in T villages)
 - Non-voucher students in private schools (in T and C villages).
- They use this to look at :
 - The individual student-level private school effect
 - The spillover community effect
 - The aggregate effect of treatment relative to control villages.
- Background characteristics difference:
 - private schools hire more teachers, pay substantially lower wages, have less qualified teachers, longer school days, longer school years, less teacher absence, lower rates of multigrade teaching, higher rates of actually teaching
 - Students attending private schools come from significantly better socio-economic background
 - However, award-winning students show no significant difference in time spent doing homework and household spending. → Any difference found rather due to changes in school rather than household factors
- No effect found on test scores in the local language Telugu, math, English and science and social science (for English, actually small effect at 10% sign.); but private schools have significantly less instructional time on Telugu and math, so they seem to be more

efficient. Private schools additionally teach Hindi which is not taught in public schools, and there is (naturally therefore) a substantial effect on test scores in Hindi. No joint significance if Hindi is excluded

	TEST SCORE IMPACTS										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Year 2 assessments			Year 4 assessments							
	Telugu score	Math score	English score	Combined across tests	Telugu score	Math score	English score	EVS score	Combined across tests excluding Hindi	Hindi score	Combined across tests
Panel A: Impact of winning a voucher (intention to treat effects)											
Offered voucher	-0.079 (0.055)	-0.053 (0.065)	0.185** (0.079)	0.016 (0.061)	-0.017 (0.051)	-0.031 (0.052)	0.116* (0.070)	0.083 (0.060)	0.036 (0.048)	0.545*** (0.068)	0.133*** (0.045)
Total observations	4,620	4,620	4,525	13,765	4,385	4,385	4,217	4,243	17,230	1,696	18,926
Treatment observations	1,778	1,778	1,738	5,294	1,674	1,675	1,607	1,628	6,584	867	7,451
Control observations	2,842	2,842	2,787	8,471	2,711	2,710	2,610	2,615	10,646	829	11,475
Panel B: Average treatment on the treated (ATT) effect of attending a private school (scaling up intention to treat effect by inverse of voucher take-up rate)											
Voucher recipient in private school	-0.156 (0.108)	-0.104 (0.128)	0.364** (0.156)	0.032 (0.120)	-0.033 (0.100)	-0.061 (0.102)	0.229* (0.138)	0.164 (0.118)	0.071 (0.095)	1.074*** (0.134)	0.262*** (0.089)
Total observations	4,620	4,620	4,525	13,765	4,385	4,385	4,217	4,243	17,230	1,696	18,926
Voucher recipients	997	997	982	5,294	945	946	911	920	6,584	510	7,451
Nonrecipients	3,623	3,623	3,543	8,471	3,440	3,439	3,306	3,323	10,646	1,186	11,475

SCHOOL TIME USE: INSTRUCTIONAL TIME BY SUBJECT (MINUTES PER WEEK)

	(1) Private schools	(2) Public schools	(3) Difference
Telugu	307.72 (6.36)	511.52 (3.60)	-203.81*** (6.99)
Math	339.75 (7.50)	500.69 (3.36)	-160.94*** (8.63)
English	322.68 (7.96)	235.52 (5.39)	87.17*** (9.69)
Social studies	239.21 (6.29)	173.24 (6.89)	65.96*** (9.84)
Science	205.52 (9.09)	104.58 (5.78)	100.94*** (9.44)
Hindi	215.78 (6.08)	0.01 (0.89)	215.77*** (6.41)
Moral science	16.85 (4.82)	20.11 (3.20)	-3.26 (5.56)
Computer use	46.7 (6.50)	0.51 (1.02)	46.19*** (6.80)
Other	311.66 (14.55)	250.29 (6.70)	61.37*** (16.20)
Total instructional time	2,005.87 (13.73)	1,796.47 (6.86)	209.4*** (14.46)
Break	461 (9.14)	473.18 (3.05)	-12.18 (10.58)
Total school time	2,466.87 (17.46)	2,269.65 (8.25)	197.22*** (19.79)
Observations	325	200	

- Cost per student is much lower (a third) in private schools, again speaking in favour of the efficiency argument.
- No evidence of any spillover effects, in particular no adverse effects found on three dimensions: (i) students starting out in a private school, or (ii) voucher losers, or (iii) non-voucher applying students

Angrist et al. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. AER.

- They use the fact that Colombia used voucher lotteries for private schooling that covered *half the cost* (not entirely free) to evaluate the causal effect of private schooling. The programme is called PACES.

- Short-run effect: higher test scores by 0.2 standard deviations
 - Only 60% took the test though! This was a test that students were asked to take on top of their normal exams. Quite likely there is a selection into who took the test
- The effect of having ever used a private-school scholarship on the prob. of finishing 8th grade, where the PACES is used as an instrument for the regressor, is 15 ppt (on a base of 63%)
- Caveats:
 - Private schools could still select students → most able students pick up which show greatest effect
 - Some private schools opted out
 - Students were required to maintain academic performance to keep the voucher → this conditionality could drive the result
 - Only half the cost covered → most benefitting students took up and estimates are an upper bound?
- Essentially, such a lottery clearly addresses a potential credit constraint, but may not address other externalities
- It is quite unclear what is driving the results here
 - Parents teaching better?
 - Peer effects of better students within private schools?
 - Incentive from conditionality of the voucher?
 - Selection into taking the test?
 - Education expenditure increase by parents through the voucher?

Hsieh & Urquiola (2006). *The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program.* J Pub Econ

- Chile introduced a school voucher programme in 1981 by which everyone could get a voucher for private school
- 10000 new private schools were built since then and private school enrolment increased to 40%. Private schools were able to turn down applicants! Intended/expected effect may be twofold:
 - Increased incentive for performance improvement in *all* schools through more competition
 - If private schools do indeed provide better education, higher test should be a result of students moving to private schools
- Resulting from these two effects, the outcomes for public schools may change through
 - A change in their teaching
 - A different composition of students in public schools
- In a similar spirit as Duflo, they compare areas which experience a faster growth of private schools (typically urban) with those that experience slower growth and observe whether outcomes *change* differently over time in the former areas relative to the latter. This works as long as endogenous factors driving the faster growth are constant over time. Controls for possible trends seem to validate this assumption.

- Result: average test scores did *not* rise faster in communities which experienced quicker private school growth and other performance measures (average repetition, grade-for-age measures) even *worsened*.
 - The result of slight worsening (if anything) is robust to an international comparison. Over the period between 1970 and 1999 (during which the programme was introduced), Chile's performance in test scores relative to 12 other countries did not change. Controlling for GDP p. worker, enrolment and the ratio of spending per student to GDP p.c., Chile even got slightly worse.
- Interpretation: This and other results point towards sorting: Households/ parents use the vouchers to choose the school with the best peer group for their children. Schools have an incentive to be very selective and provide the best peer group. Therefore, we will have separation/ stratification: Good students in schools with good peer groups will have good results (without the teaching quality being necessarily better there; parents just mistake selection for causality and try to send their child to private school) and bad students will be in bad schools with bad peers.
 - Table 5 presents the other piece of evidence. The fact that public schools perform relatively worse to the average of all schools as private enrolment goes up can have two roots: (i) sorting, s. above. (ii) private schools providing better education.
 - The fact that *overall* performance does not improve in addition to the evidence in the below table suggests that sorting is at place.
 - Panel B with changes is done to control for regional FE (more private schools going in areas where public schools are weakest)
 - The overall average effect is negative because apparently the negative effect in bad schools dominates the positive one in good schools.

Table 5
Sorting among communes, 1990's cross-section and 1982–1988 changes

	Dependent variable—within commune observations of average characteristic in public schools/average characteristic in all schools									
	SES index ^a		Income ^b		Language ^a		Mathematics ^a		Repetition ^c	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A—1990's cross sections^d										
Private enrollment ^e	−0.20*** (0.02) [−0.58]	−0.16*** (0.03) [−0.46]	−0.37*** (0.07) [−0.43]	−0.33*** (0.09) [−0.38]	−0.08*** (0.02) [−0.39]	−0.08*** (0.02) [−0.39]	−0.09*** (0.02) [−0.42]	−0.09*** (0.03) [−0.42]	0.42*** (0.07) [0.44]	0.28*** (0.07) [0.29]
Commune controls ^e	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Thirteen regional dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	296	296	184	184	296	296	296	296	299	299
R ²	0.313	0.493	0.171	0.285	0.188	0.396	0.215	0.346	0.193	0.447
Panel B—1982–1988 changes										
Change in private enrollment ^e					−0.21** (0.10) [−0.24]	−0.22** (0.10) [−0.26]	−0.14* (0.08) [−0.17]	−0.19** (0.08) [−0.23]	0.51** (0.24) [0.24]	0.38* (0.24) [0.18]
Controls: concurrent trends ^f					No	Yes	No	Yes	No	Yes
N	84	84	84	84	84	84	84	84	163	163
R ²	0.060	0.065	0.027	0.097	0.027	0.097	0.054	0.054	0.100	

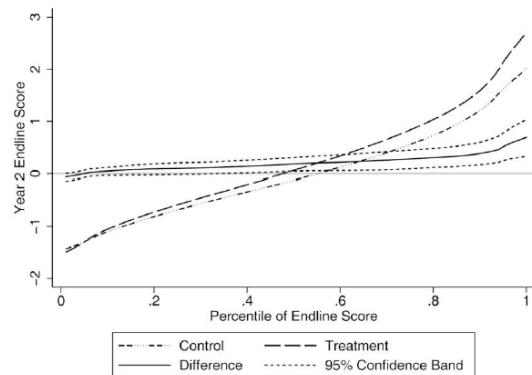
15.4 Incentives in Schools

Muralidharan & Sundararaman (2011). Teacher performance pay: experimental evidence from India. JPE.

- Address the issue of principal-agent issues within public schools (if giving up on the market, the government has to think about moral hazard issues). Conflict of interest: Teacher may have different incentives from parents or the government

- They run an RCT with government-run primary school in Andhra Pradesh, India.
- RCT with three arms in 300 schools (100 each)
 - Individual teacher incentives
 - School group-level incentives
 - Comparison group
- Potential problems with performance pay:
 - Are test scores the desirable result?
 - Does performance pay crowd out social motivation? (some evidence to the contrary, see section on motivation of public servants)
 - Fairness and insurance for teachers – some years would pay lower just because students are worse
- Other research seems to show that experience or a master's degree in education is a very poor predictor of better student outcomes; incentive schemes might be more suitable.
- Students in incentive schools outperform those in control schools by 0.27 and 0.17 sd in math and language tests after two years respectively.
- The result is robust to splitting up by repeat vs. non-repeat question and multiple choice vs. non-multiple choice question.
- Incentives improve test scores at all quantiles
- Spillovers in science and social science as well where no incentives were provided

	IMPACT OF INCENTIVES ON STUDENT TEST SCORES			
	Dependent Variable: Normalized End-of-Year Test Score			
	YEAR 1 ON YEAR 0		YEAR 2 ON YEAR 0	
	(1)	(2)	(3)	(4)
A. Combined (Math and Language)				
Normalized lagged test score	.503*** (.013)	.498*** (.013)	.452** (.015)	.446*** (.015)
Incentive school	.149*** (.042)	.165*** (.042)	.219*** (.047)	.224*** (.048)
School and household controls	No	Yes	No	Yes
Observations	42,145	37,617	29,760	24,665
R ²	.31	.34	.24	.28
B. Math				
Normalized lagged test score	.492*** (.016)	.491*** (.016)	.414*** (.022)	.408*** (.022)
Incentive school	.180*** (.049)	.196*** (.049)	.273** (.055)	.280*** (.056)
School and household controls	No	Yes	No	Yes
Observations	20,946	18,700	14,797	12,255
R ²	.30	.33	.25	.28
C. Telugu (Language)				
Normalized lagged test score	.52*** (.014)	.510*** (.014)	.49*** (.014)	.481*** (.014)
Incentive school	.118*** (.040)	.134*** (.039)	.166*** (.045)	.168*** (.044)
School and household controls	No	Yes	No	Yes
Observations	21,199	18,917	14,963	12,410
R ²	.33	.36	.26	.30



IMPACT OF INCENTIVES ON NONINCENTIVE SUBJECTS
Dependent Variable: Normalized End Line Score

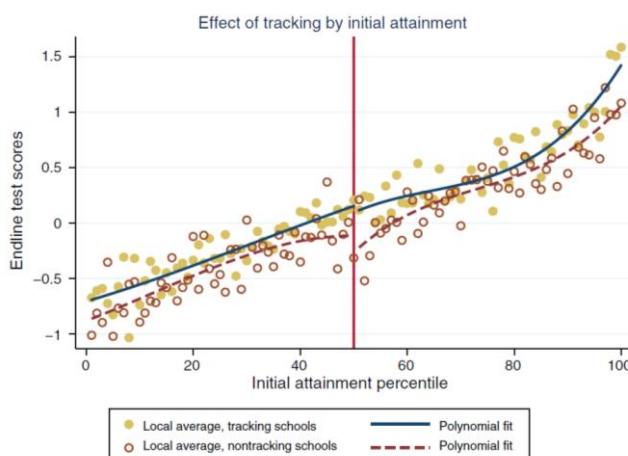
	YEAR 1		YEAR 2	
	Science	Social Studies	Science	Social Studies
A. Reduced-Form Impact				
Normalized baseline math score	.215*** (.019)	.224*** (.018)	.156*** (.023)	.167*** (.024)
Normalized baseline language score	.209*** (.019)	.289*** (.019)	.212*** (.023)	.189*** (.024)
Incentive school	.112** (.052)	.141*** (.048)	.113** (.044)	.18*** (.050)
Observations	11,786	11,786	9,143	9,143
R ²	.26	.31	.19	.18

- School-level and teacher-level incentives perform equally well in the first year, but the second seem to perform better after two years (0.28sd vs. 0.15sd higher than in the control group)
- Teacher activity measured in classroom observations and interviews
 - Classroom observations may have an effect? This may explain why there is no difference found in observable activities between incentive and control groups. Irrespective of the incentives, teachers may try and do their best when observed in the classroom by someone

Duflo et al. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. AER.

- In a setting with 121 primary schools with a single class in Kenya, new teachers were employed. In (randomly selected) half the schools, the class was split up into two new classes with random assignment of students. In the other half, students were sorted in either of the two new classes based on a pre-test (tracking).
- Possible mechanisms:
 - Peer effect: Higher-achieving class should benefit from better peers on average, lower achieving class should lose out from worse peers on average
 - Teacher tailoring: teachers can more precisely teach to the level of the class (less dispersion in your own level and the targeted level by the teacher). This should benefit both groups
 - ➔ Which effect dominates for the lower-achieving group?
- Tracked students do better in *both* classes (conditional on the initial score; in fact, along the whole distribution).

	Total score				Math score		Literacy score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Short-run effects (after 18 months in program)</i>								
(1) Tracking school	0.139 (0.078)*	0.176 (0.077)**	0.192 (0.093)**	0.182 (0.093)*	0.139 (0.073)*	0.156 (0.083)*	0.198 (0.108)*	0.166 (0.098)*
(2) In bottom half of initial distribution \times tracking school			-0.036 (0.07)		0.04 (0.07)		-0.091 (0.08)	
(3) In bottom quarter \times tracking school				-0.045 (0.08)		0.012 (0.09)		-0.083 (0.08)
(4) In second-to-bottom quarter \times tracking school				-0.013 (0.07)		0.026 (0.08)		-0.042 (0.07)
(5) In top quarter \times tracking school				0.027 (0.08)		-0.026 (0.07)		0.065 (0.08)
(6) Assigned to contract teacher	0.181 (0.038)***	0.18 (0.038)***	0.18 (0.038)***	0.16 (0.038)***	0.161 (0.037)***	0.16 (0.038)***	0.16 (0.038)***	0.16 (0.038)***
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,795	5,279	5,279	5,279	5,280	5,280	5,280	5,280
<i>Total effects on bottom half and bottom quarter</i>								
Coeff (Row 1) + Coeff (Row 2)			0.156		0.179		0.107	
Coeff (Row 1) + Coeff (Row 3)				0.137		0.168		0.083
F-test: total effect = 0			4.40	2.843	5.97	3.949	2.37	1.411
p-value (total effect for bottom = 0)			0.038	0.095	0.016	0.049	0.127	0.237
p-value (effect for top quarter = effect for bottom quarter)				0.507		0.701		0.209
<i>Panel B. Longer-run effects (a year after program ended)</i>								
(1) Tracking school	0.163 (0.069)**	0.178 (0.073)**	0.216 (0.079)***	0.235 (0.088)***	0.143 (0.064)**	0.168 (0.075)**	0.231 (0.089)**	0.241 (0.096)**
(2) In bottom half of initial distribution \times tracking school			-0.081 (0.06)		-0.027 (0.06)		-0.106 (0.06)	
(3) In bottom quarter \times tracking school				-0.117 (0.09)		-0.042 (0.10)		-0.152 (0.085)*
(4) In second-to-bottom quarter \times tracking school				-0.096 (0.07)		-0.073 (0.07)		-0.091 (0.07)
(5) In top quarter \times tracking school				-0.028 (0.07)		-0.04 (0.06)		-0.011 (0.08)
(6) Assigned to contract teacher	0.094 (0.032)***	0.094 (0.032)***	0.094 (0.032)***	0.061 (0.031)**	0.061 (0.031)**	0.102 (0.031)***	0.103 (0.031)***	
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,490	5,001	5,001	5,001	5,001	5,001	5,007	5,007
<i>Total effects on bottom half and bottom quarter</i>								
Coeff (Row 1) + Coeff (Row 2)			0.135		0.116		0.125	
Coeff (Row 1) + Coeff (Row 3)				0.118		0.126		0.089
p-value (total effect for bottom = 0)			0.091	0.229	0.122	0.216	0.117	0.319
p-value (effect for top quarter = effect for bottom quarter)				0.365		0.985		0.141



- They specifically estimate peer effects by estimating the effect of the average score of peers in a class on your score, where the average score of others is instrumented by the average initial score of others (to break the reflection problem)
 - (substantial) increase of 0.445sds on own test score found when average test score of others (instrumented) increases by 1 sd.
- They also find that teachers (both government teachers and those hired through the experimental programme) perform sign. better in terms of absenteeism and actual teaching in tracked schools

16 Health

16.1 General Summary / Remarks

- Distinguish:
 - Curative health care
 - Preventive health care: much more likely to have information problems on true benefits etc.
- Market failures (on the demand side):
 - Most prominently: Externalities (epidemiology)!
 - Lack of information
 - Credit market failures
 - Behavioural biases
 - Local (natural) monopoly
 - Lack of commitment
- Supply side issues:
 - Typical R&D fixed cost issue: R&D is very costly, but once a drug or vaccine is developed, copiers will drive the price to MC which means the developer incurs a loss
- Incentive options on the supply side
 - push incentives, e.g. government grants. May be due to corruption and lobby capture.
 - pull incentives, e.g. patents or advance market commitments – A government or organisation agrees to buy quantity X at price P (quite recent idea)
- Patent issues
 - consumer side: many beneficiaries may be too poor to afford the product. Governments would need to commit to a subsidy ex-ante.
 - producer side: not all benefits can be captured due to single world prices, limited patient time, opportunities to work around patients
- For returns, much of the literature considers economic returns. Naturally, health in and of itself enters the utility function. However, economic returns much easier to capture and measure (in terms of welfare)

16.2 Returns to Health

Miguel & Kremer (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. Ecta.

- Around $\frac{1}{4}$ of the world population (NB: 2004 statement) is infected by worms. This leads to anemia, malnutrition, pain. School children account for 85-90% of heavy infections.
- MK provide randomised⁷ deworming medicine at the school level (randomised phased in at three different points in time: group 1 – 1998/99, group 2 – 1999, group 3 – 2000)
 - Innovation: school level enables the authors to estimate overall programme effects including (cross-school) externalities as opposed to individual-level randomisation which the literature has carried out before
 - For the estimation of within-school externalities, they must still rely on non-experimental methods
- Deworming is a classic example where the social benefits from inhibiting a disease are large relative to the private benefits (large externalities) and the benefits emerge gradually while the cost is immediate (really bad for hyperbolic discounters)
 - This makes an estimation of externalities all the more relevant!
 - Cost is not only financial cost but also the fact that the deworming medicine can be painful and have unpleasant side effects
- Estimation equation:

$$Y_{ijt} = \alpha + \beta_1 T_{1it} + \beta_2 T_{2it} + X'_{ijt} \delta + \sum_d \gamma_d N_{dit}^T + \sum_d \phi_d N_{dit} + u_i + e_{ijt}$$

Where Y_{ijt} is the outcome for school i , student j , time t , T are treatment indicators for the first and second round, X are controls (to increase statistical precision), N is the number of primary school students within distance d of school i and N^T is the number of *treatment* primary school students within distance d of school i .

- Note that γ_d will provide estimates of cross-school externalities, *given that we control for N_{dit}* . The reason is that the general student density is itself correlated with worm burden (higher density = higher infection risk). If we did not control for the general density, the density of students in treatment schools clearly reflects the general density of students and we only capture an attenuated externality (externality effect + negative effect of density in general).
- The externality would then be $\sum_d (\gamma_d \overline{N_{dit}^T})$
- The effects on health go in the expected direction (positive treatment effect, positive cross-school externality, negative effect of general student density)

⁷ They used alphabetical randomisation. This is not considered high quality by Cochrane, the major Public Health meta-analysis provider and led to exclusion of the MK study in the review on deworming that generally included many studies finding much more muted results than MK. However, they usually don't study externalities. The debate / worms fight is ongoing...

TABLE VII
DEWORMING HEALTH EXTERNALITIES WITHIN AND ACROSS SCHOOLS, JANUARY TO MARCH 1999^a

	Any moderate-heavy helminth infection, 1999			Moderate-heavy schistosomiasis infection, 1999			Moderate-heavy geohelminth infection, 1999		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Indicator for Group 1 (1998 Treatment) School	-0.25*** (0.05)	-0.12* (0.07)	-0.09 (0.11)	-0.03 (0.03)	-0.02 (0.04)	-0.07 (0.06)	-0.20*** (0.04)	-0.11** (0.05)	-0.03 (0.09)
Group 1 pupils within 3 km (per 1000 pupils)	-0.26*** (0.09)	-0.26*** (0.09)	-0.11 (0.13)	-0.12*** (0.04)	-0.12*** (0.04)	-0.11** (0.05)	-0.12* (0.06)	-0.12* (0.07)	-0.01 (0.07)
Group 1 pupils within 3–6 km (per 1000 pupils)	-0.14** (0.06)	-0.13** (0.06)	-0.07 (0.14)	-0.18*** (0.03)	-0.18*** (0.03)	-0.27** (0.06)	0.04 (0.06)	0.04 (0.06)	0.16 (0.10)
Total pupils within 3 km (per 1000 pupils)	0.11*** (0.04)	0.11*** (0.04)	0.10** (0.04)	0.11*** (0.02)	0.11*** (0.02)	0.13*** (0.02)	0.03 (0.03)	0.04 (0.03)	0.02 (0.03)
Total pupils within 3–6 km (per 1000 pupils)	0.13** (0.06)	0.13** (0.06)	0.12* (0.07)	0.12*** (0.03)	0.12*** (0.03)	0.16*** (0.03)	0.04 (0.04)	0.04 (0.04)	0.01 (0.04)
Received first year of deworming treatment, when offered (1998 for Group 1, 1999 for Group 2)		-0.06* (0.03)			0.03** (0.02)			-0.04** (0.02)	
(Group 1 Indicator) * Received treatment, when offered		-0.14* (0.07)			-0.02 (0.04)			-0.10*** (0.04)	
(Group 1 Indicator) * Group 1 pupils within 3 km (per 1000 pupils)			-0.25* (0.14)			-0.04 (0.07)			-0.18* (0.08)
(Group 1 Indicator) * Group 1 pupils within 3–6 km (per 1000 pupils)			-0.09 (0.13)			0.11 (0.07)			-0.15 (0.10)
Grade indicators, school assistance controls, district exam score control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2328	2328	2328	2328	2328	2328	2328	2328	2328
Mean of dependent variable	0.41	0.41	0.41	0.16	0.16	0.16	0.32	0.32	0.32

- School absenteeism in treatment schools was reduced by a quarter relative to control schools and externalities across schools can be found as well. They find 0.14 more years of schooling per dewormed child while the treatment effect is 0.051, that is, the externality accounts for the major part of the total effect! (left table shows treatment effect and cross-school externality)
- To estimate within-school treatment externalities (right table), note that there is a selection of individuals within treatment schools as to whether they accept treatment or not. MK compare group 1 treated students in treatment schools to group 2 treated students in treatment school (for the ATT) and group 1 *untreated* students in treatment schools to group 2 *untreated* students in treatment schools (*before* the group 2 intervention took place, i.e. in 1998). This is supposed to deal with the selection whereby the identifying assumption is that group 1 and 2 untreated students follow the *same* selection when deciding not to be treated within a treatment school

TABLE IX
SCHOOL PARTICIPATION, DIRECT EFFECTS AND EXTERNALITIES^a
DEPENDENT VARIABLE: AVERAGE INDIVIDUAL SCHOOL PARTICIPATION, BY YEAR

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	IV-2SLS (7)
	May 98– March 99						
Moderate-heavy infection, early 1999							-0.028*** (0.010) -0.203* (0.094)
Treatment school (T)	0.051*** (0.022)						
First year as treatment school (T1)	0.062*** (0.015)	0.060*** (0.015)	0.062* (0.022)	0.056*** (0.020)			
Second year as treatment school (T2)	0.040* (0.021)	0.034* (0.021)					
Treatment school pupils within 3 km (per 1000 pupils)		0.044** (0.022)		0.023 (0.036)			
Treatment school pupils within 3–6 km (per 1000 pupils)		-0.014 (0.015)		-0.041 (0.027)			
Total pupils within 3 km (per 1000 pupils)		-0.033** (0.013)		-0.035* (0.019)	0.018 (0.021)	0.021 (0.019)	
Total pupils within 3–6 km (per 1000 pupils)		-0.010 (0.012)		0.022 (0.027)	-0.010 (0.012)	-0.021 (0.015)	
Indicator received first year of deworming treatment when offered (1998 for Group 1, 1999 for Group 2)				0.100*** (0.014)			
(First year as treatment school Indicator) * (Received treatment, when offered)				-0.012 (0.020)			
1996 district exam score, school average	0.063*** (0.021)	0.071*** (0.020)	0.063*** (0.020)	0.058 (0.032)	0.091** (0.038)	0.021 (0.026)	0.003 (0.023)

TABLE VI
DEWORMING HEALTH EXTERNALITIES WITHIN SCHOOLS, JANUARY TO MARCH 1999^a

	Group 1, Treated in 1998	Group 1, Untreated in 1998	Group 2, Treated in 1999	Group 2, Untreated in 1999	(Group 1, Treated in 1998) – (Group 2, Treated in 1999)	(Group 1, Untreated in 1998) – (Group 2, Untreated in 1999)
<i>Panel A: Selection into Treatment</i>						
Any moderate-heavy infection, 1998	0.39	0.44	–	–	–	–
Proportion of 1998 parasitological sample tracked to 1999 sample ^b	0.36	0.36	–	–	–	–
Access to latrine at home, 1998	0.84	0.80	0.81	0.86	0.03	-0.06 (0.04) (0.05)
Grade progression (= Grade – (Age – 6)), 1998	-2.0	-1.8	-1.8	-1.8	-0.2* (0.1) (0.2)	0.0
Weight-for-age (Z-score), 1998 (low scores denote undernutrition)	-1.58	-1.52	-1.57	-1.46	-0.01 (0.06) (0.11)	-0.06
Malaria/fever in past week (self-reported), 1998	0.37	0.41	0.40	0.39	-0.03 (0.04) (0.06)	-0.01
Clean (observed by field worker), 1998	0.53	0.59	0.60	0.66	-0.07 (0.05) (0.10)	-0.07
<i>Panel B: Health Outcomes</i>						
<i>Girls <13 years, and all boys</i>						
Any moderate-heavy infection, 1999	0.24	0.34	0.51	0.55	-0.27*** (0.06) (0.10)	-0.21**
Hookworm moderate-heavy infection, 1999	0.04	0.11	0.22	0.20	-0.19** (0.03) (0.05)	-0.09*
Roundworm moderate-heavy infection, 1999	0.08	0.12	0.22	0.30	-0.14** (0.04) (0.07)	-0.18**
Schistosomiasis moderate-heavy infection, 1999	0.09	0.08	0.20	0.13	-0.11 (0.06) (0.06)	-0.05
Whipworm moderate-heavy infection, 1999	0.12	0.16	0.16	0.20	-0.04 (0.16) (0.09)	-0.05
<i>Girls ≥13 years</i>						
Any moderate-heavy infection, 1998	0.31	0.28	–	–	–	–
Any moderate-heavy infection, 1999	0.27	0.43	0.32	0.54	-0.05 (0.17) (0.09)	-0.10
<i>Panel C: School Participation</i>						
School participation rate, May 1998 to March 1999 ^c	0.872	0.764	0.808	0.684	0.064** (0.032) (0.039)	0.080**

- Impressive cost-benefit ratio: Cost per additional year of schooling estimated to be 3.50 USD
- No significant gain in test scores, but MK argue that correlation between school presence and test scores wouldn't allow them to find any significant effect in their setting given the effects on school presence they find
- Relative to school-level randomisation, individual-level randomisation seems to doubly underestimate the effect
 - No externalities are estimated
 - The fact that there are spillover effects means the SUTVA is violated: The control group is better off than the true counterfactual should be
- MK argue that the effect is large enough effects to justify full subsidy, if not even paying people for being treated.

Baird et al. (2016). Worms at work: Long-run impacts of a child health investment. QJE

Follow-up on MK-2004 paper

- Impact of 2004 worm treatment on young adults ten years later considered
- Results:
 - Both men and women are significantly and considerably less likely to suffer from a moderate-heavy worm infection: -19ppt and -14ppt on bases of 32% and 34% respectively.
 - Men in treatment schools work 3.5 hrs more per week (base: 20.3 hrs), spend more time in non-agric. self-employment and are more likely to hold manufacturing jobs with higher wages. They also eat one more meal per week
 - Women in treatment schools spend more time in school, are 25% more likely to have passed the secondary school entrance exam and reallocate time from traditional agric. to non-agric. self-employment. They are also more likely to grow cash crops. Wages and living standards are larger for those >12 yo. in 2004.

- Government revenue gained from increased work hours exceeds direct subsidy cost. Estimated internal rate of return of 32%.
- Kernel estimates of the distribution of wage earnings and hours worked show a right shift for both men and women

Miguel & Kremer (2007). *The Illusion of Sustainability*. QJE.

See technology adoption chapter

Almond (2006). *Is the 1918 Influenza Pandemic Over? Long-term effects of in Utero Influenza exposure in the post-1940 US population*. JPE

- use the 1918 Q3 – 1919 Q1 short influenza pandemic in the US as a natural experiment to assess the impact of fetal health on life outcomes⁸
 - Mostly time series evidence: Look at outcomes for quarters between 1916 Q1 and 1920 Q4 and how the influenza has an additional effect beyond the time trend (those in utero during the influenza were born throughout the year 1919)
- $$y_i^c = \beta_0 + \beta_1 I(\text{birth in 1919}) + \beta_2 \text{yearofbirth} + \beta_3 \text{yearofbirth}^2 + \epsilon_i$$
- Almond considers heterogeneous effects for region/state, skin colour and gender, but finds the results in all subclasses
 - Two concerns:
 - if influenza caused more people to die among the affected cohorts, this systematically changes the pool of people we are comparing to the 1918 and 1920 born people. But mortality rates from influenza were very low and this goes against their results – the weakest people should have died making the 1919 sample a pool of stronger people.
 - people who were suffering from influenza in utero were treated differently during WWII (less likely to benefit from the opportunities of military service) and this explains the life outcomes – robustness check with a dummy for that and also birth states and quarter fixed effects upholds the results.
 - Results (very similar for men and women – distinction provides some robustness against the second concern as well):
 - 2ppt less likely to graduate from high school (on mean around 50%)
 - .15 fewer years of educ. (on mean of 10 years)
 - 1ppt more likely to be poor (on mean of around 20%)
 - .5ppt more likely to have a work-preventing disability (on mean of around 15%)
 - Almond also uses the variation of maternal infection within the affected cohort (men born 1918-20) and including state of birth fixed effects. The maternal infection rate was 1/3, so all coefficients have to be multiplied by 1/3 to get the effect on the outcome variable.

⁸ This goes back to Barker's (1998) ["Mothers, babies and health in later life"] fetal origins hypothesis: Shocks during the fetal period can have long-lasting impacts for life outcomes. There is some recent more recent evidence on this.

MEN BORN 1918–20: 1960 STATE OF BIRTH FIXED-EFFECTS RESULTS

DEPENDENT VARIABLE	INDEPENDENT VARIABLE			
	Maternal Infection (ψ)	Infant Mortality Rate	Attrition	State of Residence Dummies
Years of education	-.756*** [.259] -.793*** [.229] -.818*** [.228] -.759*** [.233]	-.0059* [.003] -.0553* [.0308] -.0329 [.0313]	-.426 [.318] -.392 [.314]	No No No Yes
High school graduate	-.101*** [.070] -.103*** [.035] -.109*** [.039] -.105*** [.0378]	-.0003* [.00005] -.0026 [.0051] -.00046 [.0052]	-.091* [.050] -.0908* [.0499]	No No No Yes
Log income	-.165*** [.0719] -.176*** [.060] -.172*** [.062] -.166*** [.0623]	-.0018** [.00086] -.0181* [.0086] -.0139 [.0085]	.0629 [.0557] .0707 [.0599]	No No No Yes
Poverty status (below 150% of poverty level)	.0424 [.0259] .0461** [.224] .0429* [.0233] .0386 [.0255]	.00059 [.00040] .0064* [.0039] .0041 [.0040]	-.0529 [.0398] -.0533 [.0397]	No No No Yes
Socioeconomic status (Duncan's socioeconomic index)	-2.711 [1.735]			No
		-.0150 [.03057]		No
		-.142 [.307]	-.9441 [2.372]	No
		-.023 [.298]	-.1075 [2.361]	Yes
Observations	16,566	16,566	16,566	

- channel/interpretation/mechanism unclear?
 - is it inherently worse conditions for influenza affected newborns that determine their life outcomes or their parents' situation as being affected from influenza during the time of childbirth? But parents with 1918 and 1920 born kids are also affected by influenza – so, if anything, it must be the combination of giving birth as a mother while being affected

Acemoglu & Johnson (2006). Disease and Development: The effect of life expectancy on economic growth. JPE.

- AJ estimate of the effect of life expectancy on economic growth as a cross-country exercise, instrumenting for life expectancy with a self-constructed instrument: predicted mortality.
- Theoretical idea: health may improve productivity and increase the population. The population increase first decreases GDP per capita (mechanically), but the productivity effect increases it in the longer run. But in the presence of inelastic supply of physical capital and land, we should expect decreasing returns to effective labour (productivity x

labour). This GE effect is not captured in micro estimates of health on performance. Therefore, such micro-estimates are exaggerating the true effect.

- Motivating observation: (log) life expectancy (left) seems to show some degree of convergence over the last couple of decades while (log) GDP p.c. (right) does not:

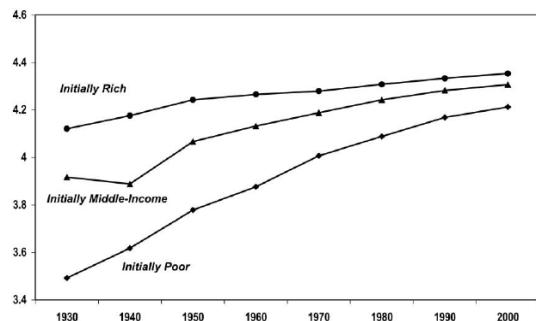


FIG. 1.—Log life expectancy at birth for initially rich, middle-income, and poor countries in the base sample.

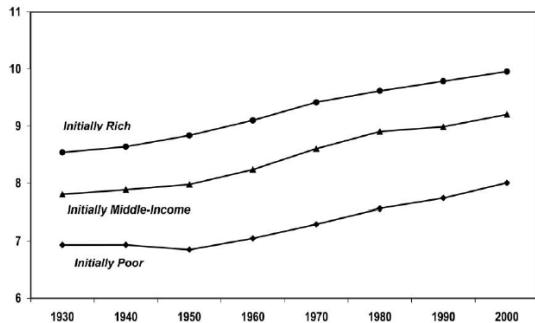


FIG. 2.—Log GDP per capita for initially rich, middle-income, and poor countries in the base sample.

- Underlying for the instrument is a variable capturing baseline mortality from 15 different diseases in a specific country from 1940 before many medical interventions/innovations took place M_{di40} (disease d , country i). They also have dummies for these major interventions for particular disease I_{dt} and mortality rates for a disease after a major intervention M_{dFt}
- The instrument is then constructed as:

$$M_{it}^I = \sum_d ((1 - I_{dt})M_{di40} + I_{dt}M_{dFt})$$

so it either picks up the baseline mortality rate as long as intervention hasn't happened yet or the health frontier mortality rate, common across countries (taken as zero in the baseline investigation)

- The idea is that this is not the actual mortality rate, but only predicted after 1940 for all time periods depending on whether some major intervention happened – which then changes the mortality rate for all countries – or not. But it changes the mortality rate *more* for countries for which the disease that the intervention addresses is relatively *more* important. This variation is used in a first difference setting to instrument changes in life expectancy in different countries to work around the obvious endogeneity.
- Issue / question: Are interventions for particular diseases endogenous? An intervention at all may be an exogenous shock to the world, but for which disease it happens may not. If research happens predominantly in rich countries, they are going to allocate more funding for diseases important in their context which in turn means that an intervention for these diseases which changes the predicted mortality rate of rich countries *more* relative to poor countries is more likely. This in turn may co-move with *larger* growth in economic prosperity in rich countries.
- They get a strong first stage and a placebo – the relationship between the predicted mortality and life expectancy before 1940 – displays a very flat line of fit (close to zero correlation indeed and insignificant)

TABLE 7
FALSIFICATION EXERCISE AND REDUCED FORMS: OLS REGRESSIONS

	Base Sample (1)	Low- and Middle-Income Countries (2)	Base Sample (3)	Low- and Middle-Income Countries (4)	Base Sample (5)	Low- and Middle-Income Countries (6)	Base Sample (7)	Low- and Middle-Income Countries (8)
A. Falsification Exercise								
	Dependent Variable: Change in Life Expectancy from 1900 to 1940	Dependent Variable: Change in Log Population from 1900 to 1940		Dependent Variable: Change in Log GDP from 1900 to 1940		Dependent Variable: Change in Log GDP per Capita from 1900 to 1940		
Change in predicted mortality from 1940 to 1980	.13 (.11)	.21 (.16)	-.17 (.15)	-.13 (.24)	.009 (.24)	.05 (.36)	.02 (.17)	.04 (.23)
R ²	.04	.06	.03	.01	.0001	.0008	.0005	.0008
Number of countries	47	36	45	34	31	20	31	20
B. Reduced Forms								
	Dependent Variable: Change in Life Expectancy from 1940 to 1980	Dependent Variable: Change in Log Population from 1940 to 1980		Dependent Variable: Change in Log GDP from 1940 to 1980		Dependent Variable: Change in Log GDP per Capita from 1940 to 1980		
Change in predicted mortality from 1940 to 1980	-.44 (.06)	-.30 (.08)	-.74 (.15)	-.62 (.21)	-.14 (.22)	.11 (.28)	.58 (.15)	.71 (.20)
R ²	.5	.27	.29	.17	.008	.004	.18	.18
Number of countries	47	36	47	36	47	36	47	36

- Results:

- fairly large effect on population – 1% increase in life expectancy increases the population by 1.7-2 % over a 40-60 year horizon.
- no statistically significant effect on GDP, but large standard errors and not very large p values – not a precise zero estimate.

TABLE 8
EFFECT OF LIFE EXPECTANCY ON POPULATION, TOTAL BIRTHS, AND POPULATION UNDER AGE 20: 2SLS ESTIMATES

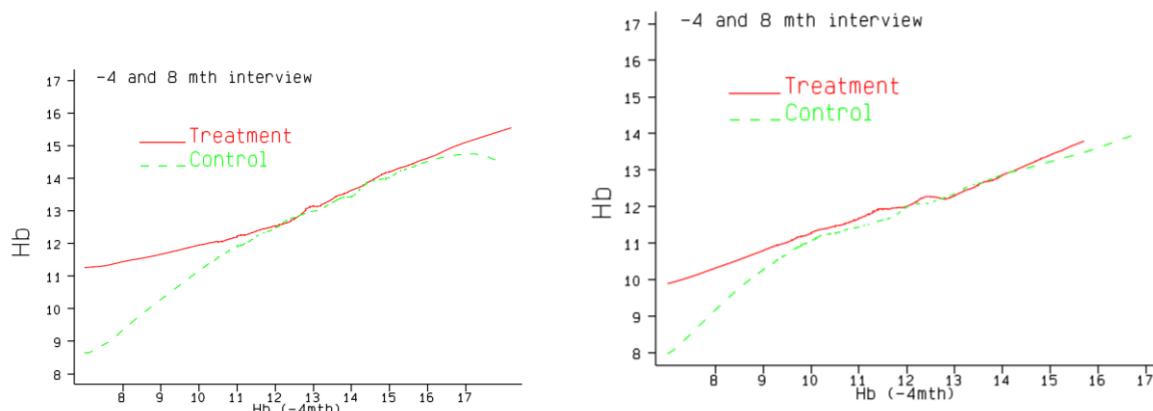
	BASELINE PREDICTED MORTALITY INSTRUMENT						GLOBAL MORTALITY INSTRUMENT
	Base Sample (1)	Base Sample (2)	Low- and Middle-Income Countries Only (3)	Low- and Middle-Income Countries Only (4)	Base Sample: Interaction with Institutions (5)	Base Sample: Interaction with Initial (1930) Value of Dependent Variable (6)	Base Sample (7)
A. Dependent Variable: Log Population							
Just 1940 and 1980	Just 1940 and 2000	Just 1940 and 1980	Just 1940 and 2000	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980
Log life expectancy	1.67 (.50)	1.96 (.53)	2.04 (1.01)	2.18 (.93)	1.63 (.73)	1.68 (.44)	1.70 (.48)
Postyear dummy × institutions or initial log population					-.006 (.05)	-.055 (.03)	
Number of countries	47	47	36	36	47	47	47
B. Dependent Variable: Log Total Births							
Just 1940 and 1980	Just 1940 and 1990	Just 1940 and 1980	Just 1940 and 1990	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980
Log life expectancy	2.53 (.70)	2.15 (.64)	2.92 (1.36)	2.67 (1.20)	2.40 (1.09)	2.53 (.70)	2.52 (.72)
Postyear dummy × institutions or initial log of total births					-.018 (.09)	-.056 (.05)	
Number of countries	45	45	34	34	45	44	45

TABLE 9
EFFECT OF LIFE EXPECTANCY ON GDP, PER CAPITA GDP, AND GDP PER WORKING AGE POPULATION: 2SLS ESTIMATES

	BASELINE PREDICTED MORTALITY INSTRUMENT					GLOBAL MORTALITY INSTRUMENT	
	Base Sample: Just 1940 and 1980 (1)	Base Sample: Just 1940 and 2000 (2)	Low- and Middle-Income Countries Only: Just 1940 and 1980 (3)	Low- and Middle-Income Countries Only: Just 1940 and 2000 (4)	Base Sample: Interaction with Institutions: Just 1940 and 1980 (5)	Base Sample: Interaction with Initial (1930) Value of Dependent Variable: Just 1940 and 1980 (6)	Base Sample: Just 1940 and 1980 (7)
A. Dependent Variable: Log GDP							
Log life expectancy	.32 (.84)	.42 (.52)	-.39 (1.44)	-.58 (1.09)	-.11 (.99)	-.069 (.73)	.46 (.73)
Postyear dummy \times institutions or initial log GDP					-.063 (.055)	-.109 (.059)	
Number of countries	47	47	36	36	47	47	47
B. Dependent Variable: Log per Capita GDP							
Log life expectancy	-1.32 (.56)	-1.51 (.57)	-2.35 (1.13)	-2.70 (1.40)	-1.64 (.77)	-1.59 (1.22)	-1.21 (.52)
Postyear dummy \times institutions or initial log per capita GDP					-.049 (.060)	-.073 (.278)	
Number of countries	47	47	36	36	47	47	47

Thomas et al. (2006). Causal Effect of Health on Labor market Outcomes: Experimental Evidence. California Center for Population Research

- RCT in Java, Indonesia: Treatment group receives 120mg of iron every week (known iron deficiency), control receives placebo
- 17,000 individuals, six months endline results presented for men aged 30-70, randomisation at household level
- Attrition: 2%
- First stage: The treatment only increased the haemoglobin levels for those with low initial levels of haemoglobin 4 months after the begin of the treatment phase (for both men, left, and women, right)



- Results (ITT):
 - better physical health, psycho-social health, economic success
 - working more, sleeping less, losing less work time to illness, more energetic, more able to conduct physically arduous activities
 - a few are not statistically significant, some effects quite small, others quite substantial.

- results for women are generally in the same direction, but less pronounced – quite substantial differences for many measures between men and women
- There is no return calculation! So, it's unclear whether people should do this or not.

Table 5: Intent to treat effects on subjects stratified by Hb status at baseline
Work, earnings and hours work

Indicator	Sample	Change in Treatments - Change in Controls		
		If low Hb @baseline DinD (1)	If high Hb @baseline DinD (2)	Low-High Hb @baseline DinDinD (3)
1. Pr(not working in month of survey interview)	Male	-0.036 [0.012]	-0.003 [0.007]	-0.033 [0.014]
	Female	-0.020 [0.014]	0.029 [0.020]	-0.049 [0.024]
2. $\sqrt{Earnings}$ (Rp 000) (last 4 months)	Male	0.576 [0.299]	-0.012 [0.173]	0.582 [0.346]
	Female	0.163 [0.091]	0.033 [0.127]	0.130 [0.156]
3. Hours spent working (last 4 months)	Male	-12.968 [36.368]	-44.185 [21.027]	31.217 [42.013]
	Female	9.644 [15.264]	30.137 [21.425]	-20.493 [26.309]
4. $\sqrt{Hrly\ earnings}$ (Rp 000) (last 4 months)	Male	0.126 [0.066]	0.007 [0.038]	0.119 [0.076]
	Female	0.034 [0.025]	-0.009 [0.035]	0.043 [0.043]
5. $\sqrt{Hrly\ earnings}$ (Rp 000) conditional on being non zero (last 4 months)	Male	0.113 [0.069]	-0.006 [0.040]	0.119 [0.080]
	Female	0.056 [0.026]	-0.021 [0.037]	0.077 [0.046]
6. $\sqrt{Earnings}$ (Rp 000) if self-employed (last 4 months)	Male	1.091 [0.445]	-0.386 [0.285]	1.477 [0.528]
	Female	0.177 [0.214]	0.101 [0.305]	0.076 [0.373]
7. $\sqrt{Hrly\ earnings}$ (Rp 000) if self-employed (last 4 months)	Male	0.230 [0.093]	-0.078 [0.059]	0.308 [0.110]
	Female	0.031 [0.052]	-0.036 [0.074]	0.067 [0.090]

Adhvaryu et al. (2015). Management and Shocks to Worker Productivity. WP

- Structural approach: Model how productivity in a garment factory (production per hour, very specific measure) is determined by a variety of factors, one of them is managerial quality of the manager (Bloom/van Reenen management survey used).
- Interesting effect for health: The effect of productivity shocks – here differential pollution exposure argued to be exogenous – on productivity in the presence of good or bad managers.
- They show that – as a first stage – there is a strong relationship between managerial quality and the regularity task reassignment as well between managerial quality and line productivity
- Average pieces produced per hour is 50 in the sampled garment factories.

- Controlling for year, month, day of week, hour of day FE and coarse particle matter (PM), there is a negative relationship between fine PM and productivity (pieces produced and efficiency defined as ratio of actual pieces produced to target quantity):
 - Doesn't seem to be a large effect though: one sd. Increase in fine PM decreases pieces produced by 0.7 (on a mean of 50!)

Figure 3A

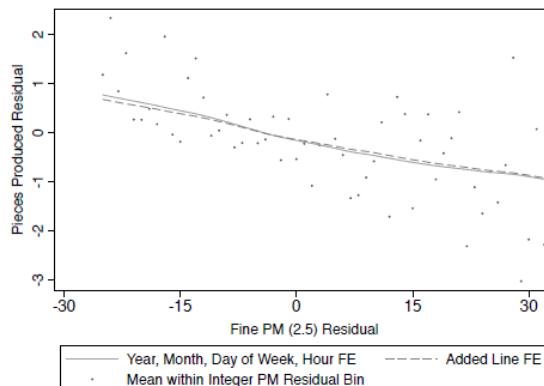
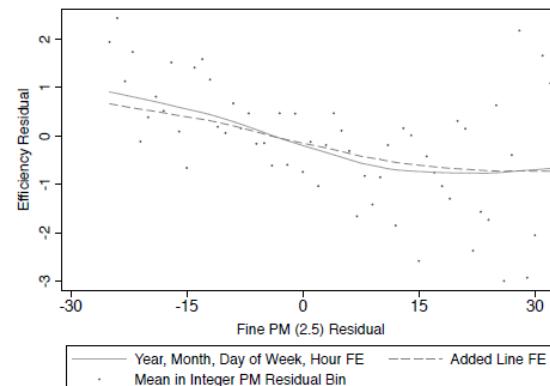


Figure 3B

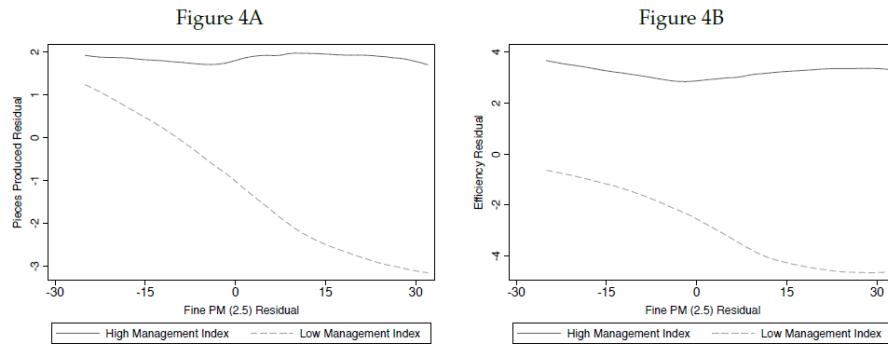


Figures 3A and 3B depict relationships between residuals of fine PM exposure and pieces produced (3A) or efficiency (3B). Residuals are from regressions of each variable on year, month, day of week, and hour of day fixed effects as well as coarse PM. Fine PM residual trimmed at 5th and 95th percentile. Scatter depicts mean residual of productivity measure within integer fine PM residual bins, Lines depict local polynomial smoothing with and without production line fixed effects included in regression producing residuals.

Table 5: Pollution and Worker Productivity

	(1)	(2)	(3)	(4)	(5)
	Pieces Produced pieces produced by worker this hour			Efficiency (pieces produced / target pieces)	
Fine PM (Std)	-0.69843*** (0.17564)	-0.72892*** (0.17309)	-0.19781 (0.19612)	-0.66694*** (0.16508)	-0.22024 (0.19072)
Task Difficulty (SAM Std) X Fine PM (Std)		-0.28229*** (0.07371)			
Seat in Line X Fine PM (Std)			-0.01135*** (0.00167)		-0.01021*** (0.00160)
Production Target					
Coarse PM and Interaction Controls					
Year, Month, Day-of-Week, Hour-of-Day FE	Yes	Yes	Yes	Yes	Yes
Line FE	No	Yes	Yes	No	Yes
Observations	1,054,962	1,054,962	1,054,962	1,054,962	1,054,962
Mean of Dependent Variable	50.60				49.14

- Good managers are better at reallocating tasks and mitigating the effects of productivity shocks through health.
 - They are particularly more likely to re-allocate workers with above-average pollution exposure and workers with below-average pollution (probably reshuffling between those which is why you get the U curve below). This results in no productivity losses



Figures 4A and 4B depict relationships between residuals of fine PM exposure and pieces produced (4A) or efficiency (4B) by high and low managerial quality. High management corresponds to above median Composite Index; low management to below median. Residuals are from regressions of each variable on year, month, day of week, and hour of day fixed effects as well as coarse PM. Fine PM residual trimmed at 5th and 95th percentile. Lines depict local polynomial smoothing.

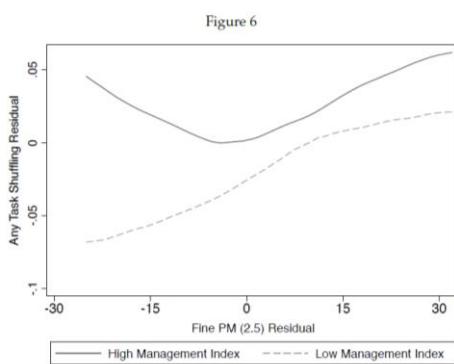


Figure 6 depicts relationship between residuals of fine PM exposure and dummy for any task shuffling by high and low managerial quality. High management corresponds to above median Composite Index; low management to below median. Residuals are from regressions of each variable on year, month, day of week, and hour of day fixed effects as well as coarse PM. Fine PM residual trimmed at 5th and 95th percentile. Lines depict local polynomial smoothing.

16.3 Market Failures in Health

- Some authors find the lack of information a problem leading to under-utilisation of health products (in particular preventive ones)
 - However, it's not clear to what extent information provision can achieve the optimal level of health. Where is the line between information provision and marketing? How do we measure the optimal level of health?
- Subsidies as a response to market failures can have adverse effects as well
 - Screening, sunk-cost effect, signalling effect → see chapter on *Pricing and WTP*
- The above worms papers by Miguel-Kremer & co. clearly also provide evidence for externalities as market failures on top of their returns to health estimates.

Madajewicz et al. (2007). *Can information alone change behaviour? Response to arsenic contamination of groundwater in Bangladesh.* JDE

- They examine whether providing information on the share of arsenic in water from a well in Bangladesh changes the behaviour as to what well to use
- They carried out a baseline survey in Araihazar, a district south of Dakar, and tested all 6500 wells in villages there. They provided information – both in words and pictures/songs for the illiterate population – on whether the well was safe or not (exceeded

a certain concentration of arsenic deemed to be unsafe). Finally, they collected endline data a year later. They also collected endline data in four control regions in other districts.

- Concerns: (i) the control regions are not randomly chosen and quite different on observable characteristics; (ii) for the endline survey in controls, they asked households whether they had changed wells within the past year since they did not have information on the original well used – huge measurement error? Suggestive question?
- Even if they control for district FE, no evidence for common trend among control regions and Araihazar
- Result: Having been provided information on an unsafe well changes the probability of changing well by 0.37. They increase the time walking fifteen-fold (by 4.3 min in value). Communication by the media is no less effective (although only 2/3 have access to a medium). Their results table on the media question shows that the likelihood of being aware of arsenic is 0.2 higher (20ppt) if living in Araihazar relative to anywhere else (where media is reporting about this problem everywhere) – this doesn't seem to be a small difference!?
- Concern about biased reporting when having unsafe arsenic levels to please the reporter mitigated by the fact that the urinary arsenic levels among those households are lower by pretty much the same percentage.

Banerjee et al. (2015). *Movies, Margins and Marketing: Encouraging the adoption of iron-fortified salt*. NBER WP.

- They randomly selected 200 out of 400 Indian villages to sell a new fortified salt at a 50% discount. Within the treatment group, other intervention arms were run:
 - increase in retailer margin for one or several shopkeepers (subgroup of treatment where one shopkeeper got a discount on the wholesale price, another subgroup where all got the discount and a third one where no one got the discount)
 - screening of an edutainment movie on the benefits
 - flyer informing households of availability
 - free distribution
- Results:
 - in control group, no one buys the salt
 - if provided for free, half the households take it up
 - at half price, 20% try and 10% still use it after about three hours
 - movie makes the purchase increase by 5.5 ppt
 - when all retailers got the better margin (and only when all got it), purchases increased by at least as much as the movie intervention achieved

16.4 Incentives for Public Delivery

Björkman & Svensson (2009). *Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda*. QJE.

- Study how motivation and capacity strengthening of local communities to hold their health providers accountable and monitor them impacts on the exertion of effort and quality of health providers in an RCT in Uganda.
- Theoretical problem of moral hazard – no information and no enforcement mechanism. It is very hard for the (distant) government to acquire good information on the status of health provision in rural areas and enforce certain behaviour. Almost certainly, there is a lack of political will and resources to do this
 - However, local communities have
 - better information
 - ability of enforcement through social punishment
 - the will to enforce since they are the beneficiaries
- Local communities just face a co-ordination issue with the typical free-rider problems. The Coase Theorem clearly fails (empirically, otw. we would observe better health provision). Are transaction costs too high?
 - The intervention is really about paying the transaction cost. The NGO facilitates coordination...
- Intervention: Local NGOs facilitated (i) village meetings in which community members discussed baseline information on the status of health and were encouraged to develop a plan identifying problems and steps to be taken to improve health provision. It was up to the community to continue this monitoring process. The local NGOs also met with the (ii) health service provider to contrast their information on service provision with the household survey and facilitated a (iii) final meeting between community members – chosen in the first community meeting – and health workers.
 - disseminate information on report cards
 - action plan developed in the meetings based on what steps should be taken, including actions to be taken by the health provider and how improvements (or the lack thereof) could be monitored
 - In the final (interface) meeting, the action plan was presented and discussed and a “community contract” set, specifying what had to be done, by whom, and how
- The intervention was carried out in 25 treatment dispensaries, 25 were control. All dispensaries (a bit like tiny hospitals with usually one doctor and 2-3 nurses) were in rural areas.
- various local NGOS – community-based organisations – focus primarily on health education
- pilot identified that lack of relevant information on service delivery and community's entitlements as well as failure of coordination on what to expect from the provider as reasons holding back effective monitoring. Individuals knew about their own situation with the provider, but not the community-wide one. Essentially two constraints that were relaxed: (i) lack of information, (ii) free-riding problem in community monitoring.
- Health workers in Uganda do not have effective incentives to exert effort (paid on seniority and education mainly, not on job performance)
- Result: Increase in weight of infants by 0.14 z scores, 33% reduction in <5 yo. mortality (U5MR); a huge number of other health outcomes considered (in the fields of treatment practices, immunisation, coverage), giving more pregnancies, less waiting time, way more immunisations for babies

TABLE II
PROGRAM IMPACT ON MONITORING AND INFORMATION

Dependent variable	Suggestion box (1)	Numbered waiting cards (2)	Poster informing free services (3)	Poster on patients' rights (4)	Average standardized effect (5)	Discuss facility in LC meetings (6)	Received information about HUMC (7)
Specification:							
Program impact	0.32*** (0.08)	0.16* (0.09)	0.27*** (0.09)	0.14 (0.10)	2.55*** (0.55)	0.13*** (0.03)	0.04*** (0.01)
Mean control group	0	0.04	0.12	0.12	—	0.33	0.08
Observations	50	50	50	50	50	3,119	4,996

TABLE III
PROGRAM IMPACT ON TREATMENT PRACTICES AND MANAGEMENT

Spec.	Dep. variable	Model	Program impact	2005	Mean control group 2005	Obs.
(1)	Equipment used	DD	0.08** (0.03)	-0.07*** (0.02)	0.41	5,280
(2)	Equipment used	OLS	0.01 (0.02)		0.41	2,758
(3)	Waiting time	DD	-12.3* (7.1)	-12.4** (5.2)	131	6,602
(4)	Waiting time	OLS	-5.16 (5.51)		131	3,426
(5)	Absence rate	OLS	-0.13** (0.06)		0.47	46
(6)	Management of clinic	OLS	-1.20*** (0.33)		0.49	50
(7)	Health information	OLS	0.07*** (0.02)		0.32	4,996
(8)	Importance of family planning	OLS	0.06** (0.02)		0.31	4,996
(9)	Stockouts	OLS	-0.15** (0.07)		0.50	42

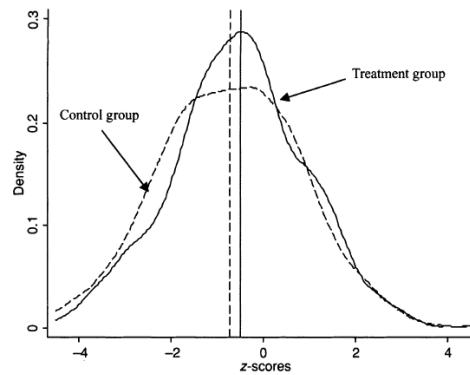


FIGURE II
Distributions of Weight-for-Age z-Scores for Treatment and Control Groups

TABLE IV
PROGRAM IMPACT ON IMMUNIZATION

Group Specification:	Newborn (1)	Under 1 year (2)	1 year old (3)	2 years old (4)	3 years old (5)	4 years old (6)
Average standardized effect	1.30* (0.70)	1.44** (0.72)	1.24** (0.63)	0.72 (0.58)	2.01*** (0.67)	0.86 (0.80)
Observations	173	929	940	951	1,110	526

TABLE VI
PROGRAM IMPACT ON HEALTH OUTCOMES

Dependent variable	Births (1)	Pregnancies (2)	U5MR (3)	Child death (4)	Weight-for-age z-scores (5)	Weight-for-age z-scores (6)
Specification:						
Program impact	-0.016 (0.013)	-0.03** (0.014)	-49.9* (26.9)		0.14** (0.07)	0.14** (0.07)
Child age (log)						-1.27*** (0.07)
Female						0.27*** (0.09)
Program impact × year of birth 2005					-0.026** (0.013)	
Program impact × year of birth 2004					-0.019** (0.008)	
Program impact × year of birth 2003					0.003 (0.009)	
Program impact × year of birth 2002					0.000 (0.006)	
Program impact × year of birth 2001					0.002 (0.006)	
Mean control group 2005	0.21	0.29	144	0.029	-0.71	-0.71
Observations	4,996	4,996	50	5,094	1,135	1,135

17 Risk Sharing & Insurance

17.1 General Summary

17.2 Papers

Townsend (1994). *Risk and Insurance in Village India*. Ecta.

- He uses detailed data on three village economies in southern India to test a model of full insurance.
- There are various risk-bearing institutions that could function within village economies:
 - Diversification of land-holdings across space and crops
 - Storage of grains
 - Asset purchases and sales
 - Borrowing and lending / credit markets
 - Gifts and transfers within the family network
- In theory, if preferences are time separable, display weak risk aversion, all individuals discount at the same rate, and all information is held in common, then the optimal allocation of risk means that individual consumption is entirely determined by aggregate consumption.
- Consider a pareto problem to determine the efficient allocation in an economy:

$$\max \sum_{i=1}^N \mu_i U_i \quad s.t. \quad \sum_i c_{ist} \leq \sum_i y_{ist} \quad \forall s, t \quad (LM: \lambda_{st})$$

Where μ_i are pareto weights, $s \in \{1, \dots, S\}$ states of the world and the rest follows the obvious notation. Assume concave utility functions.

- The FOC becomes:

$$u'(c_{ist}) = \frac{\lambda_{st}}{\mu_i} \quad \rightarrow \quad \frac{u'(c_{ist})}{u'(c_{jst})} = \frac{\mu_j}{\mu_i}$$

Under CARA utility $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, summing across households and averaging gives:

$$c_{ist} = \bar{c}_{st} + \frac{1}{\sigma} \left(\ln \mu_i - \frac{1}{N} \sum_j \ln \mu_j \right)$$

That is, individual consumption does *not* depend on individual income.

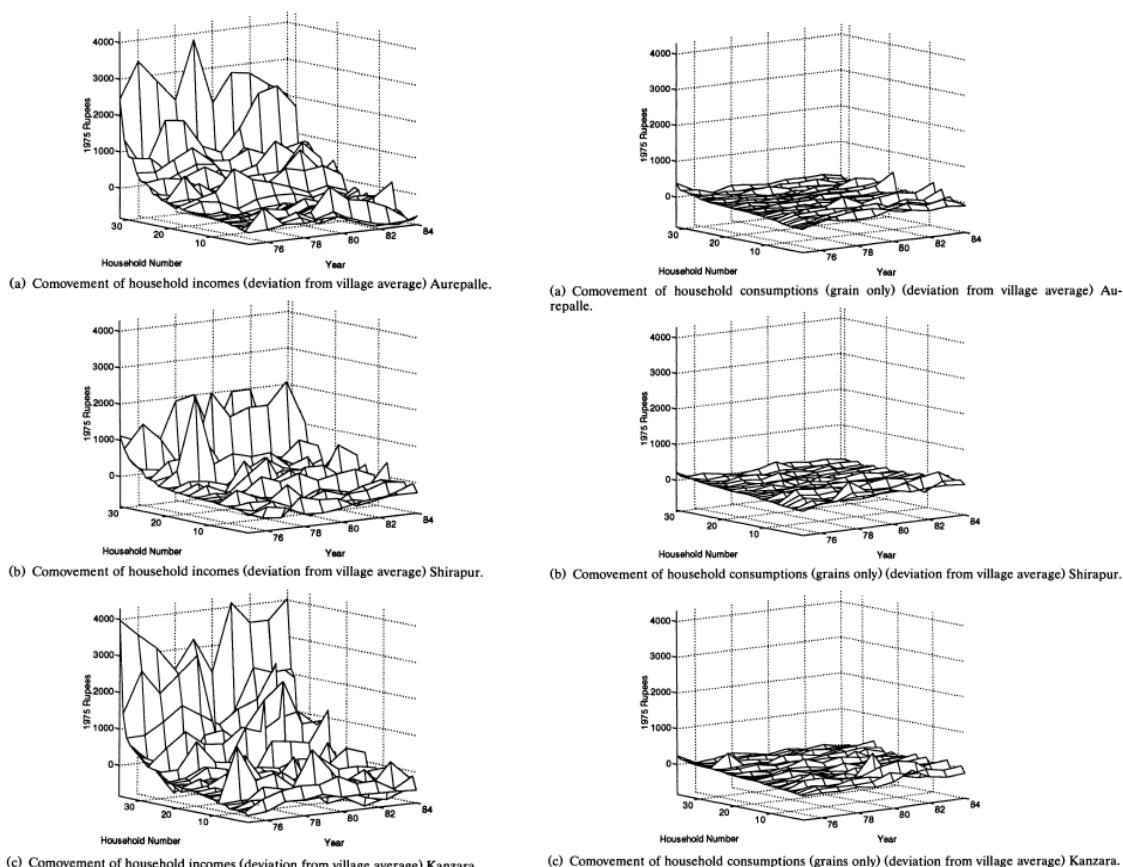
- They take two approaches to estimate whether this model holds:
 - (i) time series. Estimate *across time within household*

$$c_{it} = \alpha_i + \gamma_i \bar{c}_t + \delta_i \bar{A}_t + \zeta_i X_{it} + u_{it}$$

where A comes from the original utility function and X_{it} are different potential co-variates. The test to be run is whether $\gamma_i = 1$ and $\zeta_i = 0$ if income variables are taken as X_{it} . Note that, by definition, $\frac{1}{N} \sum_i \gamma_i = 1$, i.e. the average estimates for γ from the regressions per household should be one. However, the dispersion

around one from the different regression is of some interest and is for two villages rather small. The estimate of ζ for income is indeed closely around 0 for all villages, although with substantial standard errors for one of them.

- (ii) panel. Estimate the same equation including y_{it} as household income in the cross section. Now, it is of interest to test whether the coefficient β on income is different from zero. This is not the case for some sorts of income, however, the coefficient is always rather small.
- In general, there is definitely a lot more variation in income (below, left) in these villages than in consumption (below, right), indicating that the village economies are actually quite good at risk sharing:



- The issue with the tests is that – if we are not able to reject the null – we still don't learn where the smoothing comes from. Complete set of Arrow-Debreu markets? Permanent income hypothesis / lifetime consumption smoothing? Formal insurance? Gifts? (s. also above on possible institutions)
 - The authors can test whether the estimated household fixed effects or pareto weights are correlated with actual wealth and reject. This constitutes mild evidence against altruism and complete markets (apparently... no idea why)
 - They conclude that credit markets and gifts seem to be the main drivers in their setting.

Udry (1994). Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria. RevEconStud

- Studies the credit market within villages in northern Nigeria and how flexible and informal loan contracts provide risk pooling
- Standard model of credit contracts involves information asymmetries that have to be dealt with. In the village context at hand, this is not a problem – there seems to be a free flow of information. Much to the contrary, informal loans without written contract are used as an insurance device.
 - When borrowers experience an adverse shock, they tend to pay less interest. Note that this is standard in loan contracts and does not necessarily come from a state-contingent payment risk pooling story. The adverse shock might force the borrower to (partly) default.
 - When lenders experience an adverse shock, they tend to receive more interest payment (both, see table below). This is extra-ordinary and not part of the standard loan story. However, it may be driven by the fact that wealthier lenders are the ones demanding higher interest and holding more risky projects. Udry can exclude this on a CE model, s. below

Realized terms vs. borrower and lender shocks received

Adverse shock received by	Sample means		
	Monthly interest rates ^a	Simple interest rates ^b	Repayment period in days
(A) Borrower			
—no shock	0·5%	20·4%	67
—adverse shock	-4·0%	-0·6%	72
Impact of shock			
—on mean:	lower	lower	longer
—(t)	(1·58)	(2·20)	(1·03)
(B) Lender			
—no shock	-7·5%	-5·0%	89
—adverse shock	2·6%	11·8%	80
Impact of shock			
—on mean:	higher	higher	shorter
—(t)	(4·56)	(3·06)	(1·89)

- Udry uses two models: (i) a competitive equilibrium model with state-contingent claims and (ii) bilateral interaction model between a lender and borrower. Both give rise to the same conclusion in the model that state-contingent payments are embedded in the loan repayments, meaning that risk-pooling is at work here
 - Udry needs (i) to exclude that the correlation in the table for lenders is simply due to the fact that wealthy lenders are those that charge higher interest and have riskier projects (more adverse shocks) at the same time. In the CE model, wealth is controlled for and the relationship persists.
 - The CE model (i) with perfect (full) risk sharing is rejected; however, there is a considerable amount of risk sharing going on.

Ligon et al. (2002). Informal insurance arrangements with limited commitment: Theory and evidence from village economies. RevEconStud

- Departing from Townsend's (94) full risk-pooling model, LTW try and explain why risk pooling only happens to some extent in village economies: There is limited commitment – an “insurance scheme is only feasible if the long-term benefits of making a transfer in terms of future insurance exceed the short-term costs, that is, promises of future reciprocation are perceived to be credible and sufficiently attractive.”
- They consider two types of limited commitment that impose two different restrictions:
 - (i) static LC models: transfers must be stationary, i.e. in the same state of the world, the same transfers are always made. This looks more like gifts than loans
 - (ii) dynamic LC models: current transfers can depend on the whole *history* of past states, not only the current states. This is a more flexible specification and looks more like loans.
- They find that in particular the dynamic LC models are much better at explaining the actual degree of risk-pooling observed in three Indian village economies (the same as Townsend) than either a fully autarkic model or a full risk-sharing model.
- The structure of the model is the following (assume two households for simplification):
 - A state draw $s \in S$ is realised each period and consumers receive income $y_i(s)$
 - Income follows a Markov process with transition prob.s π_{sr} between state s, r
 - Households are infinitely lived, risk-averse and joint utility maximisers
 - A contract specifies transfer $\tau(h)$ from 1 to 2 as a function of the history h
 - If a household breaches the agreement, it is excluded from the mutual insurance scheme in the future and suffers a loss $P_i(s) \geq 0$ (e.g. shame)
- Relative to the full risk-sharing framework, a key additional LC constraint is: $U_{i,t}(h_t) \geq -P_i(s_t)$. This essentially means that for both households, utility must be within a closed interval each period (it directly imposes a lower bound and 1's lower bound translates into 2's upper bound and vice versa): $U_i^s \in [\underline{U}_i^s, \overline{U}_i^s]$. Let $U_2^s(U_1^s)$ describe the pareto frontier. In the dynamic LC model, this is given by:

$$U_2^s(U_1^s) = \max u_2(y_2(s) + \tau_s) - u_2(y_s) + \delta \sum_{r \in S} \pi_{rs} U_2^r(U_1^r)$$

Subject to

$$\begin{aligned} u(y_1(s) - \tau_s) - u(y_1(s)) + \delta \sum_{r \in S} \pi_{rs} U_1^r &\geq U_1^s, \\ U_1^r &\geq \underline{U}_1^r, \quad U_2^r(U_1^r) \geq \underline{U}_2^r \quad \forall r \in S \end{aligned}$$

- This admits a relatively straight-forward solution in which the slope of the pareto frontier λ which is the LM on the first constraint (by the Env. Theorem) stays the same within a certain limit. That is:

$$\lambda(h_{t+1}) = \begin{cases} \underline{\lambda}_r & \text{if } \lambda(h_t) < \underline{\lambda}_r \\ \lambda(h_t) & \text{if } \lambda(h_t) \in [\underline{\lambda}_r, \overline{\lambda}_r] \\ \overline{\lambda}_r & \text{if } \lambda(h_t) > \overline{\lambda}_r \end{cases}$$

The bounds come directly from the IC constraints (second line of constraints) and, in the solution, you essentially jump in consumption as little as possible along the pareto frontier. No jumps would be full risk-sharing

The only difference in the static LC model is that – if λ is within the bounds – it always stays λ_0 , the original λ from the initial state.

- It can be shown that this admits no solution for certain parameter values, in particular sufficiently low δ (high impatience) and low punishments P . For the opposite case, this can even degenerate to the full risk-sharing case. In this sense, the model *nests* both the autarky case and Townsend's full risk-pooling case.
- After the setup of the model, LTW calibrate it by assuming CRRA utility with parameter γ and finding the intervals for λ for this utility function. They simulate the consumption path, given by the model, and then estimate the parameters $\theta = (\gamma, \delta, P)$ in two different ways (for the dynamic LC case):
 - (i) [level estimator $\bar{\theta}$]: They match levels of (log) consumption in an ML procedure, i.e. estimating

$$\log c_{i,t} = \log \hat{c}_t(\lambda_{i,t}, \bar{c}_t | \theta) + u_{i,t}$$

where \hat{c} is predicted consumption from the model

- (ii) [changes-in-shares estimator $\dot{\theta}$]: They match changes in shares of consumption σ in an ML procedure, i.e. estimating

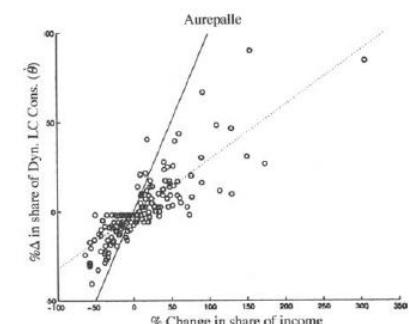
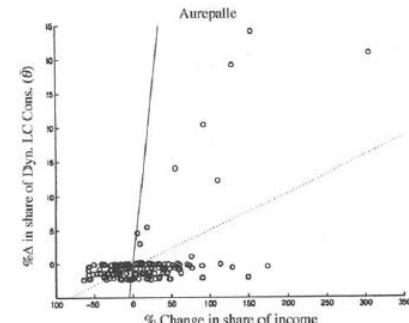
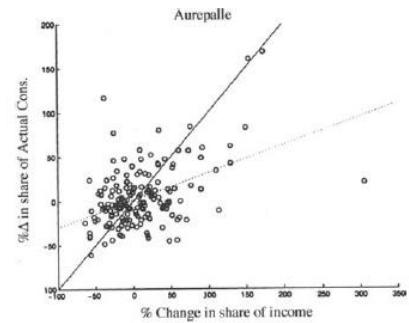
$$\log \frac{\sigma_{i,t+1}}{\sigma_{i,t}} = \log \frac{\hat{\sigma}_{i,t+1}(\theta)}{\hat{\sigma}_{i,t}(\theta)} + v_{i,t+1} - v_{i,t}$$

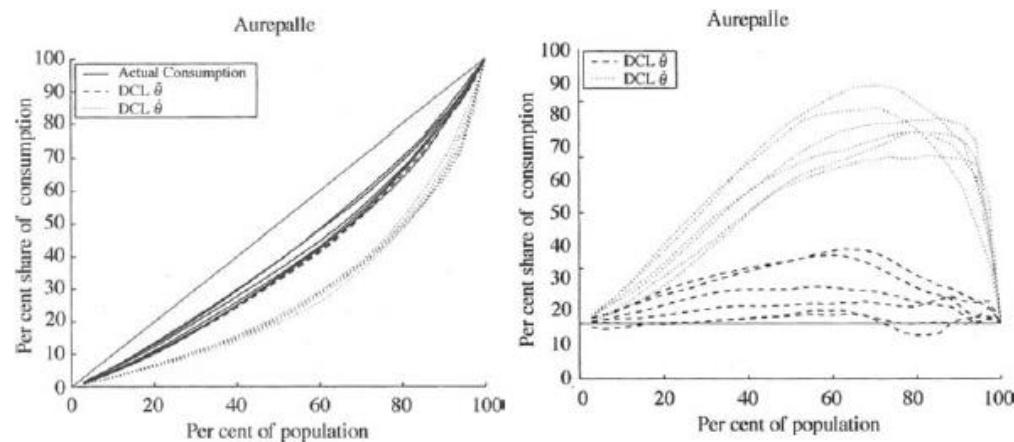
- They estimate the autarky case, just the utility parameter (pareto optimal case), the static LC model and (two estimations of) the dynamic LC model and the changes-in-shares dynamic LC model gives the best fit (the log likelihood is the “criterion”, the measure of best fit – by definition, the more parameters are estimated, the better the fit so the ranking is autarky, pareto optimal case, static LC model, dynamic LC model:

Village	Model	γ	$P(1 - \delta)$	δ	Criterion
Aurepalle	Autarky	—	—	—	-207.0320
	Pareto Optimal	26.5659	—	—	-95.0120
	Static LC	1.5576	0.0044	—	-65.0657
	Dynamic LC ($\bar{\theta}$)	1.5439	0.0052	0.7048	-64.1456
	Dynamic LC (θ)	0.9501	0.0006	0.8525	—
Shirapur	Autarky	—	—	—	-197.6770
	Pareto Optimal	5.4774	—	—	-76.4276
	Static LC	1.6081	0.0027	—	-70.2262
	Dynamic LC ($\bar{\theta}$)	1.4982	0.0082	0.7093	-63.2487
	Dynamic LC (θ)	0.8420	0.0000	0.9487	—
Kanzara	Autarky	—	—	—	-153.5230
	Pareto Optimal	0.0100	—	—	-30.6478
	Static LC	1.5764	0.0043	—	-13.1712
	Dynamic LC ($\bar{\theta}$)	1.4393	0.0166	0.9485	-6.3231
	Dynamic LC (θ)	0.8435	0.0000	0.9385	—

Note: For ease of interpretation, the column labelled $P(1 - \delta)$ reports the estimated punishment scaled by $1 - \delta$, so that it has the interpretation of a per-period punishment, measured in utils. For the static limited commitment model, δ is taken to be zero.

- While the changes-in-shares estimated LC model provides a good fit to the actual data for the relationship between changes in shares of consumption and changes in shares of income, the level estimated LC model provides a much better fit for the distribution of consumption levels observed. This is natural: The level estimator tries to match the distribution while the changes-in-shares estimator tries to match consumption shares.
- The top figure are the actual values observed. The middle one displays the predictions by the levels estimations (looks like almost full insurance) while the bottom one displays the predictions by the changes-in-shares estimator (very close to the actual one).
- In the second figure, the Lorenz curve is displayed. The left one shows the actual straight line, levels prediction dashed, and changes-in-shares prediction dotted. The right one shows the difference between actual and predictions for both cases. The lines are all the different time periods.
- Results for the other two villages are very similar.





- The authors also verify whether deviations by the predictions from the actual consumption patterns are systematically related to income. They find that this is to some degree the case for the levels estimator, but not for the changes-in-shares estimator.
- Main conclusions:
 - LC model provides much better fit than full risk-sharing.
 - However, there is no estimation that can fit both the distribution and changes in consumption shares very well at the same time. Potential reasons:
 - Utility is not CRRA
 - Household should not be the unit of analysis, but individuals within the household because of intra-household bargaining issues etc.
 - Introduce savings to the model. This sophisticates the model considerably but may be worth it. Basic trade-off: More risk reduction, but less risk sharing
 - The interaction with formal / government sponsored schemes may be worth looking into. Existing research suggests that while such schemes may help insuring against aggregate village-level risk, they may have a negative impact on the extent of private risk-pooling since they increase autarky utility.

Karlan et al. (2014). Agricultural Decisions after Relaxing Credit and Risk Constraints. QJE.

- RCT in northern Ghana with different insurance and cash treatment to investigate the effect on farming decisions.
- Piloting revealed that not only lack of capital but also risk of unpredictable rainfall seemed to be core reasons why farmers did not intensify farm investment
- three treatment arms and control:
 - cash grants (117 HHs)
 - grants or access to purchase rainfall index insurance (as opposed to crop insurance which would create moral hazard and adverse selection issues since rainfall is observable and not influenceable by the farmer) (125 HHs)
 - both (95 HHs)
 - control (155 HHs)
- Theoretically, the responses to cash grants and insurance grants may be different depending on whether capital markets or risk markets are missing or not

- KOO explore this in a model (details in the online appendix). There are two time periods 0 and 1, two states of the world $s \in \{G, B\}$, a cash grant k paid in period 0, insurance payouts k_s paid in period 1 and farmers can invest in $x = (x_r, x_h)$, a risky and hedging asset, as well as save amount a with interest R for the next period. The production function with x depends on the state of the world:

$$f_s(x) = \begin{cases} f_G(x_r) & \text{if } s = G \\ f_B(x_h) & \text{if } s = B \end{cases}$$

That is, in the good state of the world, the marginal return to the hedging assets is zero and in the bad state of the world, the marginal return to the risky asset is zero

- Farmers may also engage in perfect *risk pooling* within the village (if risk markets are *not* missing) which means that they always actually consume the expected consumption in period 1 and pay transfers to each other depending on their state realisation.
- They solve:

$$\begin{aligned} & \max_{x_r, x_h, a} u(c^0) + \beta \sum_s \pi_s u(c_s^1) \\ \text{s.t. } & \begin{cases} c^0 = Y - x_r - x_h - a + k, \quad a \geq 0 & \text{if credit markets missing} \\ c^0 = Y - x_r - x_h - a + k & \text{if credit markets not missing} \end{cases}; \\ & \begin{cases} c_s^1 = f_s(x) + Ra + k_s & \text{if risk markets missing} \\ c_G^1 = c_B^1 \equiv c^1 = \sum_s \pi_s (f_s(x) + Ra + k_s) & \text{if risk markets not missing} \end{cases} \end{aligned}$$

where obvious notations are used and $\beta R = 1$ is assumed.

- Case 1: credit markets and risk markets – equilibrium:

$$\begin{aligned} (1) \quad & u'(c^0) = \beta(\pi_G)^2 f'_G(x_r) u'(c^1) \\ (2) \quad & u'(c^0) = \beta(\pi_B)^2 f'_B(x_h) u'(c^1) \\ (3) \quad & u'(c^0) = \beta R u'(c^1) = u'(c^1) \end{aligned}$$

- Separation result holds: In the interior solution, the production result on x is completely pinned down by the prod. function f and probabilities π . It is independent of preferences, income, and in particular the credit grant and insurance grant, that is:

$$\frac{\partial x_r}{\partial k} = \frac{\partial x_h}{\partial k} = \frac{\partial x_r}{\partial k_s} = \frac{\partial x_h}{\partial k_s} = 0 \quad \forall s$$

- Case 2: no (incomplete) credit markets, but risk markets – equilibrium:

$$(1), (2), (3^*): u'(c^0) \geq u'(c^1)$$

- In this case, it can be shown from (1) and (2) that

$$\frac{\partial x_r}{\partial k}, \frac{\partial x_h}{\partial k} > 0 > \frac{\partial x_r}{\partial k_B}, \frac{\partial x_h}{\partial k_B}$$

- That is, the cash grant increases investment while the insurance payout decreases it. Intuition: The cash grant increases wealth today and thus the marginal value of future consumption leading to more investment. The insurance grant in-

creases payouts and expected wealth tomorrow and thus decreases its marginal value, thereby decreasing investment today. If both increase, the net effect is probably positive since the cash grant's expected value in their context is considerably larger than the insurance grant's one.

- Case 3: No (imperfect) risk markets, but credit markets – equilibrium:

$$(1') u'(c^0) = \beta \pi_G f'_G(x_r) u'(c_G^1)$$

$$(2') u'(c^0) = \beta \pi_B f'_B(x_h) u'(c_B^1)$$

$$(3') u'(c^0) = \beta R \sum_s \pi_s u'(c_s^1) = \sum_s \pi_s u'(c_s^1)$$

- It can be shown that:

$$\frac{\partial x_r}{\partial k} \cdot \frac{\partial x_r}{\partial k_B} > 0 > \frac{\partial x_h}{\partial k} \cdot \frac{\partial x_h}{\partial k_B}$$

- That is, a cash or insurance grant increases risky investment and decreases hedging investment where the effect of credit depends on the utility function. It is as described for DARA, but there is *no* effect for CARA. Intuition: Insurance response is standard insurance effect – invest in more risk. Credit response comes through a wealth channel in DARA: If more wealthy, go for more risk
- Case 4: No (imperfect) risk markets, no (incomplete) credit markets: The effect of binding capital constraints dominates, so that the predictions are the same as in case 2

TABLE I
SUMMARY OF IMPLICATIONS OF MARKET IMPERFECTIONS

		Market environment		Predicted change in investment					
		Perfect capital markets		Capital grant treatment only		Insurance grant treatment only		Capital & insurance grant treatment	
		Perfect capital markets	Perfect risk markets	Risky asset	Hedging asset	Risky asset	Hedging asset	Risky asset	Hedging asset
1	Yes	Yes	0	0	0	0	0	0	0
2	No	Yes	++	++	-	-	+ ^a	+ ^b	
3	Yes	No	+ ^c	- ^d	++	--	++	--	
4	No	No	+	+	-	-	+	+	

Notes. ^aThe model prediction is ambiguous, but in practice in our experiment the expected value of the insurance treatment was considerably smaller than the value of the cash grant, thus the net predicted effect in our setting is positive. ^bThe model prediction is ambiguous, but in practice in our experiment the expected value of the insurance treatment was considerably smaller than the value of the cash grant, thus the net predicted effect in our setting is positive. ^cSmall and positive via wealth effect, if DARA; zero if CARA. ^dSmall and negative via wealth effect, if DARA; zero if CARA.

- Results are largely consistent with imperfect risk markets; insurance generally tends to lead to a portfolio with more risky and less safe assets.
- Insurance take-up is quite sensitive to the price but still on the order of 50% at the actuarially fair price (GHS 6-9) [left]. However, the coverage, conditional on buying insurance, is quite low at the actuarially fair price [right]. The latter fact could be optimal or just mean people want to try the product because they don't really trust it.

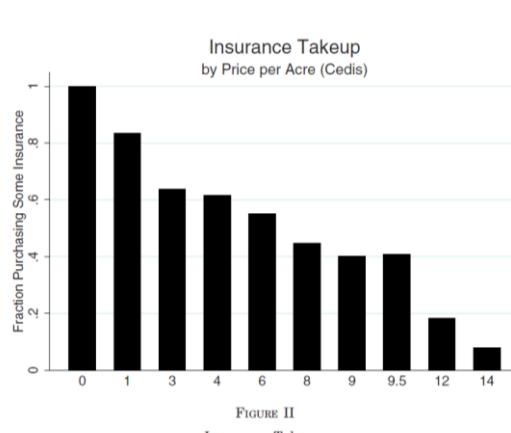


FIGURE II
Insurance Take-up

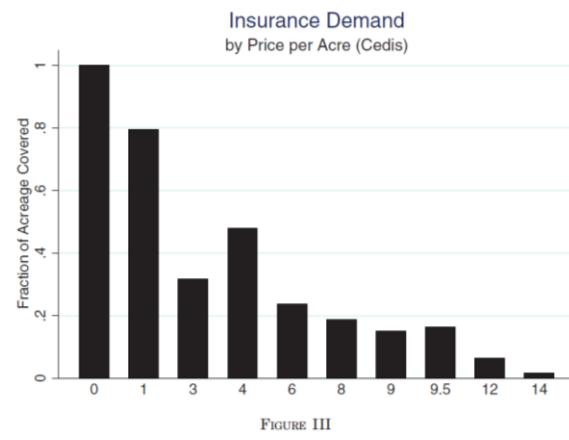


FIGURE III
The Demand for Acres Insured

- Insurance demand is sensitive to social network experience with insurance (trust as a key issue) and the number of payouts. "Irony": Insurance offers largest benefits for low-probability high-loss events, yet rare payouts harm demand.
- Little effect of cash-only effect
- Large effect of insurance provision at actuarially fair price: This seems to be a main hindrance to investment. When insured, farmers still find money to invest in agricultural activity, so capital constraint do not seem to be as binding as typically thought.
- People also actually seem to engage in risk taking. The output value curve with respect to rain is steeper for the insured, meaning they go for high yield high risk crops knowing that an insurance will pay out in case of bad weather. In table IV and V and Fig. I, land preparation cost, total cost # area cultivated, chemicals used and proportion of maize crops are riskier (more rain-sensitive) investments in their context while non-farm activity and orchard income would be the hedging (less rain-sensitive) activity

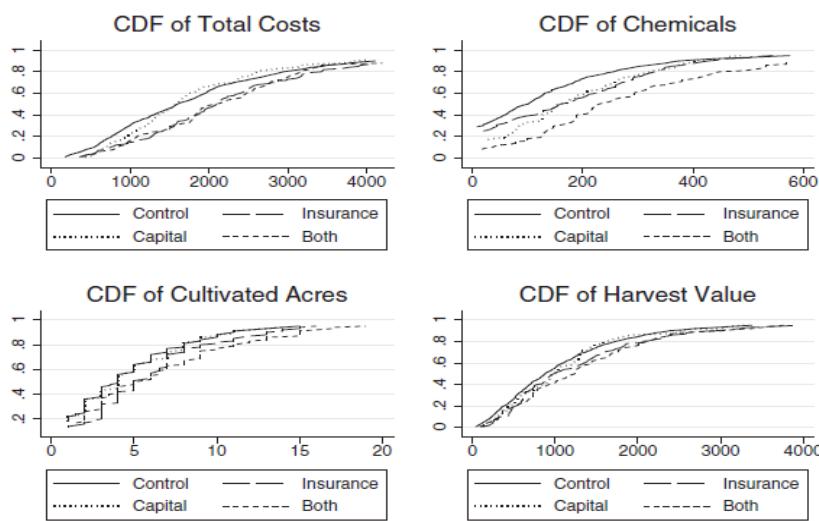


FIGURE I
Effect of Insurance and Cash Grants on Investment and Output

TABLE IV
IMPACT ON INVESTMENT AND HARVEST (INSTRUMENTAL VARIABLES)

Dependent variable:	(1) Land preparation costs	(2) # of Acres cultivated	(3) Value of chemicals used	(4) Wages paid to hired labor	(5) Opportunity cost of family labor	(6) Total costs	(7) Value of harvest
Insured	25.53** (12,064)	1.02** (0.420)	37.90** (14,854)	83.54 (59,623)	98.16 (84,349)	266.15** (134,229)	104.27 (81,198)
Insured * capital grant treatment	15.77 (13,040)	0.26 (0.445)	66.44*** (15,674)	39.76 (65,040)	-52.65 (86,100)	72.14 (138,640)	129.24 (81,389)
Capital grant treatment	15.36 (13,361)	0.09 (0.480)	55.63*** (17,274)	75.61 (68,914)	-130.56 (92,217)	2.44 (148,553)	64.82 (89,764)
Constant	169.38*** (10,603)	8.12*** (0.399)	171.70*** (13,804)	201.88*** (45,383)	1,394.58*** (84,786)	2,033.11*** (124,294)	1,417.52*** (90,635)
Observations	2,320	2,320	2,320	2,320	2,320	2,320	2,320
R-squared	0.017	0.143	0.041	0.005	0.006	0.009	0.012
Mean for control	189.1	5.921	158.3	327.9	1,302	2,058	1,177
Chi ² test of insured and insured + capital grant treatment	8.889	7.125	36.15	3.136	0.239	5.091	6.618
p-value	.003	.008	.000	.077	.625	.024	.010

TABLE V
REALLOCATION OF INVESTMENTS (INSTRUMENTAL VARIABLES)

Dependent variable:	(1) Value of harvest	(2) Proportion of land planted with maize	(3) Average weekly orchard income	(4) Household has nonfarm income generating activity (binary)	(5) # of HH members working in nonfarm income generating activity	(6) Average weekly enterprise income
Insured	-1,069.13* (596,208)	0.09*** (0.031)	-1.59* (0.876)	-0.06* (0.033)	-0.11* (0.061)	-8.64 (7,151)
Insured * capital grant treatment	1,324.48 (821,152)	0.04 (0.029)	0.65 (0.776)	0.07** (0.033)	0.16** (0.062)	3.77 (9,126)
Capital grant treatment	-879.77 (642,233)	0.12*** (0.034)	-0.19 (0.926)	-0.04 (0.038)	-0.08 (0.066)	-2.83 (4,530)
Insured * total rainfall	156.82** (76,291)					
Insured * capital grant treatment * total rainfall	-155.36 (105,649)					
Capital grant treatment * total rainfall	124.95 (83,589)					
Total rainfall (hundreds of millimeters)	2,247.39*** (624,545)					
Total rainfall squared	-146.65*** (40,970)					
Constant	-7,154.76*** (2,375,086)	0.23*** (0.016)	2.42*** (0.613)	0.17*** (0.027)	0.22*** (0.038)	5.79 (4,363)
Observations	2,320	2,782	2,316	2,320	2,320	2,350
R-squared	0.021	0.090	0.001	0.007	0.010	0.007
Chi ² test of joint effect of insurance and insurance + capital	0.138	15.52	0.906	0.132	0.388	0.449
p-value	.710	8.16e-05	.341	.717	.534	.503
Mean for control	1177	0.309	2.587	0.261	0.405	6.604

- Calculating the value of harvest effect in Table V for average rainfall gives a positive result, so it seems to be the case that the risky investment has a higher expected return. Also, other welfare measures show a positive impact:

TABLE VI
INCOME AND HOUSEHOLD WELFARE (INSTRUMENTAL VARIABLES)

Dependent variable:	(1) Total farm revenue (inc. insurance payouts, net of premiums)	(2) Postharvest assets (livestock + grain)	(3) Household reports having missed a meal in past 12 months (binary)	(4) Total expenditure in 12 months	(5) Utility expenses in past 12 months	(6) School expenses in past 12 months	(7) Borrowed in past 12 months from any source (binary)
Insured	284.98*** (82,991)	530.74** (230,839)	-0.08** (0.033)	46.39 (58,767)	0.36 (7,102)	-0.71 (15,872)	-0.00 (0.025)
Insured * capital grant treatment	109.13 (84,446)	310.66 (229,150)	-0.03 (0.030)	2.44 (58,568)	19.96** (8,444)	25.83 (16,111)	-0.13*** (0.033)
Capital grant treatment	66.93 (90,585)	606.12** (266,636)	-0.08** (0.037)	7.14 (61,540)	10.30 (8,268)	24.04 (18,841)	-0.06 (0.040)
Constant	1,386.17*** (91,209)	1,782.29*** (223,471)	0.37*** (0.035)	470.10*** (43,073)	37.72*** (5,768)	107.94*** (12,632)	0.46*** (0.035)
Observations	2,320	2,265	2,304	2,316	2,316	1,940	3,756
R-squared	0.023	0.007	0.013	0.015	0.050	0.032	0.203
Chi ² test of joint effect of insurance and insurance + capital	17.97	10.68	9.830	0.581	5.192	1.984	13.39
p-value	0.0000225	0.00108	0.00172	0.446	0.0227	0.159	0.000253
Mean for Control	1,179	1,756	0.229	585.6	41.93	115.2	0.313

Mobarak & Rosenzweig. Selling Formal Insurance to the Informally Insured. Bread WP.

- They investigate the link between formal insurance and informal insurance. Theoretically, they may be both substitutes and complements
 - substitutes: informal insurance may crowd out formal weather insurance if informal insurance covers weather
 - complements: since formal weather insurance entails a degree of basis risk (s. below), informal insurance may step in to cover this basis risk. This is more in line with Karlan et al. (2014)
- Index insurance advantage: moral hazard and adverse selection problems mitigated since payout depends on *exogenous* and *publicly observable* index. Disadvantage: basis risk through the potential mismatch between index realisation and actual loss incurred (eg. weather insurance: loss incurred is not perfectly proportional to rainfall observed at the weather station, especially if the weather station is far from your farm)
- They consider the Indian subcaste *jati* and, using rich data on a sample, they estimate the degree of informal insurance individuals are involved in
 - Estimate the amount of transfers that arrives as a response to shocks from family and caste members and construct an index: η for idiosyncratic shocks and ι for aggregate village-level shocks
- They then randomise the offer of formal index insurance and – within the treatment group – further randomise the placement of automatic rainfall stations. Basis risk can be proxied by the distance to the rainfall station. This setting allows them to address several questions:
 - how does informal risk sharing affect formal insurance take-up?
 - how does index insurance affect risk taking?
 - how does basis risk affect index insurance demand?
 - how does the effect of basis risk on index insurance demand interact with informal insurance arrangements?
- The proxy for basis risk seems to work – log output value per acre *decreases* in the distance from the rainfall station for a given amount of rainfall (less insured people):

Table 5
Testing for basis risk: effects on log of output value per acre

Variable	Uttar Pradesh		Two States (UP and AP)	
Rain per day	0.16516 [1.32]	0.30169 [2.2]	0.13937 [1.4]	0.23606 [2.09]
Distance to aws (km)		0.12483 [2.4]		0.08460 [1.92]
Rain per day x Distance to aws		-0.02231 [3.53]		-0.01673 [2.81]
N	945	936	1,459	1,418

Absolute values of t-ratios in brackets, clustered at the village level. Two state regressions include state fixed effects. Sample includes all cultivators.

- [note that strangely the coefficient on distance to AWS is significant although the placement ought to be randomised]
- Theory:

- When individuals in a group face both idiosyncratic and aggregate risk, informal networks lower the demand for formal insurance only if the network indemnifies against aggregate risk.
- When the formal insurance product is imperfect due to mismatches between the rainfall-index-based payouts and the actual losses incurred by the policy holders (basis risk), however, informal risk sharing, by covering household losses that are the consequence of basis risk, may enhance the benefits from formal index insurance contracts
- Empirical results are in line with the model.
 - Distance to automatic weather stations (AWS) is an impediment to take-up of formal insurance (around 2 ppt decrease per km). This effect is attenuated if informal insurance is successful in covering individual losses.
 - Those already informally providing insurance coverage for aggregate shocks are less likely to purchase the index product: per sd increase in ι , take-up decreases by 3.6 ppt or 9% (no idea how they compute this number, ι might be measured on a weird scale...)

Table 6
Fixed-Effect Estimates: Determinants of Formal Insurance Take-up

Variable/Est. Method	Three States		Two States (AP and UP)		
	FE-State		FE-Caste		
η_i	0.125 [0.56]	0.151 [0.61]	0.0228 [0.07]	-	-
$\eta_i \times$ Distance to aws	-	-	0.151 [3.42]	0.139 [2.55]	0.157 [2.31]
ι_j	-198 [1.71]	-209.6 [1.28]	-209.7 [0.94]	-	-
$\iota_j \times$ Distance to aws	-	-	-	-	-18.6 [-0.528]
Distance to aws (km)	-	-	-0.0254 [3.50]	-0.0246 [2.63]	-0.019 [1.50]
Agricultural laborer	-0.0343 [2.19]	-0.0341 [2.13]	-0.028 [1.58]	-0.0238 [1.49]	-0.0379 [1.43]
Agricultural laborer \times Distance to aws	-	-	-	-	0.00333 [0.797]
Actuarial price	-0.00143 [2.07]	-0.00159 [2.07]	-0.00167 [2.40]	0.00154 [2.14]	0.00157 [2.14]
Subsidy	0.389 [3.38]	0.355 [2.86]	0.35 [3.10]	0.376 [3.26]	0.372 [3.20]
Owned land holdings	0.000405 [0.14]	0.000445 [0.14]	0.000648 [0.20]	0.00353 [1.42]	0.0035 [1.42]
Village coefficient of variation, rainfall	0.523 [2.16]	0.751 [2.89]	0.747 [2.77]	0.874 [2.92]	0.908 [3.04]
N	4,260	3,338	3,338	3,338	3,338

Absolute values of t-ratios in brackets, clustered at the caste level. Standard errors are bootstrapped to account for the fact that η_i and ι_j are estimated regressors. Specifications also include scheduled tribe or caste indicator and whether non-Hindu

- They also find that farmers with insurance seem to substitute away from drought-resistant crops towards high yield crops in their portfolio, ie. they do engage in more risk-taking. Consistent with general theory on insurance and Karlan et al. paper.

Table 10
Intent-to-Treat Caste Fixed-Effects Estimates of Index Insurance on Risk and Yield:
Proportion of Planted Crop Varieties Rated "Good" for Drought Tolerance and Yield,
Tamil Nadu Kharif Rice Farmers

Crop Characteristic: Variable	Good Drought Tolerance		Good Yield (1)
	(1)		
Offered insurance	-0.0593 [2.67]	0.0519 [1.93]	
Owned land holdings	0.0000934 [0.02]	0.00056 [0.12]	
Village coefficient of variation, rainfall	0.351 [0.88]	-0.516 [0.81]	
N	325	325	

Absolute values of t-ratios in brackets, clustered by caste/village (because the randomized insurance treatment was stratified at the caste/village level).

18 Psychology and Development

18.1 Behavioural Economics

- Behavioural Economics studies departures from the standard rationality assumption. Although it is obvious that it doesn't quite apply, the focus is on **whether this has any implications** (that's not so clear). Study if this has implications for development
- Examples of non-rational behaviour observed include
 - Endowment Effect**: people's preferences depend on **reference points**. Depending on status quo and expectations, people will have different utility changes for the same gain/loss. Empirically, these differences in utility observed at small differences in status quo or expectations cannot be explained by risk theory assuming reasonable risk aversion. This goes **against the Coase theorem** which says that the allocation of resources will *not* depend on the initial assignment of property rights.
 - Example: Mug Experiment (Kahnemann et al., 1991)
 - Gift Exchange** (List, 2006): A buyer makes a price offer to the seller. The seller can decide whether to accept or not and the quality at which she delivers where quality is costly (SPE: seller accepts any positive price and chooses lowest possible quality, buyer chooses lowest possible price). Finding: some sort of **reciprocity** – quality provided increases in the price offered.
 - Uncertainty Effect** (Gneezy et al., 2006): People experimentally value lotteries over some (different) positive outcomes less than getting the worst outcome for sure (which should be strictly dominated by the lottery).
 - Relative price effect** [my name] (Tversky & Kahnemann, 1981): People value price differences relative to the original price. Empirically, most people are willing to walk 20 minutes to save USD 5 on a USD 20 calculator and *not* willing to walk 20 minutes to save USD 5 on a USD 200 jacket.
- Critiques** of Behavioural Economics:

- Results often only hold for **small stakes**
- Results often have **undergraduates** as test subjects
- Some of them are **consistent with non-behavioural stories**
- Results have **no meaningful implications**
- **Markets** can fix the problem of seemingly individual irrationalities
- **Lab** outcomes are **different** from the real world
- However, some models do have some explanatory power addressing **weird behaviour** by poor people we observe. Examples:
 - **No investment on a high-return technology**, such as fertiliser (addressed by Duflo et al., 2011; see 6.2)
 - **Poverty situations** out of which agents should be able to save themselves out quickly – *The Fable of the Market Woman*: Karlan & Mullainathan (?) find the following problem in India and the Philippines: Women take out a loan for 1000 Rupees (around USD 20) every morning to buy vegetable from a wholesaler and sell it on the street to make revenue of 1200R. They have to repay the credit on the same day at 10% which yields a profit of 100R. The credit is incredibly expensive and being able to invest 1000R herself, the woman could double her daily income (and then expand). Saving 1R per day and using this for own investment continuously would get women out of the necessity to borrow within 49 days (because of the strong compound interest rate effect)! In an **experiment**, the authors have two treatments: One-time debt repayment (enabling the women to invest herself from tomorrow, not having to save and reinvest 49 days) and training. **Training has no significant effect** while **debt repayment does reduce borrowing**, but **not** in the “long”-run (couple of months later).
 - The treatment effect show that savings constraints cannot be the reason as it takes (much more) time to reduce such a bulk of saving if having lived on 100R per day before.
 - 6.3 and 6.4 have behavioural models which can address this.

18.2 Time Inconsistency

- Time inconsistency arises when people make **different decisions over future consumption today than they would make in the future**. More formally, you are *time consistent* if you make the same choice for time period $t+x$ in time period t as in time period $t+y$ for any y .
- Typical formulation:

$$U(c_1, \dots, c_T) = \sum_t D(t) u(c_t)$$

It can be shown that behaviour is *time consistent* iff $D(t) = \delta^t$ for $\delta \in (0,1)$. Sufficiency is easy to see. Consider the choice between two time period j and $j+x$ that you make in any time period t . The relevant expression is

$$MRS_{j,j+x}^t = \frac{D(j-t) \frac{\partial u}{\partial c_j}}{D(j+x-t) \frac{\partial u}{\partial c_{j+x}}} = \delta^x \frac{\frac{\partial u}{\partial c_j}}{\frac{\partial u}{\partial c_{j+x}}}$$

Which is clearly independent of t

- Typical formulation for **present bias**:

$$U(c_1, \dots, c_T) = u(c_t) + \beta \sum_{\tau \geq t+1} \delta^{\tau-t} u(c_\tau)$$

Clearly, here:

$$MRS_{j,j+1}^{t \neq j} = \frac{\frac{\partial u}{\partial c_j}}{\delta \frac{\partial u}{\partial c_{j+x}}} \neq MRS_{j,j+1}^j = \frac{\frac{\partial u}{\partial c_j}}{\beta \delta \frac{\partial u}{\partial c_{j+x}}}$$

Duflo et al. (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. AER.

See also EC535 presentation on this.

- DKR look at **fertiliser use in Kenya** and provide a model with present biased farmers explaining no investment and procrastination.
- The Savings and Fertiliser Initiative (**SAFI**) programme offered free delivery on the purchase of fertiliser **immediately after harvest** (when farmers had high income)
- Drawing on the insights from behavioural economics, this addressed two points:
 - Farmers were to make a **decision when being relatively rich** and not impeded by the “exhaustion of poverty”
 - Farmers were offered a **simple option** and **taken away the cost of going to the shop**, thinking through the products, etc.
- Present-biased farmers would buy the product immediately after harvest because the same discount today is more worth than tomorrow as opposed to rational farmers who would buy it only immediately before they need the fertiliser
- They find a **significant increase** in the **use of fertiliser** through SAFI.
- Conclusion: A **paternalistic libertarian** policy such as offering a **time-limited discount** such as SAFI may be welfare-increasing.
- Principally, for perfectly rational farmers, a subsidy is only distortive.

18.3 Temptation

- Consider the following model of **accumulation**: agents with income y in two periods and investment possibility I yielding return RI maximise $u(y - I) + \delta u(y + RI)$. So:

$$u'(y - I^*) \geq \delta R u'(y + RI)$$

holding with strict equality if $I^* \neq 0$, i.e. $\delta R > 1$ (for $I^* = 0$, FOC yields $1 \geq \delta R$)

- In an infinite period setting in which the agent lives from investment return from period 1 onwards after a one-off income in period 0 with $R = f'(I)$ for concave f such that $f'(I) < 1/\delta$ for some I and $f'(I) \rightarrow \infty$ for $I \rightarrow 0$, the agent's FOC becomes:

$$u'(c_t) = \delta f'(I_t^*) u'(c_{t+1})$$

- For $\delta R = \delta f'(I) > 1$ and concave u , $c_{t+1} > c_t$, i.e. consumption increases over time → there are no poverty traps regardless of the initial level of wealth. This essentially becomes from **investment returns becoming very large as investment becomes very small**.

- Banerjee & Mullainathan (2010) develop a **temptation** model. Consider two selves in two time periods consuming a regular good x and a temptation good z with

$$U(x_2, z_2) = u(x_2) + v(z_2)$$

$$U(x_1, x_2, z_1, z_2) = u(x_1) + v(z_1) + \delta u(x_2)$$

- The temptation good in period 2 is worth nothing to the agent today. For income y in both periods and investment opportunity I , the agent maximises $u(y - z_1 - I) + v(z_1) + \delta u(y + RI - z_2(y + RI))$ over I and z_2 where z_2 is the solution to $u'(x_2(c_2)) = v'(z_2(c_2))$ from the maximisation in the second period (obviously $c_1 = y - I$ and $c_2 = y + RI$). This gives the FOCs:

$$u'(x_1) \geq \delta R(1 - z'_2(c_2))u'(x_2)$$

$$u'(x_1) = v'(z_1)$$

$$BCs: x_t + z_t = c_t, t = 1, 2$$

- For the corner solution $I = 0$, the second period solution and the FOCs give $x_1 = x_2$ and therefore $1 \geq \delta R(1 - z'_2(y))$ which includes a "**temptation tax**" $z'_2(y)$ now!
- For concave z , a **poverty trap can arise** as this admits that $\exists \hat{y}$:

$$1 > \delta R(1 - z'_2(y)) \quad \forall y < \hat{y}, i.e. I^* = 0$$

$$1 < \delta R(1 - z'_2(y)) \quad \forall y > \hat{y}, i.e. I^* > 0$$

- The **concavity of z_2** can be questioned here. Otherwise this is a nice way of thinking about temptation goods and people **not saving because they are afraid they are going to "waste" the gained money tomorrow.**

18.4 Scarcity

- Main idea: **Attention, concentration** and **self-control** are **scarce resources**. Poor people are constantly thinking hard about subsistence, (even small) consumption decisions, etc. that they do not make intelligent or rational decisions when it comes to investment but behave affectively.
- Trade-off experiment** shows that low income people think much harder about buying small articles (e.g. a toaster, mixer) in terms of what consumption they would forego instead.
- Shiv and Fedorikhin (1999) made an experiment where half the subjects had to remember a **7 digit number** while the other half had to remember a **2 digit number** while walking between two rooms. On the way, they were asked to **choose between a chocolate cake and a fruit salad**
 - Assumption: affective, less cognitive, behaviour would induce chocolate cake choice which people might impulsively crave but not like on dietary rational grounds.
 - Finding: those who had to remember the 7 digit number went much more often for the chocolate cake (as an aside: this result is only found when the snacks are presented as real products, not only as products, which reveals another fact about how our brain works)

- Interpretation: the **7 digit guys had lower processing resources available and went for the more affective less cognitive choice**. Of course, as the experiment shows, this is also heavily influenced by the type of person.
- Translation into the scarcity argument: **poverty is like remembering a 7 digit number**
- Mani et al. (2014) find that there is a significant **difference in the performance** in two **intelligence tests for poor people** between those (random allocation) asked to think about where to get **USD150** from for a car repair relative to where to get **USD1500** from for a car repair (the latter performing much worse). There is **no significant difference** for the two treatments among **rich** people.
- **Caveat** for many experiments in this area: There is **no good record of replication success!**

19 Transfers (UCTs, CCTs, IKTs)

19.1 General Summary

- Basic argument in favour of unconditional cash transfers (UCTs) as opposed to conditional cash transfers (CCTs) and in kind transfers (IKTs): money is fungible and recipients take the optimal decision what to use the money for. CCTs and IKTs will necessarily lead to a sub-optimal solution with a distorted consumption decision
- Basic argument in favour of CCTs and IKTs as opposed to UCTs: behavioural reasons – recipients of transfers do not take the optimal decision as to what to consume because of behavioural poverty traps, myopia, parents not fully internalising the optimality for their children (particularly relevant for schooling decisions), etc.
 - E.g. temptation goods

19.2 Papers

Evans & Popova (2014). Cash transfers and temptation goods. A review of global evidence. WB Policy Research Paper.

- Review literature on the impact of transfer programmes on tobacco and alcohol expenditure, considering studies carried out between 1997 and 2014.
- Almost without exception, across 44 estimates in 19 studies, no significant or significantly negative impact of transfers on expenditures on tobacco and alcohol found (either level or proportion of spending)
 - The only two studies finding a positive significant effect have other estimates within the same papers that are inconsistent with the positive finding, the findings are only weakly significant and the magnitude is trivial.
 - For levels only, there are again only two studies finding positive significant effects (should be the same two studies therefore)
 - Among RCTs only, no studies find significant positive effects
- Several effects at play with transfers:
 - Income effect
 - Substitution effect under conditionality: e.g. conditioning on schooling makes schooling more valuable (not only the private returns to schooling, but also the future income returns from satisfying the condition)

- Flypaper effect: No matter whether CCT or UCT, the money is either earmarked or advertised for something. Advertising alone in the case of UCTs can lead to spending where spending is supposed to happen
- Household bargaining effects if differential bargaining power between men and women and money is given to a specific gender only
- Concerns:
 - Systematic under-reporting? / "social desirability bias"

Bastagli et al. (2016). *Cash transfers: What does the evidence say? A rigorous review of programme impact and of the role of design and implementation features.* ODI report.

- Concise executive summary on several outcomes: monetary poverty, education, health and nutrition, savings / investment / production, employment, empowerment

Outcome	Effect	Gender diff.?	HH head diff.?
Education	Mostly positive on school attendance, less clear on learning outcomes	Effects on girls found too	No clear pattern
Health and nutrition	Greater effect on use of health service and dietary diversity, smaller on anthropometric measures		Mixed results
Savings / investment / production	Positive impact on savings, impact on borrowing mixed. Only some studies have impact on agricultural productive assets, most find an increase in livestock ownership and value. Mixed evidence on the impact on business engagement and expenditure.		Three studies find significant impact only for female HH head. Two find different sorts of investment.
Employment	More than half the studies find no sign. Effect on external and internal margin. Those finding one find a positive effect. For child labour, more than half don't have significance. The significant ones are all reductions	Women tend to increase domestic work, men different work.	
Empowerment	Seems that physical abuse of women is reduced, but non-physical abuse (controlling behaviour) may increase. Mechanisms: lower stress-related abuse and bargaining out of abuse. Reduced risky sexual behaviour and early marriage.		

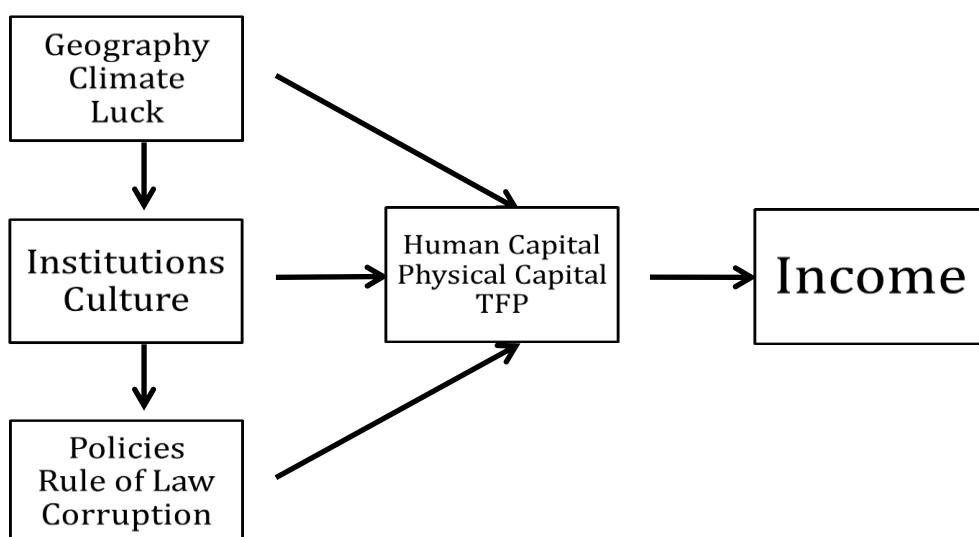
- For most indicators among studies varying the gender of the main recipient, there is *no* difference.
- Effect with regards to timing of transfer found, e.g. if around the time that school fees have to be paid, more likely that education investment is greater
- Conditionality (expectedly) gives greater impact on the expenditures that are conditioned on (e.g. education, health).
 - However, is this optimal behaviour?
- Complementary interventions (e.g. training) strengthen the intended impacts

Banerjee et al. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. Science.

20 Institutions and Development

20.1 The Importance of Institutions

- **Definition:** “Institutions are the rules of the game in a society or, more formally, are the human devised constraints that shape human interaction” ... “in consequence they structure incentive in human exchange, whether political, social, or economic.” (Douglas North, 1990)
 - **Economic institutions** include **property rights, markets, monopolies** etc.
They are **determined by political power**
 - **Political power** in turn is **de jure** determined by institutions such as **democracy, constitution**, etc. **De facto**, some argue, it may be **guns, tanks and money**
- An overview of causes of (under)development could look like this:



- **Institutions hypothesis:** institutions are primary fundamental cause of underdevelopment (mainly supported by Acemoglu, Robinson & Johnson in different papers and books)
 - Countries with **extractive institutions** do badly
 - Countries with **inclusive institutions** do well
- How institutions are determined by the **long arm of history**. At some point in colonial history, institutions were set up and **remain sticky**.
 - Institutions are sticky because the **political Coase theorem does not hold**. This is mainly because of a **lack of commitment**
- Stylised facts on the importance of institutions
 - There is a positive correlation between corruption level of a state and NYC violations per diplomat
 - North Korea's GDP is much lower than South Korea's although these countries are very similar in a lot of regards – one major difference: institutions

20.2 Political Coase Theorem and Commitment

- **Political Coase Theorem:** It should be possible to trade institutions / institutional setting in order to achieve maximal efficiency, i.e. there should be a **convergence in institutions**
 - Modified: since countries have different beliefs on what are efficient institutions, there may be some variations in institutions
- If institutions are main driving force for development and we observe **great development differences**, the political **Coase theorem obviously fails**.
- A reason may be a **lack of commitment**: A **farmer** and a **king** live in a country and the farmer chooses effort $e \in [0,1]$ (that is, his cost) to produce output A with probability \sqrt{e} and 0 otherwise, i.e. expected output $A\sqrt{e}$.
 - The (first best) efficient level of effort in this setting follows from $\max_e A\sqrt{e} - e$ and is $e = \left(\frac{A}{2}\right)^2$ and with efficient level of output $x = \frac{A^2}{2}$.
 - This could be achieved if the King can *commit* to a lump-sum tax $T \geq \frac{A^2}{4}$
 - The optimal tax *rate* under commitment is:
$$\tau^* = \arg \max \tau(x(\tau) + z) = \arg \max \tau \left(\frac{(1-\tau)A^2}{2} + z \right) = \frac{A^2 + 2z}{2A^2} \geq \frac{1}{2}$$

where x follows from the modified farmer's maximisation and z are natural resources

- In a setting without commitment, the farmer chooses effort first, then the king chooses a tax and the gains are realised.
 - **SPE: Tax = 100%, effort = output = 0.** Pareto inferior to the commitment solution.
- In an **infinite setting without commitment**, a tax rate $\hat{\tau} < 1$ may be one SPE following a **Grimm trigger strategy** by the farmer, i.e. in case of a King's deviation from $\hat{\tau}$, punish by setting $e = 0$ forever after (the punishment outcome off the equilibrium path is a NE, so no profitable deviation for anyone). On the equilibrium path, the farmer maximises utility, so he has no incentive to deviate. The king has no incentive to deviate iff his payoff from deviation V^d is smaller than staying in the equilibrium V^p . By one deviation principle, check a one-time deviation: $V^p(\hat{\tau}) = \frac{\hat{\tau}Y(\hat{\tau})}{1-\delta} \geq Y(\hat{\tau}) + \frac{\delta}{1-\delta}z = V^d(\hat{\tau})$ which yields:

$$\hat{\tau} \geq (1 - \delta) + \delta \left(\frac{z}{z + x(\tau)} \right)$$

- Observations:
 - This is **more easily fulfilled for high δ** . This parameter can be interpreted as the **stability of the environment** or how much rulers care about successors. **Stable dynasties** should therefore yield commitment outcomes.
 - The **higher natural resources z relative to farmer's output x , the higher the tax and the harder to fulfil the constraint** (effectively less power for farmer). This may be an interpretation of the resource curse.
 - The model **leaves out a paranoia effect**: rulers may fear too much power in the hands of the farmers if the pie gets bigger and bigger.

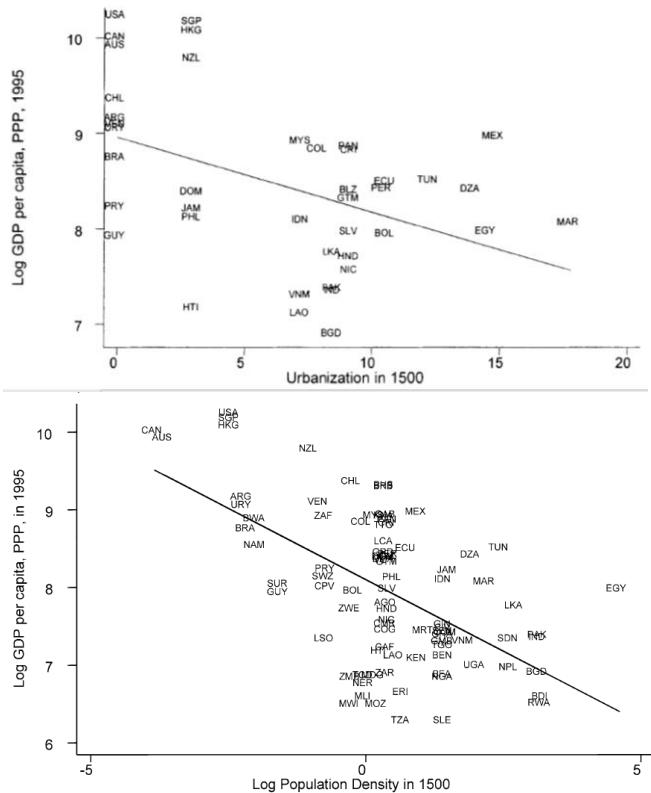
- Extensions to **inclusive institutions**: Suppose the king cares about the farmer and includes the farmer's utility with weight λ in his own utility function. This may **capture cultural affinity or democracy** (government by the people). Ignoring z , the constraint to ensure no deviation by the king in the SPE now follows from $V^p(\hat{\tau}) = \frac{(\hat{\tau} + \lambda(1 - \hat{\tau}))Y(\hat{\tau}) - \lambda e(\hat{\tau})}{1 - \delta} \geq Y(\hat{\tau}) - \lambda e(\hat{\tau}) = V^d(\hat{\tau})$ which yields (not immediately):

$$\hat{\tau} \geq 1 - \frac{\delta}{1 - \lambda \left(1 - \frac{\delta}{2}\right)}$$

- In comparison to the solution without λ and ignoring z above, i.e. $\hat{\tau} \geq 1 - \delta$, this is **more easily fulfilled** and yields lower taxes. **The higher λ , the stronger this effect.**
- Note that λ **changes nothing for the static problem**. Maximising $(\tau + \lambda(1 - \tau))A\sqrt{e} - \lambda e$ in a one-shot problem still gives 100% tax, no effort and no output as SPE.

20.3 Macro Evidence: Sticky Institutions vs. Geography

- The **geography hypothesis** postulates that differences in geography are the main cause for differences in income.
 - A positive correlation between latitude and GDP p.c. can be found but the same positive (and less variable) correlation can be found for institutions and GDP p.c.
- The institution holds the **reversal of fortune** (Acemoglu et al., 2002) **against the geography hypothesis**: In 1500, proxies for GDP (urbanisation, population density) were high for countries that are poor today and vice versa
 - Given geography is a constant feature of a country, this does not match.
 - Reversal around the 18th century for many countries suggests it has something to do with the technical revolution
 - Institutions hypothesis explanation: **Colonising powers set up extractive institutions in rich countries to get wealth and inclusive institutions in poor countries to build up wealth**. These institutions are **persistent** (sticky institutions).
- Acemoglu et al. (2001) look into sticky institutions by looking at the type of institutions colonial powers set up. For obvious endogeneity problems in determining their influence on development/income (reverse causality, OMV), they **instrument** institutions by the **settler mortality rate**.



- Argument: Settler mortality rate is a **proxy for the disease environment** which was a main **determinant** of whether settlers **planned to stay in the short or long run**. For the short term, extractive institutions would be the optimal choice while inclusive institutions would be the optimal choice for the short run.
- They find a **sign. pos. effect**, even **controlling for religion, latitude, current disease environment and size of the coastline** which all turn out to be without effect
- If we are to believe them, the numbers are **quite substantial**: Institutions matter a lot for long-run development. But...
- Concerns:
 - Still **extremely crude**. More of a correlation-based argument than in any way neat causal identification
 - **Data problem**: old data may be extremely bad
 - **Measurement of institutions**: they use average protection against risk of expropriation (instrumented by settler mortality rate), but you can think of so many variables. Glaeser et al. (2004) suggest to focus more on actual law, rules and compliance regulations.
- **Critique** (Glaeser et al., 2004): The main cause is human capital: **Human Capital → Growth → Institutions**.
 - Human capital is indeed something which may very well **violate the exclusion restriction** of the instrument (settler mortality rate driving both human capital and income/development)

20.4 Micro Evidence

Goldstein & Udry (2008): The profits of power: Land rights and agricultural investment in Ghana, JPE

- Goldstein and Udry (2008) explore **two** questions:
 - **What effect do tenure rights have on agricultural investment (fallowing durations)?**
 - **What effect does this investment have on profitability?**
- **Trade-off** by farmers: Longer fallowing is productively efficient but also increases the probability of losing the plot.
 - Strong tenure rights should decrease the probability of losing the plot
- Theory: Farmers maximize their expected present value of profits:

$$\max_{\{\tau_p\}_{p=1}^{\infty}} \sum_p \pi_p(\tau_p) \sum_{n=1}^{\infty} \left(\frac{1 - \omega_p}{1 + \rho_h} \right)^{\tau_p}$$

where p is a plot indicator, τ the fallow duration, ω the constant probability of losing a plot each period and ρ a household (h) discount factor. From the ln specification, this gives for a representative plot p :

$$\frac{\pi'_p(\tau_p^*)}{\pi_p(\tau_p^*)} = - \frac{\ln \left(\frac{1 - \omega_p}{1 + \rho_h} \right)}{1 - \left(\frac{1 - \omega_p}{1 + \rho_h} \right)^{\tau_p^*}}$$

- **The optimal fallow duration τ_p^* falls in ω_p and ρ_h**

- Econometrics: Tenure rights are **proxied by the social/political role** within the community (officeholders vs. non-officeholders). Tenure rights defined as such are taken as an **instrument for fallow durations** for the second question
- Findings:
 - **Officeholders fallow their plots for years longer** (optimal fallow duration said to be 7 years). This effect is even stronger for plots obtained through the traditional matrilineage allocation system instead of commercially
 - One more year of **fallowing increases profitability** substantially. Inefficient fallowing has a **cost** of around **1% of 1997 GDP**.
- **Concerns:**
 - **Reverse causality:** fallow duration may have increased the standing in the community and the likelihood of getting an office. However, there are no significant differences between **inherited and non-inherited offices**.
 - **ω may not be constant across time.**
 - **Exclusion restriction:** While tenure rights per se may be plausibly exogenous unless via fallow durations, social/political role, the proxy, may not, but influence fallow durations in different ways
 - **Skills** may be correlated with both the likelihood of holding office and profitability.
 - **Effort** may be another channel via which officeholders differ systematically from non-officeholders and which impacts on profitability (they may have less time precisely because of their role)
 - Through social/political role, plotholders may have had **better information** on optimal fallow durations and thus fallow longer (this has nothing to do with tenure rights).
 - **Lower opportunity cost of capital** and/or **bargaining process** between officeholders and non-officeholders within the same household. This is **controlled for**.
 - **External validity:** very specific setting in agriculture in a concrete place in Ghana and one specific measurement of tenure rights. Need further verification.

Acemoglu et al. (2014), *Chiefs: Economic development and elite control of civil society in Sierra Leone*, JPE

- They construct historical documentation of the chiefdoms in Sierra Leone and the institution of paramount chiefs for chiefdoms and record the number of ruling families in each chiefdom
- Arguing that the number of ruling families is quasi-random, they relate it to measures of current development and find that a greater number of ruling families leads to largely better development outcomes, namely literacy, educational attainment, the share of people working outside of agriculture, child health, asset wealth, housing quality.
 - Their interpretation is in an understanding of the number of ruling families capturing political competition and accountability. The political decision makers are more constrained institutionally if there are more ruling families competing for political power.

- For this interpretation of political power, they construct a Herfindahl concentration index (from IO literature on market concentration) using the number of chieftaincy seats and ruling families as input. The idea is to measure how concentrated the occupation of seats is with one or a few families. This index is indeed highly correlated with the number of ruling families.
- Interestingly, they find that fewer ruling families are *positively* related with measures of social capital (attendance of community meetings, participation in social groups, undertaking collective action) which goes against the standard understanding of more autocratic power and less accountability implying lower social capital in communities.
- They provide three strategies to support their exogeneity claim:
 - A case study of six chiefdoms documenting that the number of ruling families seems to result to a large extent from historical coincidence such as the availability of a male heir
 - they show that there is no correlation between the British colonial tax assessments in the 1890s (taken as a measure of the level of development) and the # of ruling families
 - they include some detailed geographic controls that don't change the results much
- For their main result, they control for amalgamation, age, age squared, gender, ethnicity, the number of chiefs that could be remembered in a chiefdom, district FE and report regressions with and without an additional six geographic controls (presence of mining permissions in the 1930s, distance to coast, distance to river, distance to historic trade route, distance to historic railroad and minimum distance to one of the big cities)
- Educational result: Moving from the bottom to the top quartile of the ruling family distribution (1.8 to 7.7 ruling families) increases the likelihood of being literate, having primary and secondary school education by about 7 ppt. These are all statistically significant.
- Health result: Same shift would lead to a reduction in the rate of moderate to severe anemia in children by about 13 ppt.
- For economic outcomes, they consider
 - Share in non-agric. Employment: Same shift leads to a 2.3ppt increase relative to a 11% base
 - Asset wealth: same shift leads to increase by about a fifth of a std. dev.
 - Housing quality: index goes up by 0.08 units (the index is the average of three binary questions and ranges in the unit interval)
- Property rights: More ruling families lead to reduced likelihood of having to ask the chief for permission to use land and an increased right among people to sell land (they even consider the interaction with whether one is a stranger in a community relating to a historic institution of them not being allowed to sell land)
- Social capital: As the number of ruling families increases, *fewer* people agree that authority should be respected more and that only older people can lead.
- For social capital, they also consider variables from the NPS on attendance of community meetings, membership in groups (ROSCAs, labor gangs, secret societies) and a collective action index (mean of binary whether someone has participated in road brushing or another community project). For all these, there is again a *positive* relationship between

having fewer ruling families (more concentration of political power) and social capital outcomes

- Their interpretation of these results on social capital is that chiefdoms with fewer ruling families needed to assert social control in the hands of the few authorities and therefore monitor society and bring people together and tell them what to do. This increased participation in social groups and community contribution, i.e. the social capital measures.
- They rule out the alternative explanation that more ruling families simply mean more people connected to the elite and benefitting from it, but there still being some central authority/ies that extracts rent from the many for the elite. More ruling families here would mean that the patronage network of beneficiaries is a bit larger, but development outcomes for society in general are not necessarily different. They exclude this by showing that the number of ruling families is uncorrelated with the probability of being connected to the chieftaincy elites (measure from NPS) and that controlling for this measure in the main results doesn't change anything.
- A policy conclusion they offer is that investing money in community development programmes may be futile if no attention is paid to how local institutions may determine the allocation of these resources and in this particular context this may just lead to a further enrichment of local elites without the intended benefits for entire local communities.

21 Microfinance

21.1 General Summary

- A number of papers (in particular six recent RCTs: Angelucci et al. 2015 in Mexico, Attanasio et al. 2015 in Mongolia, Augsburg et al. 2015 in Bosnia, Banerjee et al. 2015 in India, Crepon et al. 2015 in Morocco and Tarozzi et al. 2015 in Ethiopia) discussed in MSc dissertation and, in more detail, in the summary document on the literature for the Diss
- Also, older and a few more non-experimental studies are discussed in the BSc Seminar Paper
- Banerjee et al. (2015): *Thus, it is an important reassurance that our results find a strong echo in five other studies that look at similar programs in different contexts (...). In short, microcredit is not for every household, or even most households, and it does not lead to the miraculous social transformation some proponents have claimed. Its principal impact seems to be, perhaps unsurprisingly, that it allows some households to sacrifice some instantaneous utility (temptation goods or leisure) in order to finance lumpy purchases, either for their home or in order to establish or expand a business. Prima facie, these marginal businesses do not appear to be highly productive or profitable, (...)*

21.2 Papers

Rajan & Zingales (1998). Financial Dependence and Growth. AER.

- address the fundamental question whether financial development leads to growth (Schumpeter) or growth leads to financial development (Robinson)
- Their innovation / identification is to carry out a within-country between-industry analysis looking at how growth rates for certain industries are influenced by their external financing need. The industry's financing need is proxied by the financing need of the same industry in the US
 - Note that this is not only done because of lacking data but – even if data on external financing need in all countries was available – it would certainly be endogenous in this question: The fact that some countries have less developed financial markets means that industries display lower external finance need
- They run the following specification for 20-30 industries in 41 countries.

$$\begin{aligned} \text{growth}_{jk} = & \beta_1 \text{country} + \beta_2 \text{industry} \\ & + \beta_3 (\text{industry } j's \text{ share of manufacturing in country } k \text{ in 1980}) \\ & + \beta_4 (\text{ext. depend. of the industry} \times \text{financ. dev. of country } k) \end{aligned}$$

- For different measures of financial development, they robustly find a negative coefficient.
- They test for reverse causality and can exclude that.

Banerjee and Mushni (2004). How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur. Review of Economic Studies

- Look at local population (Gounders) vs. newcomers (Outsiders) in Tirupur, a town with garment industry in southern India (that produces 70% of India's knitted garment exports) to investigate how social ties affect access to capital
- Basic facts emerging from the comparison (Gounders v. Outsiders) of means and growth rates, after controlling for (years of) experience in exporting and cohort effects (cohort = all firms starting business in a particular year)
 - Average Grounder firm had three times more capital than average Outsider firm initially
 - At all levels of experience (years since business start), the average Grounder firm has more fixed capital than the average Outsider firm
 - At all levels of experience, the capital-export and capital-output ratio is higher for the average Grounder firm
 - Outsider firms that start off with less capital grow faster than Grounder firms on average and exceed Grounder firms in output after about 5 years.
- These community level effects do not reflect ability / unobserved *individual* heterogeneity since Outsiders with *less* capital grow faster and produce more after a few years.
- *Within* communities, ability and capital do not seem to be substitutes since – within communities – firms that have more capital also produce more (in line with standard theory)
 - The within- and across-community results point to the fact that capital access is majorly associated with social ties.
- QNS:
 - How is capital stock measured?

- Is there a counter-factual calculation: how large are the effects? If outsiders are more able (more successful after a few years), but start off with less capital, they should – in terms of efficient allocation and functioning markets – have even more investment than Gounders to start with, ie. estimates should be lower bounds
- Is there a measure of social ties for a structural approach?
- Capital providers may be rational in giving capital to people they know because there is high uncertainty associated with giving capital to strangers – their decision may be rational. Still, this information asymmetry gives great inefficiency. Also, the degree of risk aversion you would need for the magnitude of effects they find is unrealistic.

Pitt and Khandker (1998). *The Impact of Group-Based Credit Programs on Poor Households in Bangladesh? Does the Gender of Participants Matter?* JPE

- One of the first evaluations of Grameen micro credit impact in Bangladesh, by gender. Outcomes considered are labour supply, schooling, household expenditure, assets. Women found to experience larger effects than men (eg. 18% v. 11% increase in HH expenditure)
- Very complicated quasi-experimental WESML-LIML-FE approach. Credible?
- They use the rule that HHs are no longer eligible if they own more than $\frac{1}{2}$ an acre of land as exogenous variation in an ML approach with the identifying assumption that land ownership is exogenous
 - This is likely not the case: land ownership associated with underlying ability or ability of parents and therefore success?
- RDD or DID between programme and non-programme villages using the cutoff would probably have been a better approach
- QNS:
 - Credible identifying assumption and method?
 - Are gender effects driven by selection: more able / stronger women selecting in to the programme since it's harder for woman to have access?

Morduch (1998). *Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh*

- One of the first evaluations of programmes in Bangladesh: Grameen, BRAC, BRDB
- DID strategy using the 0.5 acre of land holding cutoff as an eligibility criterion within treatment villages.
- Key identifying assumption (parallel trend): In the absence of the programme, the difference between eligible and non-eligible people (>0.5 acres of land vs. <0.5 acres of land) would have been the same across treatment and control villages.
 - Reasonable? Treatment villages may have been precisely targeted because of such a baseline non-parallel trend, e.g. target the ultra-poor with low land holdings who show a relatively larger difference to holders of more land in their villages compared to the difference between <0.5 and >0.5 acre landholders in "control" villages.

- Also, it seems that the rule was not very strictly followed. The prob. Of eligibility is indeed almost 1 for those <0.5 acre of land, but ranges between 25 and 50% (!) for people with land holdings between 0.5 and 3 acres of land in treatment villages. And there is certainly a selection of those who borrow with such land holdings.
- Main result: No effect on consumption in all three programmes:

Table 6

Average logarithm of consumption per capita,
Difference-in-difference using *de facto* classifications (n = 1798)

	Grameen	BRAC	BRDB	Control	Difference		
	(1)	(2)	(3)		(1)	(2)	(3)
“Eligible”	4.23	4.24	4.18	4.24	-.01 (.33)	0 (.16)	-.06** (1.98)
“Not eligible”	4.50	4.53	4.60	4.51	-.01	.02	.08
Difference	-.27	-.29	-.42	-.27	0 (.05)	-.02 (.12)	-.14 (1.54)

Absolute values of t-statistics of differences in parentheses; ** (*) significant with 95% (90%) confidence.

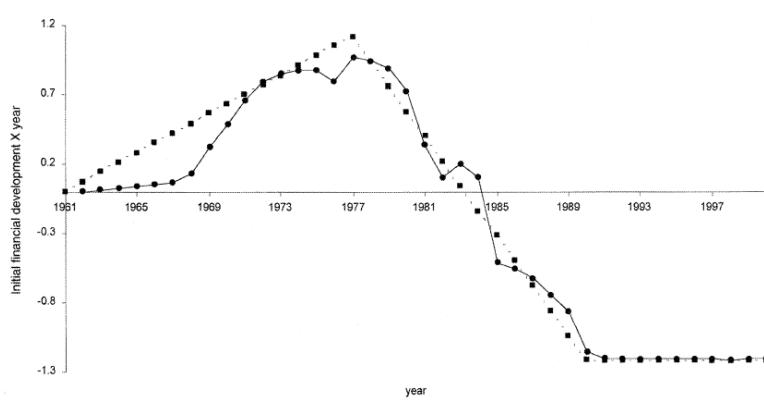
- He finds that people seemed to rather use credit to reduce vulnerability and increase consumption smoothing, but there was no positive level effect on consumption.

Burgess and Pande (2006). *Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. AER*

- Non-random placement of new banks is problematic for estimation of the causal effect of banking sector expansion on poverty reduction. The Indian government implemented a branch licensing policy between 1977 and 1990, by which a bank must open branches in four eligible unbanked locations before receiving a license to open a further branch in a banked location
- They use this policy for their identification strategy: they take deviations from the pre-programme (pre-1977) trend in branch expansion that are due to the 1:4 policy as an instrument for rural bank openings in certain areas. 1st stage:

$$\widehat{B}_{it}^R = \widehat{\alpha}_i + \widehat{\beta}_t + \widehat{\gamma}_1(B_{i1961} \times [t - 1961]) + \widehat{\gamma}_2(B_{i1961} \times [t - 1977]) + \widehat{\gamma}_3(B_{i1961} \times [t - 1990]) \\ + \widehat{\gamma}_4(B_{i1961} \times P_{1977}) + \widehat{\gamma}_5(B_{i1961} \times P_{1990}) + \epsilon_{it}$$

Where P_T are dummies equaling 1 from T (controlling for fixed effects from start and finish of policy. This shows rel-



evance clearly (the trend reversals captured in γ_2 and γ_3 are clearly seen in the graph

- The second stage is:

$$y_{it} = \alpha_i + \beta_t + \phi \widehat{B}_{it}^R + \eta_1([t - 1961] \times B_{i1961}) + \eta_2(P_{1977} \times B_{i1961}) + \eta_3(P_{1990} \times B_{i1961}) + u_{it}$$

i.e. the excluded instruments here are precisely $(B_{i1961} \times [t - 1977])$ and $(B_{i1961} \times [t - 1990])$

- They estimate overall effects at the village level and find that rural branch expansion can explain a 14-17 ppt decrease in rural headcount (no effect on urban headcount found as robustness check) which corresponds to roughly half the overall fall across the period
- This runs contrary to the oft-given concern that state controlled banks would be susceptible to elite capture
- QNS:
 - Eligibility of unbanked locations determined by the government as a source of endogeneity?
 - What about individual effects? Do we observe large effects because of marginal households benefitting, but not the ultra-poor?
 - How to square this with generally modest results of microcredit? Break-up by banking products needed? Is it savings, or the combination of all that brings about these substantial positive effects?
- These effects are at the state level. All sorts of mechanisms are going on here; it's not necessarily only the direct effect of microfinance. If small firms receive finance, they may hire more people which also reduces poverty and so on.
- The IV is a LATE estimate, so banks may still choose to go to the most promising unbanked locations where we may expect the largest effect.
- The magnitudes are not to be taken too seriously. The paper is more or less highlighting that banking matters on a very general level.

Banerjee and Duflo (2014). Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program. RevEconSt

- Using a policy that made some firms eligible for directed credit in 1998 and lose this eligibility again in 2000 while other firms were eligible for preferential credit throughout, they find that credit constraints play a large role in the setting in India considered
- Banks are required to lend 40% of their loan portfolio to a priority sector at preferential rates. The definition (cut-off point in terms of capital) of the priority sector changed in 1998 and 2000.
- Key idea: If firms are not credit constrained, they can borrow as much as they want on the credit market at the market interest rate. They will borrow as long as their ROI equals the (constant) interest rate. If they receive preferential credit, they will substitute as much of their (market) borrowing with the preferential loan at lower rates. If, however, their total borrowing exceeds the size of the preferential loan they can get (banks have to allocate their preferential loan size between all eligible firms), they would *not* increase their total borrowing (and sales) since their ROI on the last marginal dollar of investment still equals the market interest rate. Only their profits should increase because of lower cost.

- Only firms that have fully substituted market borrowing with preferential credit will expand their investment as long as the ROI exceeds the preferential credit rate.
- Estimation strategy (triple difference; DID on growth rates): compare how the rate of change (growth rate) of borrowing and sales changes pre- to post-1998 for newly eligible firms ($BIG=1$) relative to always eligible firms.

$$y_{i,t} - y_{i,t-1} = \alpha BIG_i + \beta POST_t + \gamma BIG_i \times POST_t + \epsilon_{i,t}$$

- Identifying assumption: there was no differential *trend* around the years looked at between newly eligible and always eligible firms that is unrelated to credit availability.
- If all firms use both preferential credit and market borrowing post-1998 and always eligible firms use both before, the difference in growth rates between newly eligible and always eligible firms should *not* change if there are *no* credit constraints. The fact that they do provides evidence that there are credit constraints.
 - (Post-interest) Profit should of course change if part of the credit is given at a preferential rate. However, a measure like EBIT should not change.
- They also use an IV setting and estimate a model, in particular an interest rate, to find more direct evidence of credit constraints (that of course, however, relies on the model assumptions). The posit a simple relationship between revenues and credit:

$$\Delta \log R_{it} = \Delta \log A_{it} + \theta \Delta \log k_{it}$$

where k_{it} is instrumented for using $BIG \times POST$

- They find here and above that additional resources are used to finance expansion, rather than to just substitute market borrowing. Sales grow with credit. The IV gives a remarkable rate of return (possibly over 100% for compliers; however, see LATE concern) where market interest rates are in the range between 30 and 60%.
- Note that an OLS of ΔR on $\Delta Credit$ gives a zero estimate because of reverse causality here: The effect of credit changes on revenue changes is positive. However, the effect of revenue changes on credit changes is negative: Firms experiencing decreasing or negative revenue growth resort to credit to finance their business relative to successful firms self-financing. The two effects seem to cancel.
- QNS:
 - Is the policy change endogenous: these years were chosen because of differential trend for larger v. smaller firms?
- Concerns:
 - IV gives LATE estimate. This is the treatment effect for the compliers. The LATE in this setting is essentially the weighted average of marginal returns where the weights reflect how strongly the respective firm reacts to credit. If the high return people are those taking up credit, LATE > ATE
 - There are also probably spillover effects: If banks have a fixed pool for preferential credit and the number of eligible firms increases, there is less for the always eligible firms in the control group. This is a violation of the SUTVA assumption.

Fischer (2013). Contract Structure, Risk-Sharing, and Investment Choice. Econometrica

- Comes up with a unified model of investment choice, informal risk-sharing and formal financial contracts that aims to explain why many MFI clients don't grow their businesses beyond subsistence level. The theory is that, in the environment of such a model, risky but high-expected return investments are discouraged to an inefficient extent (less risk taken than socially optimal)
- Two main effects under both formal and informal arrangements: joint liability in which you are insured against default by your project partner but do not necessarily have to pay compensations when being successful makes you free-ride and engage in risk-taking
- Equally, under peer monitoring, you always have to pay a transfer (default insurance) but do not receive compensating payments in case of the other project's success. This makes you heavily discourage risk-taking by the other.
- The joint liability of formal contracts may mean you also develop an informal risk pooling arrangement with your liability partners. Again, both effects play a role and the need to provide insurance in case the other project fails may discourage you jointly to go for risky options to prevent the case where both fail.
- An experiment in the field with MF clients in India seems to show that the discouragement of risk-taking effect seems to dominate. Also, equity-like contracts produced more risk-taking and lower defaults which means they may be the best options.

Karlan and Zinman (2010). *Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to estimate the Impacts in Manila*. WP

- Setting: For-profit micro-lender "First Macro Bank" (FMB) that gives out three-months loans at 60% APR. Sample drawn from thousands of applicants who applied at 8 out of 9 FMB branches over 18 months between 2006-07
- KZ built a model with them to distinguish creditworthy from non-creditworthy lenders. They then randomise among marginal lenders (specific group!) whether their loan is approved or not. This provides controlled risk for the lender and randomisation for the evaluator.
 - They do acknowledge that the subject group is an important caveat: 74% of them are first-time applicants. Applicants clearly above or below the line of creditworthiness may not be comparable at all.
 - There are two different risk groups for randomisation: a lower and higher bracket. In the lower bracket, there is a 60% chance of approval, in the higher bracket, there is an 85% chance of approval.
- They investigate ITT effects (351/1272 treatment assigned did not ultimately result in a loan). They also include a dummy for the two different risk groups.
- Some surprising results:
 - (i) business investment does not increase, but *shrink*. Some evidence that profits increase and the mechanism seems to be that unproductive employees are shed (credit access reduces the need for favour-trading within families and also provides a good excuse not to employ unproductive family members)
 - (ii) male and high-income borrowers display strongest treatment effect, contrary to usual targeting policy
 - (iii) there is some evidence among *male* borrowers that more is invested into children's education, contrary to other findings.

- (iv) Substitution away from formal insurance
- (v) however, increased access to (informal) risk-sharing mechanisms
- Overall picture: Microcredit seems to work through interactions between credit access and risk-sharing (often viewed as second-order by policymakers) as well as human capital investment, but in line the fact that money is fungible and not necessarily used (only) to grow one's business.
 - Therefore, focussing on business outcomes leads to a wrong assessment of MF in a welfare sense: "Business outcomes are not a sufficient statistic for household welfare"

Karlan and Zinman (2009). *Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment.* Econometrica

- They test the empirical importance of adverse selection vs. moral hazard in a field experiment in South Africa.
- They randomise along three dimensions to disentangle the effect: (i) initial interest rate offered – high or low, (ii) actual contract rate revealed only after agreement – low for low initial, in subsample dropped to low from high initial, (iii) future interest rate – low if repayment (dynamic incentive) or normal (no dynamic incentive)
- Distinguishing between borrowers selecting in at different rates, but then facing the same actual rate and those selecting in at the same rate but then facing different rates or different dynamic incentives enables them to disentangle adverse selection from moral hazard empirically

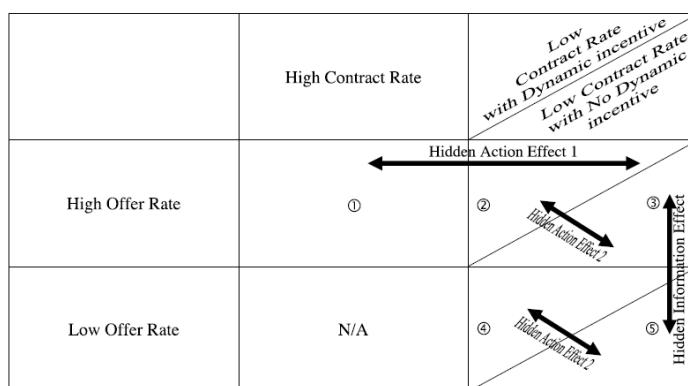


FIGURE 1.—Some basic intuition for our identification strategy.

- They find strong significant effects for moral hazard (explaining 13-21% of default), but weak (not significant) evidence for adverse selection
- HUGE problem though: their sample are former successful (repaying) borrowers, so already pre-screened and pre-selected people. Therefore, it is hardly a surprise that there is no significant effect on adverse selection. The paper is still influential since it shows a very neat way of disentangling the effects.
- There are a few assumptions in the basic model with risky types (r) and safe types (s):
 - Loan repayment is possible if the project succeeds
 - The probability of success is a concave function of effort
 - The probability of success decreases in riskier types
 - The expected return is constant across risk types (while riskier types have lower success probability, they have higher returns in case of success)

- You always pay back your loan in case of project success (the cost of default is higher than repayment for all risk types in case of success)
- Note the difference between the two types of moral hazard above
 - The first ("repayment burden", the horizontal arrow) is the distinction between different rates, conditional on having selected in at the same rate. A given set of borrowers exerts less effort at higher rates.
 - The second ("moral hazard", the diagonal arrows) is the distinction between different repayment costs, conditional on having selected in at the same rate and paying the same rate (in period 1). A given set of borrowers exerts less effort as the cost of default decreases.
- There are two channels of hidden information / adverse selection:
 - Riskier types select into borrowing at higher rates if the cost of default in case of project failure is smaller than the loan repayment. The reason is that riskier types essentially incur the default cost instead of the loan repayment more often. Conversely, if the default in case of project failure is larger than the loan repayment, safer types select into borrowing at higher rates.
 - Optimal effort decreases in type (high type = riskier type) and lower effort means lower success probability. Based on this logic, for *discrete* changes in the interest rate (as is the case here; for continuous changes, this second channel is zero by the Envelope theorem), there may be selection on anticipated effort. The additional cost of effort at the discretely higher rate may be smaller or larger than the additional benefit for the risky type. That is, the risky type may select in (on anticipated effort) at higher rates since her effort increase at higher rates has larger cost than benefit and she would therefore exert less effort and have lower success probability meaning she pays back the interest rate less often.
 - Depending on which of the four possible combinations you consider (which direction the first and second channel take depends on the parameters), the two effects either go in the same direction or in opposing directions. In particular, let's assume the classic case where the first effect is an in-selection of risky types at high rates. Now, if risky types also experience larger cost from additional effort than benefit at the high rate, the second channel goes in the same direction and reinforces the adverse selection. However, if that's not the case, the second channel goes in the opposite direction of the first and the net effect is theoretically ambiguous.
- The (imprecise) zero estimate on adverse selection is unclear to interpret: either no (significant) effect or two opposing effects.

Banerjee et al. (2015). *The Miracle of Microfinance? Evidence from a Randomized Evaluation*. AEJ: Applied

- Evaluation of MFI Spandana randomised expansion in Hyderabad.
- Main results:
 - Significant increase of business profits, mostly for existing businesses; informal borrowing is substituted for (overall borrowing did not increase)
 - Take-up is 17.8% (26.4% for all MFIs)

- No significant effect on children's schooling, child work, women independence index; highly sign. effect on the number of women managing a business
- No sign. effect on consumption, significant reduction of temptation goods, significant increase in durables
- Increase in labour supply found, especially in the own business
- ITT effects estimated, since there may be GE and spillover effects (low take-up though, so unclear what insignificance means)
- Heterogeneous effects: there is *no* significant effect on business profits between the 5th and 95th quantile – the extreme upper tail is driving the entire result.
- Important caveats:
 - Low take-up, so effects are driven by marginal borrowers
 - Women were provided with a loan, they could choose what to use the money for
 - Study took place in marginal neighbourhoods that the MFI was indifferent about working in; this is clearly a selected region where Spandana may not have expected great effects

22 Technology Adoption

22.1 General Summary

22.2 Papers

Foster & Rosenzweig (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. JPE

- Setting: Green revolution in India during which new technology in agriculture became available. The new varieties are high yield varieties (HYVs)
- Question: How do people learn? Learning is different from social pressure. The key assumption in learning is that other's behaviour influences your own productivity, not only your own behaviour. Then, we can speak of learning externalities.
- They use a target-input model of new technology in which the best use of inputs under the new technology is unknown and stochastic and changes randomly each period
 - This gives a growing profitability of any new technology over time as opposed to a simple one-time adoption of technology once the exogenous profitability is known
- The target use of technology is

$$\widetilde{\theta_{ijt}} = \theta^* + u_{ijt}$$

Where the error is mean-zero iid. normal and farmers have a prior about the average target use $\theta^* \sim N(\widehat{\theta_{j0}}, \sigma_{\theta j0}^2)$

- After each period, farmers realise the previous target use, so they can learn about θ^* . They also learn about the target use of HYVs used by neighbours, with some noise, that is, they learn $\widetilde{\theta}_{ijt} + \xi_{ijt}$. This gives them a (precision-weighted) posterior about θ^*
- Note the several levels of randomness: randomness in the actual target use each period, randomness in the belief on mean target use and randomness in the revelation of neighbour's target use.
- Payoffs are decreasing in the square of the distance between actual and target use.
- Several results arise in Nash equilibrium between farmers and their neighbours:
 - Increases in the cumulative number of HYV parcels planted up to time period t by the farmer raise the profitability of a farmer's HYV at t , as do increases in HYV parcels planted up to t by neighbours
 - Farmers will always use more than the within-period optimal amount of HYVs because of the learning gains for future profitability; that is, the current period marginal return to planting HYV is negative.
 - If there is no learning from neighbours (the variance of ξ goes to infinity), there are no effects of neighbours' assets (neighbours' amounts of parcels of land A_{-j}) on the farmer's adoption decision (use of HYVs)
 - If learning externalities (on others) are internalised (e.g. by a social planner), the effect of neighbour's assets on the amount of HYV area planted by a farmer is always positive, and precisely n times the effect of the farmer's own assets on her adoption of HYVs
 - If learning is not internalised, the effect of own assets on HYV adoption is still positive, but the effect of neighbours' assets on HYV adoption may be negative (free-riding on their experimentation)
- In a second step, they use panel data to estimate the model they set up. There are several issues:
 - Assets are likely to be endogenous: If farmers live in areas with better land, they tend to accumulate more assets including land. This gives them a greater incentive to benefit from experimentation. Therefore, the authors take FDs to get rid of farmer individual unobserved heterogeneity
 - Differenced shocks (error term) in the (differenced) profit equation may be problematic. The HYV decision (how many parcels to use for HYVs, H_{jt}) is realised before a shock in the same time period. However, a past shock may affect the HYV decision, e.g. because a positive past shock may mean farmers have more to invest. The past shock shows up in the differenced shock, so there may be correlation between it (the error term) and the regressor H_{jt} . Therefore, they instrument for this variable using inherited assets.
 - Endogeneity? Parents' ability (correlated with their assets) may be correlated with the children's ability. Hard working children may inherit more.
- Results:
 - (i) imperfect knowledge about the new technology is a barrier to adoption. Profitability increases significantly with own and neighbours' experience net of the adoption of HYVs in addition to affecting rates of adoption. This means that experience effects in part operate to let farmers make better decision about input use

- (ii) But the effect declines as experience increases, i.e. the importance of the barrier diminishes over time
- (iii) own experience and neighbours' experience matter for HYV profitability. Doubling own experience increases future profits by 36%. There are learning spillovers. Comparing the magnitudes for own vs. neighbours' experience: A given increase in average experience by a farmer's neighbour increases profitability twice as much as the same increase in own experience.
- (iv) farmers do not fully incorporate the village returns to learning in making adoption decisions. Net of own and neighbours' experience, own and neighbours' assets have opposite effects on adoption, that is, farmers tend to free-ride on others' learning.
 - Caveat: others' experimentation is measured by the average experimentation of *everyone else in the village*. This may also reflect a common village shock, rather than learning.
- Policy implication: no sufficient coordination on social learning within the village. The free-riding problem leads to too slow adoption relative to the social optimum. This gives a scope for public efforts to increase adoption through subsidies of early adopters.

Banerjee (1992). A simple model of herd behaviour. QJE.

- Considers a sequential decision model in which people have their own signal but also observe the behaviour of others and infer about their signals.
- The equilibrium is characterised by herd behaviour in the sense that people end up doing what everyone else does and give their own signal relatively little weight.
- In extreme cases, herd behaviour may lead to worse outcomes for a majority and an equilibrium that is far from the social optimum
- Simple example: choice between restaurant A and B. Prior prob.s: 51% that A is better, 49% that B is better. 100 people, each of whom receives a private signal. Private signals from any person have equal weight. 99 receive the private signal B, one person receives the private signal A. If the latter person chooses first, she will choose A for sure. The second person is key: She *knows* from the first person's behaviour that the first person must have received the private signal A. This private signal and her own private signal cancel and the second person decides in line with the prior prob.s and joins the first person in restaurant A. The third person knows the second person's decision problem, i.e. realises that the second person would have chosen A regardless of her private signal. Therefore, she essentially faces the same decision problem as the second person and chooses A. And so on: Everybody ends up in A although B is almost certainly better.
 - The second person imposes a "herd externality" on everyone else by ignoring her private signal.
- General model: N agents sequentially choose a point $i \in [0,1]$. There is a correct point i^* that yields payoff $z > 0$ and the payoff is 0 for $i \neq i^*$. Everyone receives a private signal i' which is equal to i^* w/p. β and drawn randomly on the uniform distribution between 0 and 1 w/p. $1 - \beta$. Those with incorrect signals can cause a herd of the uninformed in this setting.
- Key results in the general model:

- Equilibrium pattern may be inefficient ex ante in the welfare sense: people would like to agree before the game that the first couple of deciders are not allowed to observe others' actions. Follows directly from the herd externality
- The probability that no one chooses the right option is bounded away from zero and can be made large by making the probability that a player's private signal is correct (β) small
- The herd externality is a positive feedback effect (if someone joins the crowd, she induces others to join the crowd as well). Therefore, the equilibrium pattern is very volatile across several plays of the same game

Ellison & Fudenberg (1993). *Rules of Thumb for Social Learning.* JPE.

- Setting in which agents decide between two technologies based on the experiences of their neighbour. Two environments are considered: homogeneous population of players where the same technology is better for everyone and heterogeneous population where each technology is better for some of them
- Main result: Setting full Bayesian updating of the entire history aside (which involves extremely complex calculations) and concentrating on fairly naïve decision rules (of thumb) that ignore historical data and focus only on last period's payoffs and incorporate some popularity weighting (go for a technology that is taken by many) can be quite efficient.
- In the homogeneous environment, people simply look at the average performance of the previous period and may incorporate some popularity weighting. In the heterogeneous environment, people are indexed on the interval [0,1] and it is assumed that neighbours close to oneself's position have similar payoffs. Then, agents are assumed to focus on a certain 'window width' around their own point and apply the rules of thumb: They take the average performance within their interval and some popularity weighting to make their decision.
- The underlying question can be whether a new technology is adopted and whether it is adopted by the appropriate players in the heterogeneous setting.
 - The speed of adoption in the homogenous model is correlated with the extent of the payoff difference.
 - Low-probability high-gain and high-probability small-loss technologies will be adopted slowly, if at all. (e.g. vaccines, seatbelts)
 - In the heterogeneous model, parameters that favour long-run efficiency – small window widths and high popularity weights – lead to slow diffusion of a new technology
- Model: Technology f and g can be chosen where the latter is better if $\theta > 0$ and θ is unknown:

$$u_t^g - u_t^f = \theta + \epsilon_t$$

Where $\epsilon_t \sim U[-\sigma, \sigma]$

- A share α gets to reconsider her choice each period and let x_t denote the share of people using g in period t
- The (imposed) decision rule is choose g if $u_t^g - u_t^f \geq m(1 - 2x_t)$ where m reflects the popularity weighting. It can be chosen that $m = \sigma$ is efficient in the long-run in the sense that it always leads to $x_t = 1$ iff $\theta > 0$ in the long run.

- However, there is a trade-off between the probability of convergence and the speed of convergence. Higher m leads to higher speed of convergence at the cost of a risk of inefficient herding.

Conley & Udry (2010). *Learning about a new technology: Pineapple in Ghana. AER.*

- Inherently difficult to disentangle learning effects from unobservable common shocks that are spatially correlated.
- It is hard to define (i) the set of neighbours from whom an individual can learn and (ii) given this set, distinguishing learning from other phenomena such as interdependent preferences, technologies or being subject to a common shock
- They collect detailed data on information interconnections by asking people who they talk to, are given advice by, how often they interact. This allows them to distinguish information neighbours from geographic neighbours. They also have detailed information on credit, soil information, geographic information etc.
- Construct a local learning model where the production function is $f(x)$ and information on $f(x)$ only provides information on x , not values in the neighbourhood of x . They also construct an index for good vs. bad news
- They use a two-year HH survey of 200 HHs in southern Ghana.
- Results: A given farmer is
 - More likely to change fertiliser use upon bad news from neighbours that use similar levels of fertiliser than upon bad news from neighbours using different levels of fertiliser
 - Increases (decreases) input use if an information neighbour achieves higher profits using more (less) fertiliser. Holds up when controlling for correlation in growing conditions, common credit shocks, several information metrics
 - More responsive to news if inexperienced in pineapple cultivation.
 - More responsive to news from experienced farmers with similar wealth level (not those with similar soil level or other similarities)
- Robustness: For the established maize-cassava cultivation, they find indeed no learning effect as you would expect for established crops.

Duflo et al. (2006). *Understanding technology adoption: Fertilizer in Western Kenya. Evidence from field experiments. WP.*

(related to Duflo-Kremer-Robinson 2008 paper on returns to fertilizer and DKR 2011 paper on behavioural reasons for uptake "Nudging farmers...")

- In spite of potential to improve yields substantially, fertiliser has been used for maize in Kenya in the year preceding the survey only by 17% of farmers.
- Three main reasons often given for low adoption of fertiliser, especially in African countries:
 - (i) Fertiliser may not be appropriate in the specific context
 - (ii) It may be appropriate, but farmers don't know about it or don't know how to use it. Inefficiently low levels of experimentation due to learning externalities possible (relate to Foster-Rosenzweig)

- (iii) Issue of financing, even short-term saving for fertiliser purchase as profitable investment may be impossible
- They conducted several experiments with multiple treatment arms to address the issues:
 - Try different amounts of fertiliser to learn about returns
 - $\frac{1}{4}$ teaspoon top dressing
 - $\frac{1}{2}$ teaspoon top dressing
 - 1 teaspoon top dressing
 - Full package recommended by Ministry of Agriculture
 - Learn about technology adoption:
 - Demonstration plots where randomly chosen farmers were given information and test inputs in the first period (learning by watching)
 - Starter kits with instructions how to use fertiliser (learning by doing)
 - Invite randomly chosen 'agricultural connections' of farmers (people who discuss agriculture with them) for demonstrations and discussions (learning from others)
- Results:
 - Returns are high for $\frac{1}{2}$ teaspoon fertiliser use (so (i) fails as an explanation)
 - Learning adoption shows quite disillusioning results:
 - demonstration plots lead to roughly 10ppt higher likelihood of fertiliser use 1, 2 and 3 seasons later (significant, but small)
 - starter kit effect is small and dies out after 1 season
 - invited contacts effect is also small and dies out after 1 season
 - they find here that the inability to save seems to be a strong impediment
 - update in DKR-2011: somewhat more sophisticated behavioural story. Evidence for present biased farmers who take up if offered a specific deal of the opportunity to buy the fertiliser as commitment straight after harvest. Check the DKR 2011 paper for details

Miguel & Kremer (2007). *The Illusion of Sustainability*. QJE.

- Builds on MK 2004 paper on deworming (see health chapter)
- As opposed to the original paper focussing on the benefits of deworming, this paper focuses on the adoption aspect and whether continued high subsidies seem warranted or what mechanisms for adoption would work
- One of the questions they explore is whether – if peers play a role for the adoption decision – pushing a few people to adopt gets a community to move from a low-adoption to a high-adoption equilibrium
- They test different adoption strategies
 - Health education programmes did not affect behaviour
 - A small fee (30 cents) instead of a full subsidy led to an almost 80 % reduction in take-up (from 75% to 19%) while – once you move from a zero to a non-zero price – take-up is not very responsive to price changes. (Note that the latter is not a general result however. Other studies find quite a high price sensitivity even for small positive prices; see chapter on pricing and WTP)

- A mobilisation intervention where people were asked in advance whether they plan on using deworming drug (to induce commitment exploiting the social psychology finding that people strive for consistency between their actions and statements) had no impact on adoption
- Most striking result (perhaps) is that people with more links to early treatment schools displayed *lower* take-up going against the positive peer effects in adoption hypothesis
 - Potential explanation is that they had very high priors (particularly the more educated with more links) which they adjusted downwards when learning about the small private value and long-run benefit vs. immediate cost ratio (s. MK 2004)
 - Also: If there are benefits of treated on the untreated, those with many links may have had many people around them that were dewormed and were free-riding on them, seeing less need to take the medicine themselves
- MK conclusion: given the large externalities and issue of gradually evolving benefits facing immediate cost, there does not seem to be sustainably high adoption once you move from a full subsidy. They conclude that it is necessary for (the socially desirable) high adoption to have ongoing subsidies.

Kremer et al. (2008). *Trickle down: Diffusion of chlorine for drinking water treatment in Kenya*. WP.

- Use randomly phased-in drinking water treatment WaterGuard. It is a point-of-use chlorine treatment for drinking water that kills bacteria
- Assess both the health impact and technology adoption
- Treatment resulted in a 69% reduction of the bacteria E.coli in water. Child diarrhea fell by 35-40%
- Households with close relationships to a treatment household were significantly more likely to adopt the new technology
- Why did 20% not take up when WaterGuard was provided for free?

23 Pricing and Willingness-to-Pay

23.1 General Summary

Papers address roughly the following four questions:

- How price-sensitive is demand?
- Can information or marketing make households less sensitive to price?
- Do households appropriately use health products that are given to them for free?
- Are households willing to pay for a health product once the subsidy is lifted?

23.2 Papers

Dupas (2009). *What Matters (and What Does Not) in Household's Decision to Invest in Malaria Prevention?* AER P&P

- She studies insecticide-treated bednets (ITN) and their take-up under different treatments in Kenya. ITNs reduce overall child mortality by 18% and general population morbidity.
- The different treatments they have in addition to random variation in price:
 - Different framing (health benefits emphasised vs. cost savings emphasised vs. nothing)
 - Verbal commitment to purchase the product (“foot-in-the-door” technique)
- She finds that no intervention had an effect on uptake and the gender doesn't matter either. With regards to the price, take-up is highly sensitive in spite of large benefits.
 - However, for this and similar papers looking at framing / marketing, note that usually no professional agency is hired. The fact that there is a huge market for marketing & communication in the private business sector strongly suggests that marketing works.

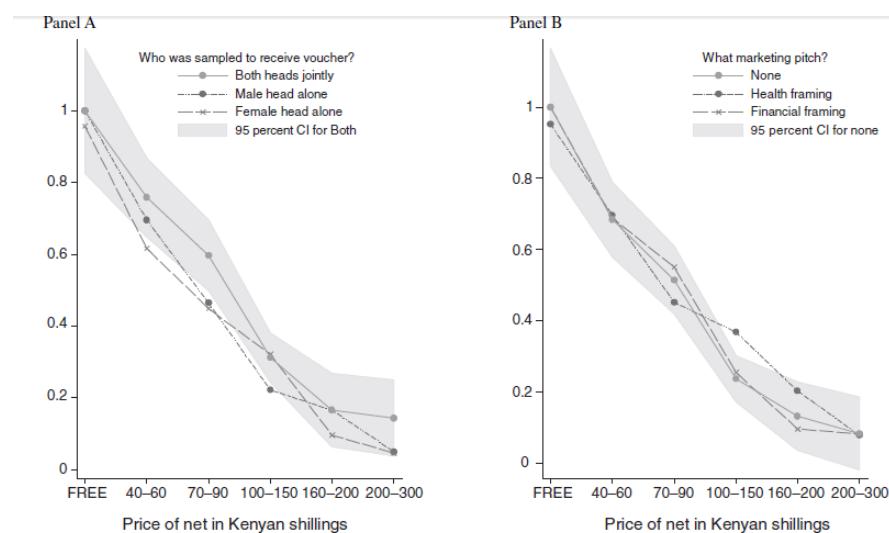


FIGURE 1. TAKE-UP OF NET, BY PRICE GROUP

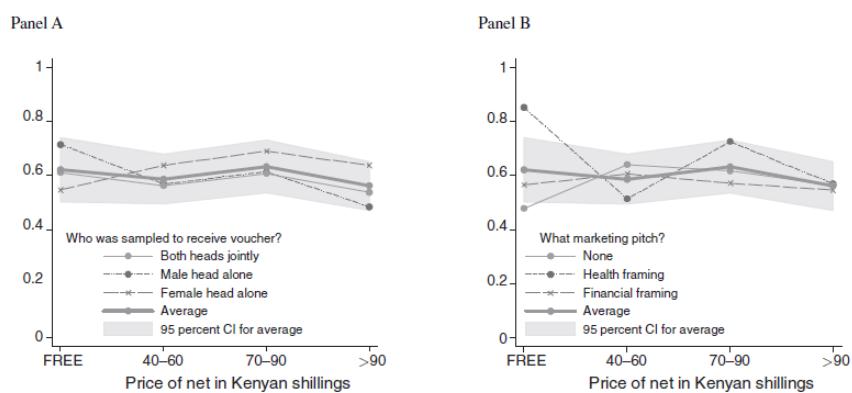
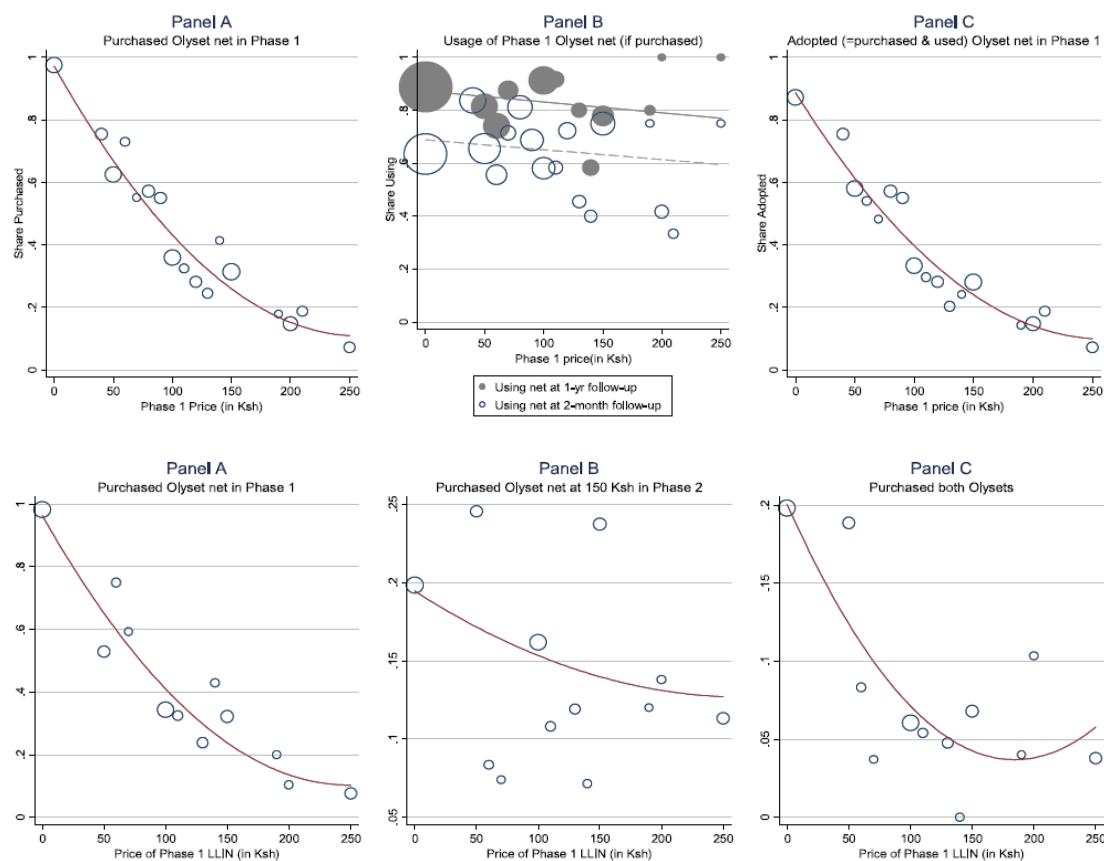


FIGURE 2. SHARE OF HOUSEHOLDS USING THE EXPERIMENTAL NET AT FOLLOW-UP

Dupas (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. Ecta.

- Addresses the question how short-run subsidies affect long-run adoption. Two opposing forces are at play

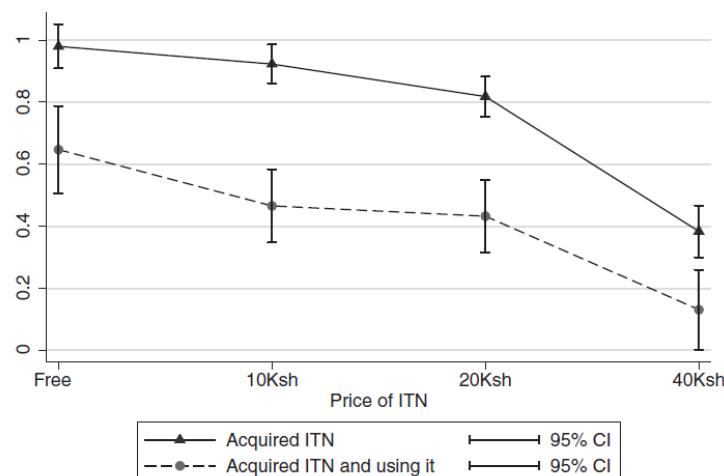
- Anchoring: If people anchor around the subsidised price, this may decrease their future WTP for the product once the subsidy is removed
- Learning: If the product needs experiencing first, a one-off subsidies and learning about the benefits may boost future take-up at the full price
- She assesses which effect is present / dominates in a two-phase design of Olyset long-lasting insecticide-treated bed net (LLIN) sale in Kenya; this product is practically a better version of an ITN.
 - Phase 1: LLINs are offered at a price between 0 and 3.80 USD for three months across households within six villages
 - Phase 2: A year later, LLINs could be purchased at 2.30 USD for everyone; on the market, LLINs were not available, but ITNs were at 1.50 USD
- They find strong price sensitivity. Reducing the price from 3.80 USD to 0.75 USD increases demand from 7% to over 60%. They also find stable usage rate across prices (slight decline even) [upper graph]
- They find evidence for the learning effect (lower prices positively affect future WTP). Those who paid a heavily subsidised price (<0.75 USD) were 7.2 ppt or 49% more likely to purchase in Phase 2. [bottom graph]



- Some suggestive evidence for spillovers: those experiencing higher exposure through spatial networks in Phase 1 were less likely to purchase in Phase 2
- Estimating an experience-good model that allows for reference-dependent preferences, Dupas finds strong evidence for learning and no evidence for anchoring

Cohen & Dupas (2010). Free distribution or cost-sharing? Evidence from a randomized Malaria prevention experiment. QJE.

- Three potential effects of positive prices (so-called ‘cost sharing’) and price variation on take-up/ usage:
 - (i) selection effect: charging positive prices could select out those who do not value or need the product (that’s the most common selection effect given; of course, there may also be negative selection at positive prices in that the poorest and potentially most needy who are most likely to use the product are selected out for affordability reasons)
 - (ii) sunk-cost psychological effect: paying a positive price or higher price may induce people to actually use the product purchased
 - (iii) signalling effect: positive or higher price may signal higher quality
- They don’t suspect (iii) to be strongly at play in their context since the real price of ITNs was well known (the price it was sold at in many shops) and it was clear that any price below it would be due to a subsidy
- They test for the presence of different effects using random variation in prices for ITNs in Kenya. In particular, to disentangle (i) and (ii), they employ a randomised two-stage pricing design (in the spirit of Karlan-Zinman 2010 and Ashraf et al. 2010). In some clinics charging a positive price for ITNs, once the participating women decided to buy, a random subsample was surprised with a lottery for an additional discount. For the given sample of selected-in woman, any variation in usage due to the random lottery can be interpreted as the sunk cost effect
- Main findings:
 - Sharp price sensitivity: even small positive prices reduce demand dramatically
 - Conditional on purchase, usage rates are constant across different prices (large SEs for higher prices though since the sample size is small due to small demand at high prices)
 - Effective coverage (usage rate x demand) sharply decreases in the price. This is precisely estimated; hence, even if usage rates were increasing in the prices (through i, ii and/or iii), this is dominated by the demand effect
 - There is no evidence for a sunk-cost effect; the lottery does not affect usage. This means that the constant usage is *not* due to *negative* selection and sunk-cost effects offsetting each other



- Policy conclusion: Cost-sharing arrangement could potentially lead out to the exclusion of the most vulnerable to malaria.
- External validity unclear since ITNs are a well-known product. What about unknown products or products where the benefit may not be clear?

Ashraf et al. (2010). *Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia.* AER

- Door-to-door sale of Clorin, a chlorine bleach solution to kill pathogens in household drinking water and reduce the incidence of waterborne illnesses. It is a well-known product in the context in Lusaka, Zambia, where 1000 households were subjects of the study. They conduct a follow-up survey on usage after two weeks and can actually chemically test for the product in drinking water (don't rely only on answer given)
- Two-stage randomisation design in the spirit of Karlan-Zinman 2010: Households were offered the product at an 'offer price' between 0 and the retail price. Once they agreed to purchase, they received a surprise lottery for a discount and a 'transaction price' resulting from the lottery outcome. This allows to disentangle the screening effect (higher price, in particular non-zero price, may mean people who don't value the product are screened out) from the psychological sunk-cost effect (if a higher price is paid for the product, I am more likely to use it) [s. also Cohen-Dupas 2010 on these effects]
- They find strong evidence for a selection effect (unlike Cohen-Dupas): Higher WTP (higher offer price) induces higher usage in the follow-up survey
- They do *not* find evidence for a sunk-cost effect. Usage is constant across different transaction prices, conditional on the same offer price.
- They find substantial variation in take-up / demand by the offer price. Therefore, the lack of a sunk-cost effect is not due to "small stakes"

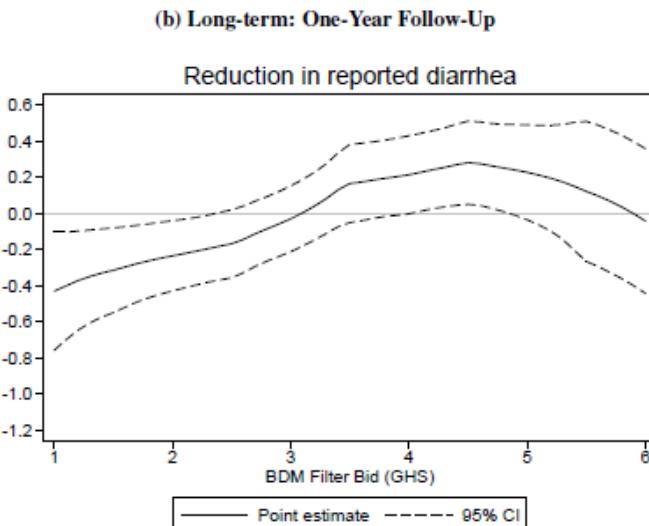
Fischer et al. (2016). *To charge or not to charge: Evidence from health product experiment in Uganda.* WP

- They study the effect of the free distribution of three different health products on subsequent demand 3 months later in Uganda to address the issue of anchoring vs. learning (cf. Dupas 2014) in a randomised setting. The three products have different characteristics:
 - Panadol: well-known painkiller that is not expensive relative to average income. Therefore, neither learning nor income effects expected (income effects in the sense that a household that had to pay a lower price initially is richer later which could drive demand changes for large price differences relative to average income; this may be a concern for ITNs). This drug is suitable to identify anchoring effects.
 - Elyzole: A "moderately well-known" deworming drug that is susceptible to *negative* learning effects b/c it has unpleasant side effects.
 - Zinkid: A largely unknown treatment for childhood diarrhea, recently recommended by the WHO, has potential for *positive* learning.

- Main result (contrary to Dupas, 2014): For all three products, prior distribution for free reduces demand in the future. The effect is insignificant for Zinkid where learning and anchoring may cancel each other off. The effect for the other two is -12ppt.
- They stress that whether learning or anchoring dominates is a function of product, market and household characteristics and therefore a general pricing policy is impossible.
 - With respect to market characteristics, they also find that take-up is significantly more likely when an NGO as opposed to a private firm offers the product at a zero price (they randomise the identity of the organisation as well)
- They develop a model that would captures these differences in characteristics and would map from them into what pricing may be optimal.

Berry et al. (2015). *Eliciting and Utilizing Willingness-to-Pay: Evidence from Field Trials in Northern Ghana.* WP

- They use the Becker-DeGroot-Marschack (BDM) mechanism to elicit willingness to pay for water filters. Idea: Subjects are asked to report a WTP b at which they would buy a product. A random price p is drawn. If $b \geq p$, they can buy the product at price p , if $b < p$, they do not get to buy. This is essentially a 2nd price auction against a random draw and the dominant strategy is to reveal true WTP.
- The BDM mechanism also gives exogenous variation in product allocation and allows them to estimate *heterogeneous* treatment effect of water filters, conditional on WTP.
- Key results:
 - Median WTP is 10-15% of the cost of manufacturing and delivery
 - Usage one year later increases slightly in WTP (small screening effect)
 - Conditional on WTP, usage is not affected by the price paid (no sunk-cost effect)
 - Benefits of the filter (measured by reductions in children's diarrhea) increase in WTP (consistent with usage pattern)
 - Small positive (non-zero) prices do not substantially reduce demand (contrary to other findings)
- BDM mechanism seemed to be well understood. Several test rounds; while there are non-negligible ex-post regret numbers, in particular for those narrowly bidding less than the drawn price, only very few (around 5%) actually offered to pay more. Also, almost everyone (269/272) actually purchased the product as agreed when $p < b$
- For the BDM WTP elicitation, hardly any household characteristics are significantly correlated with WTP. Jointly, they are significant though. Also, R^2 is low (0.053), ie. much of the heterogeneity is unexplained. In particular, E.coli (bacterium) count is uncorrelated with WTP
- Using BDM as an instrument for purchase, they estimate the effect of the water filter on health outcomes. Heterogeneous treatment effects can be found using different levels of WTP. This gives a picture of heterogeneous LATEs. There is a substantial degree of heterogeneity in the long run (one year later).



- Negative effect for low WTP may be due to moral hazard dominating the (already) low degree of usage: If you have a filter, you take dirtier water
- QNS:
 - Income or wealth measures as household characteristics are missing? This should be positive.

24 Methods and External Validity

24.1 Methods Overview

- Micro vs. Micro
 - Problem with lots of **macro** data: **endogeneity** and **selection**
 - **Natural experiments** and **RCTs** on micro-level provide a solution to this
 - Micro-level investigation (in particular RCTs) can solve **internal validity**, but **external validity** is questionable
 - A model helps for external validity
- Theory vs. Empirics
 - If you have no model, there is **no structure** and it might not be clear what predictions you want to **test**
 - Without model, **external validity** is more difficult to establish
 - A model makes sense as a **first step** before engaging in empirical research
- Conditions for a competitive equilibrium:
 - Price taking by firms and consumers
 - No externalities
 - Firms maximise profits
 - Choices reflect preferences and preferences are non-satiated
 - Preferences are individualistic
 - Markets are complete
 - All contracts are enforced
 - Goods are private

- If one of the above conditions fails, this constitutes a market failure and gives a rational for policy. Before proposing a policy, you should be able to explain and evidence the market failure that would cause a break-down of the First Welfare Theorem.
- Two problems for RCTs in the face of externalities:
 - (i) breach of Stable Unit Treatment Value Assumption (SUTVA): the control group is actually treated if there are externalities of treated people on people in the control group
 - Potential solution: randomise at village level and only estimate the *aggregate* effect
 - If wanting to explicitly measure the externality: randomly vary the ‘degree of treatment’, e.g. treat 80% in one village and 20% in another and check the difference.
 - (ii) GE effects: The externalities are part of what happens in GE. When we can somehow work around (i) and get an estimate of the ATT and ignore externalities, our finding is not reflective of the *whole* effect (ATT and ATU) of the treatment – this is an important for policy

24.2 Papers

Hunt (2015). *Site Selection Bias in Program Evaluation*. QJE

- Site selection bias in randomised evaluations (at the individual level within sites) can be a substantial problem to external validity. Often, sites chosen for the randomised evaluations are not representative for the whole or even a specific target evaluation.
 - Eg. MFIs for RCTs on Microfinance: Chosen partner MFIs differ highly significantly in 6/8 relevant observable characteristics from non-partner MFIs. Similar results from hospitals chosen for trials provide suggestive evidence that external validity may be violated in such evaluations
- Hunt evaluates the Opower programme in which context households were mailed “Home Energy Reports” to provide energy conservation tips and comparisons with neighbours. 111 RCTs in different contexts have been carried out. Taking the first 10 and trying to predict the further result overpredicts the ATE by 0.41-0.66 ppts (equivalently 560 to 920 mn. USD worth of retail electricity)
- This doesn’t necessarily mean that non-experimental studies on much more data (sites) should be run and “sacrifice internal validity for external validity”. In the Opower context, non-experimental results are very poor.
- There are two different levels:
 - In single sample to single target site predictions, external unconfoundedness means homogeneous site effects, that is the treatment effect is the same for each site. Quite unrealistic.

$$\tau_{s'}(x) = \tau_{s''}(x) \text{ for sites } s' \text{ and } s''$$
 - In multi sample to multi target site predictions, external unconfoundedness means no site selection bias, that is, the *expected value* of the treatment effect among evaluated sites is equal to the *expected value* of the treatment effect among non-evaluated sites. This is a weaker assumption.

$$E[\tau_s(x)|D_s = 1] = E[\tau_s(x)|D_s = 0] \text{ where } D \text{ is a dummy for evaluation}$$

Fischer and Karlan (2015). *The Catch-22 of External Validity in the Context of Constraints to Firm Growth.* AER P&P

- Short paper pointing towards the tension between catching specific mechanisms in an environment with many failures, highly specific evaluations to identify mechanisms and external validity and learning in a bigger context
- Small and medium-size firms report a set of challenges they are facing and would need assistance with. Sales & marketing and accounting & cash flow are reported by many but challenges display otherwise a great heterogeneity.
- Removing one obstacle may be ineffective in the presence of an “O-ring” (Kremer) of many obstacles
- Potential implications: Perform diagnostics prior to some training treatment to find out what treatment may be needed in a specific context. Problems include that the diagnostic process itself may be a treatment and very precise measures are needed both before and during the evaluation
- “External validity is gained by having an empirically-validated theory of why something is working, and that theory ought to include relevant contextual factors”. However, equally, empirically it requires a representative sample. This is a tension. Empirically, you would want a broad context. Theoretically, you need a very narrow setting to identify a specific mechanism.
 - Conclusion: don’t expect too much from one single paper. The scope for learning from one paper is very limited.

Bandiera et al. (2016). *Do Women Respond Less to Performance Pay? Building Evidence from Multiple Experiments.* WP

- They use a Bayesian Hierarchical Framework to build evidence on whether the gender gap in wages can be partly explained by women being less responsive to performance pay
- BHF uses existing data to explore the question without being restricted to a subgroup and uses the data itself to estimate to what extent it is informative about a common phenomenon
- Results: (i) slightly *positive* gender-incentive coefficient (0.07sd), so women respond even slightly *more* to performance pay; (ii) the general response to performance pay is large (0.28sd)
- typical approach in evidence aggregation:

- FE pooling model – all studies estimate the same effect η , but have idiosyncratic variation. The pooling estimator is the weighted average where the weights are the inverse sample variances. In the presence of cross-study heterogeneity, the estimated variance of the pooling estimator will be too small, though.
- RE pooling model – takes between-study variance τ into account and assumes it is constant. Also assumes that the true effect is study-specific η_s . Therefore, the estimate from a study $\hat{\eta}_s$ is normally distributed around the true estimate, and the true study estimate is normally distributed around the common true effect with variance τ . The RE pooling estimator is again a weighted average where the weights are now composed of both the within-study and between-study variance.
- τ is essentially a measure of the comparability of studies (same outcomes vs. different outcomes for example)
- The BHF goes even one step further than the RE model and assumes that even the common true effect and between-study variable are random variables, distributed in a particular way.
- Only experimental literature is considered here. However, even if you were not sure about internal validity, there are methods to deal with this.

King & Nielsen (2016). Why propensity scores should not be used for matching. WP

- Discuss propensity score matching (PSM) relative to other matching methods such as MDM, minimising the geometric distance between X characteristics in several dimensions (not just reducing the dimensions to a scalar like in PSM) and coarsened exact matching (X covariates are coarsened so much that you get an exact match between treatment and control with the coarsened data)
- Let the ‘caliper’ be δ , determining how close you want to match treatment and control units.
- The set of matches under MDM is:

$$Set_{MDM} = M(X \mid \sqrt{(X_i - X_j)S^{-1}(X_i - X_j)} < \delta)$$

- Under CEM, it is

$$Set_{CEM} = M(X \mid C_\delta(X_i) = C_\delta(X_j))$$

Where $C_\delta(X)$ is a coarsened measure of X, e.g. college dummy instead of yrs. of schooling

- For PSM, the dimensionality of X is reduced to a scalar along which treatment and control group are matched. In that sense, PSM tries to approximate a *completely randomised experimental design* (CR) while other matching methods approximate a *fully blocked randomised experimental design* (FR) which is the idea that units are first matched in pairs with the same covariates X before one is randomly assigned to treatment and the other to control. CR is taking any population and assigning a share of it randomly to treatment. A FR design gives 0 imbalance by definition, a CR design only gives that asymptotically.
- They consider several relevant criteria:

- Model dependence: The unobserved counterfactual to a given treatment unit is generated by a statistical model (e.g. regression, with or without heteroskedastic errors, under some parametric assumptions, etc.). Model dependence measures the degree to which the generated counterfactual varies depending on the statistical model used.
- Imbalance: The degree to which treatment and control units are comparable
- Bias. Suppose a number of different statistical models generate different counterfactuals and thus different estimates of the treatment effect, $\hat{\tau}$. Assume the unrealistic case where all of them are unbiased conditional on the model. Then, taking the average of all of them gives you an unbiased estimate of the treatment effect. Consider, by contrast, a 'human-in-the-loop estimator' which is essentially the researcher choosing one of the 10 treatment effects, e.g. the greatest one.
Clearly, this is a *biased* estimator of the true treatment effect
- The idea of matching is to reduce model dependence and subsequent bias. The data is pruned in an iterative procedure in which bad matches are excluded until only close matches exist. Then the statistical analysis is run on the remaining data. The PSM paradox is that PSM first reduces imbalance, model dependence and thus bias, but after a certain point of pruning increases imbalance, model dependence and thus bias. This is because essentially – unlike CEM and MDM – conditional on the propensity score, pruning using PSM is essentially random pruning (that is, just unnecessarily reducing information by reducing the dataset).
 - Two observations with the same (or very similar) propensity score may have very different X matrices. The different elements of X that are combined in a function to give the propensity score can go in different directions in exactly such a way that the propensity score is similar in spite of substantial differences for individual elements in X. Matching on the elements of X (like in CEM or MDM) does not encounter this problem.
- They show this result in simulations and by reanalysing real data used in actual publications.

Meager (2016). Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature.

s. also brief presentation from Dev Reading Group

- Uses 7 RCTs on microcredit to aggregate the evidence with a Bayesian hierarchical model on ATEs and quantile effects
- Innovations in the paper relative to previous work
 - Quantile analysis
 - Treatment effect on the average and the dispersion
 - Dealing with non-continuous variables
- For various household outcomes below the 75th percentile, negligible impact on the distribution found. No generalizable prediction above this percentile
 - This is very different from a simple (full pooling) aggregation model: The uncertainty (confidence intervals) at the tails are *much* larger

- Full pooling would suggest significantly positive effects for business profits, revenues and expenditures at the upper tail where the BHM gives just massive uncertainty
- See Dev reading group slides for a brief description of the BHM idea

Meager (2016). Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of 7 Randomized Experiments.

- Uses BHM to aggregate evidence from 7 randomised experiments on microcredit
- In line with the Bandiera et al. paper, this assumes that there is a distribution for each treatment effect estimate and on a higher level a distribution of all true treatment effects across sites.
 - This nests the “full pooling” model in which we assume there is no distribution for true treatment effects (the true treatment effect is exactly the same in each site and all variation observed is sampling variation) and the “no pooling” model (which is assuming that the distribution for true treatment effects has infinite variance, i.e. there is no common element across sites and all variation observed is heterogeneity in the treatment effect across sites)
- This is estimated using a parameterised likelihood function and prior distributions and a Bayesian updating rule
- Effect of microcredit is found to be likely positive, but small in magnitude relative to the control group average. Negative effects cannot be ruled out
- 60% pooling on treatment effects found
- 27% of households who previously operated businesses before the microcredit expansion almost entirely drive the cross-study heterogeneity.
- She uses experimental studies, not questioning the causality. Cites building a structure accounting for differing strength of exclusion restrictions as area for future research.
- Has very different results than simple reviews just providing a summary of different estimates: *“These differences demonstrate that while the econometric issues involved in evidence aggregation may seem abstract or technical, they can have a major impact on the conclusions we draw from the evidence we have. If development economists seek to produce reliable bodies of generalizable evidence for policymakers to use, then aggregating results and assessing external validity via Bayesian hierarchical models should become a routine part of translating research into policy.”*
-