Empirical Methods for Policy Evaluation II

Matteo Bobba

Toulouse School of Economics (TSE)

TSE PhD Program
Fall 2021

Part 5: Job Search Models & Diff-in-Diffs

- Context
  - **Welfare policies and informality in Mexico**
- Methods
  - Flinn and Heckman (JoE, 1982)
- Application
  - Bobba, Flabbi and Levy (IER, 2022)

Labor Markets in Latin America

- More than half of the labor force is in the informal sector
  - Workers not contributing to and not covered by the social security system
  - Informal employees and (most of the) self-employed
- Patterns in the data are not consistent with either a segmented or a competitive view of the labor market
  1. Individuals transit back and forth between formal and informal jobs
  2. Wage/productivity distributions overlap
  3. Mix of formality status within the same firm
- Informal workers have started to gain access to non-contributory social security programs

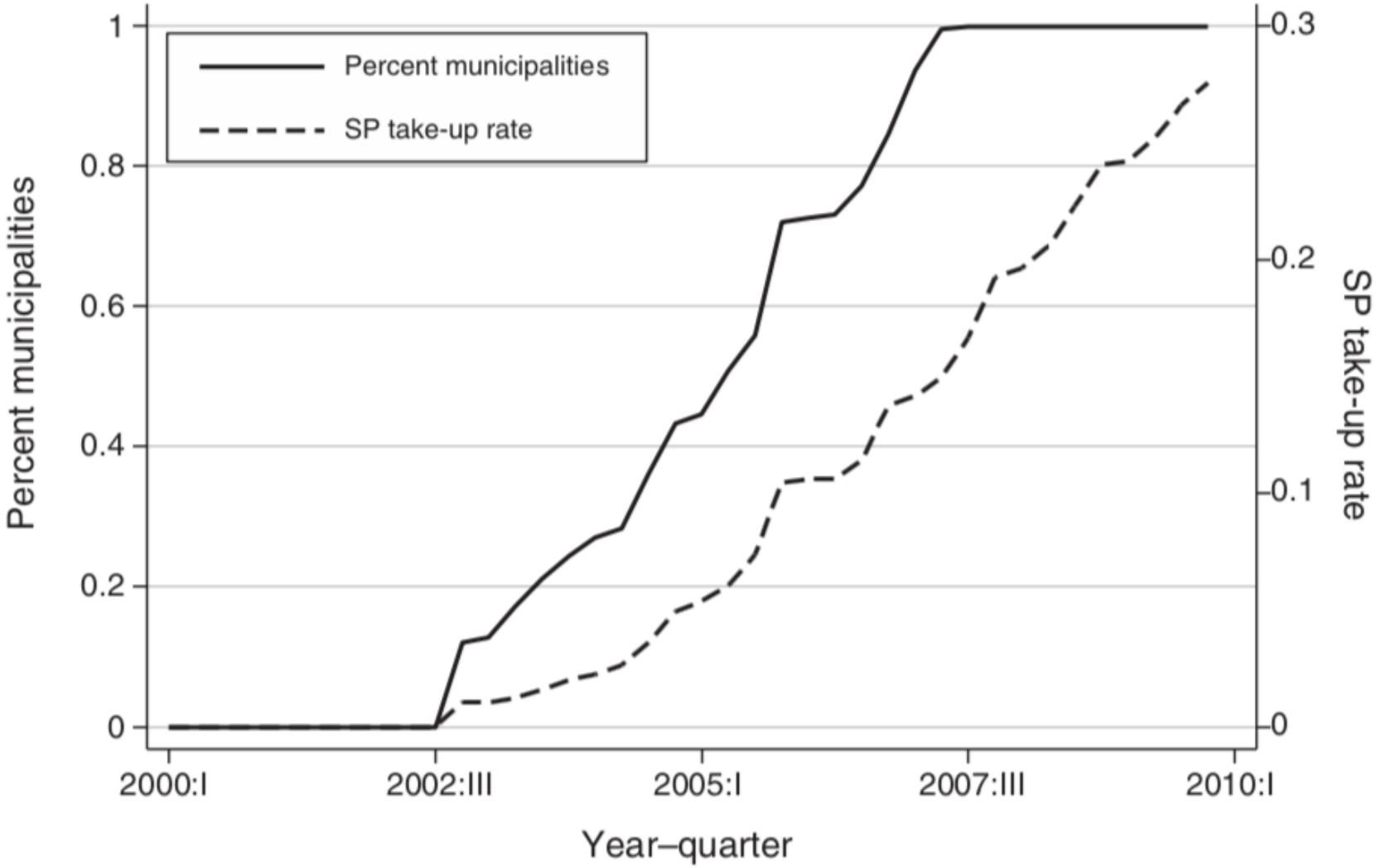
The Costs of Informality

- Informality may well be an optimal choice to a given institutional setting
  - It may provide de facto flexibility for firms and workers to cope with adverse shocks
- Still, its pervasive diffusion may generate short- and long-term costs:
  - Hinders fiscal capacity and the provision of public goods
  - Subsidy for smaller and often less productive firms
  - Worsen hold-up problems in investment decisions of firms and workers

The *Seguro Popular* Program

- In the early 2000s, implemented a health reform aimed at enhancing health insurance to individuals not covered by social security
  - Enrollment in the program is voluntary, and granted upon compliance with simple requirements
  - Health center and a family doctor plus a package of health services
- SP is a transfer to informal sector workers and to the non-employed
  - Negative impact on employment and/or formality rates
  - Wages in equilibrium might compensate for the increase in benefits in the informal sector

Staggered Rollout of the SP



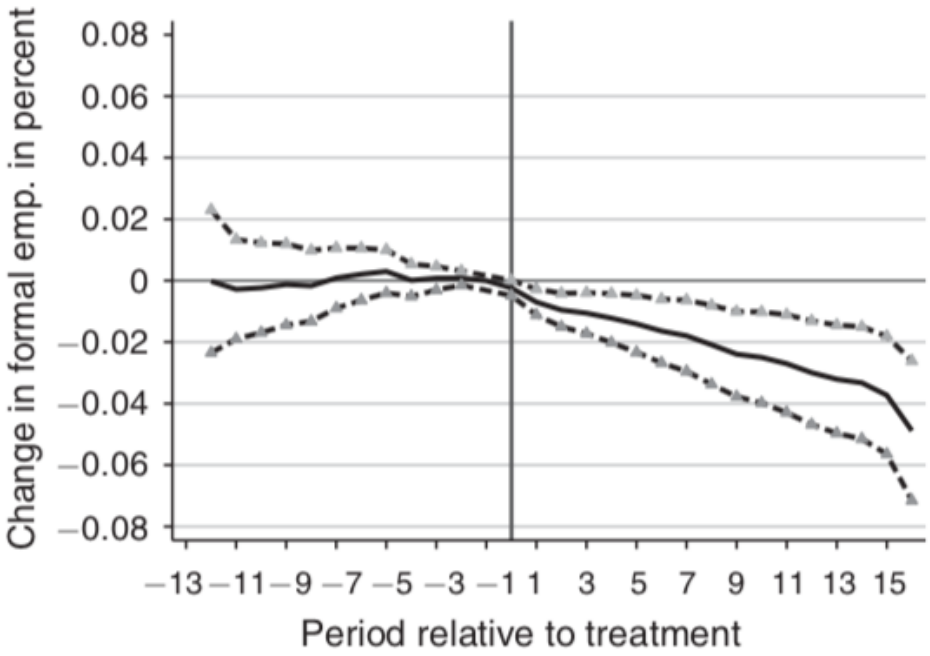
SP implementation at the municipality level

TABLE 1—DETERMINANTS OF MUNICIPALITY AFFILIATION WITH THE SP PROGRAM

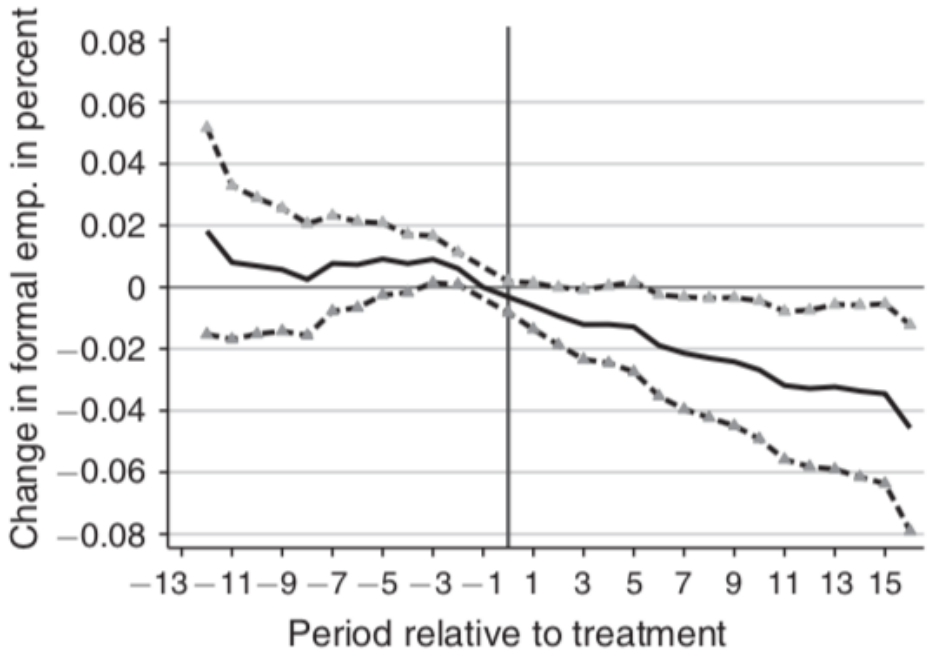
	Panel (1)	Post-pilot (2)
log population	−1.320***	−0.513**
	(0.281)	(0.219)
log state population	2.339**	0.959
	(0.924)	(0.573)
Share of insured population	1.921	0.746
	(2.793)	(2.293)
Urban	0.460	0.434
	(0.425)	(0.368)
log median wage	−0.300**	−0.117
	(0.143)	(0.132)
Years of schooling	−0.497	−0.197
	(0.486)	(0.323)
Unemployment rate	3.069	9.438
	(15.037)	(17.086)
PRD	3.418	4.092***
	(2.250)	(0.806)
PRI	−0.188	0.375
	(1.752)	(1.063)
Poverty index (food)	0.040	−0.001
	(0.056)	(0.021)
Poverty index (income)	0.029	0.026
	(0.049)	(0.022)
Share of aged < 24	−11.754	−5.677
	(7.223)	(5.108)
Share of aged > 24 and < 40	11.215	11.419*
	(8.526)	(5.662)
Share of males	1.451	1.625
	(5.979)	(5.374)
Industry shares	YES	YES
Observations	1,392	1,052
R^2	0.244	0.213

The Effects of the SP on Firms

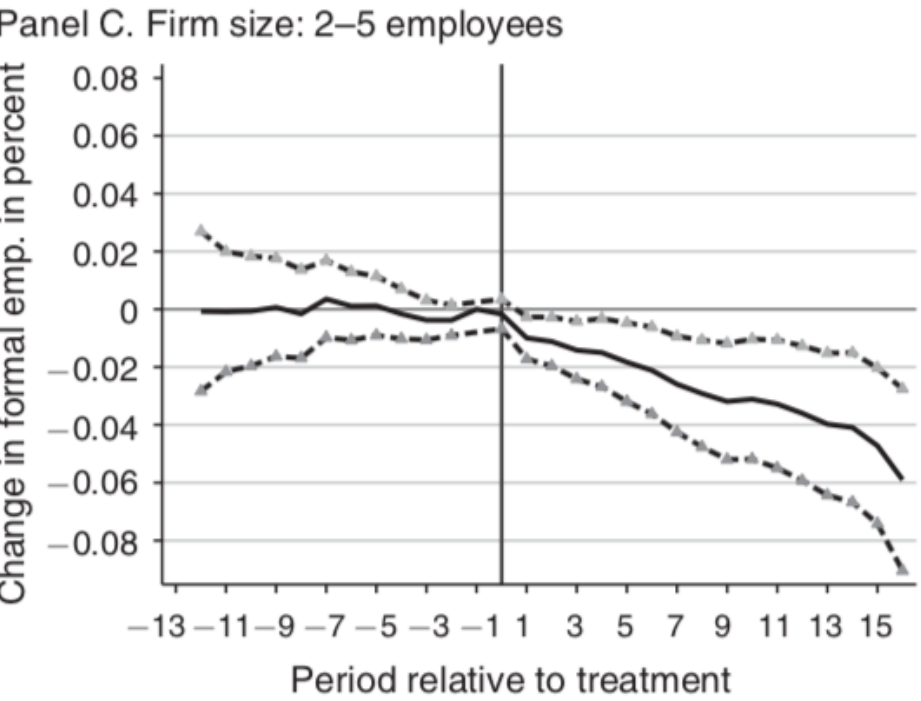
Panel A. Total number of employers

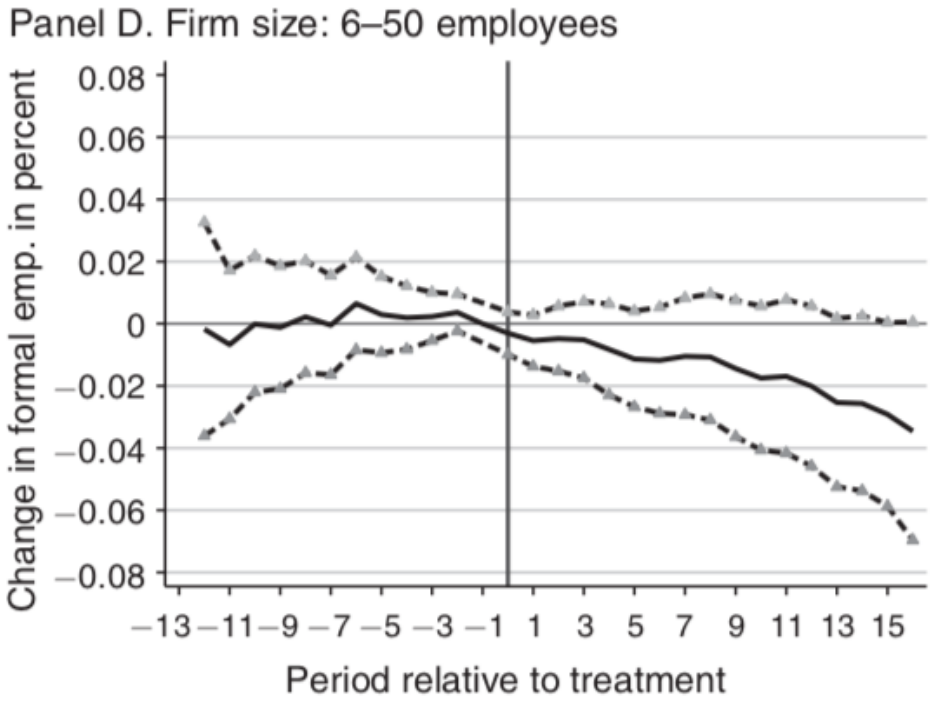


Panel B. Firm size: 1 employee



The Effects of the SP on Workers





Part 5: Job Search Models & Diff-in-Diffs

- Context
  - Social Protection Policies and Informality in Mexico
- Methods
  - **Flinn and Heckman (JoE, 1982)**
- Application
  - Bobba, Flabbi and Levy (IER, 2022)

The Optimal Stopping Model

- Risk neutral individual in discrete time with preferences in $t = 0$ given by

$$\sum_{t=0}^{\infty} \beta^t c_t$$

- Start as unemployed, with consumption equal to b
- Jobs sampled sequentially. Each job is for life
- All jobs are identical except for their wages, and wages are given by an exogenous stationary distribution $F(w)$
- At every date, the individual takes a iid wage draw $w_t \in W$, and has to decide whether to take this or continue searching

Dynamic Programming Formulation

- Value function for the agent when he has sampled a job of $w \in W$ is

$$V(w) = \max\left\{\frac{w_t}{1-\beta}, \beta V + b\right\}$$

where V is the continuation value of not accepting a job:

$$V = \int_{\omega\in\Omega} V(\omega) dF(\omega)$$

- Combine these two equations and get:

$$V(w) = \max\left\{\frac{w_t}{1-\beta}, b + \beta \int_{\omega\in\Omega} V(\omega) dF(\omega)\right\} \tag{1}$$

Reservation Wage

- $V(w)$ is non-decreasing, therefore decision rule has a reservation value property
- Reservation wage is given by

$$\frac{w^\star}{1-\beta} = b + \beta \int_{\omega \in \Omega} V(\omega) dF(\omega) \tag{2}$$

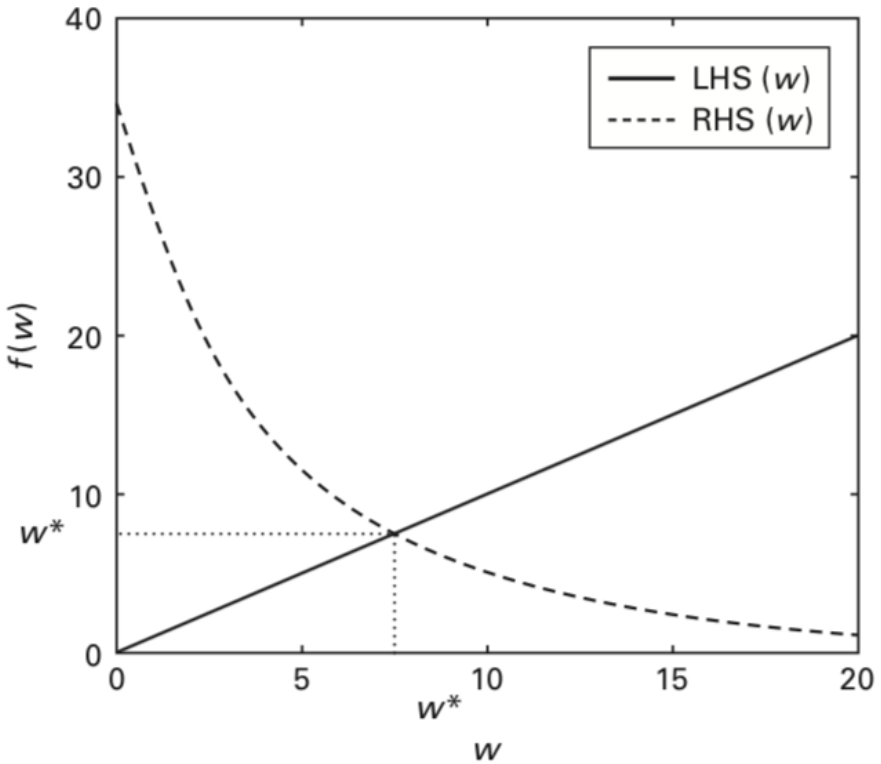
- Decision rule: $\forall w < w^\star$, $V(w) = \frac{w^\star}{1-\beta}$ and $\forall w \geq w^\star$, $V(w) = \frac{w}{1-\beta}$
- Therefore, (2) becomes:

$$\frac{w^\star}{1-\beta} = b + \beta \left[\frac{w^\star F(w^\star)}{1-\beta} + \int_{w \geq w^\star} \frac{w}{1-\beta} dF(w) \right] \tag{3}$$

Reservation Wage

- Since $\frac{w^\star}{1-\beta} = \int_{w<w^\star} \frac{w^\star}{1-\beta} dF(w) + \int_{w \geq w^\star} \frac{w^\star}{1-\beta} dF(w)$, we can rewrite (3) as:

$$w^\star = b + \frac{\beta}{1-\beta} \left[\int_{w \geq w^\star} (w - w^\star) dF(w) \right] \tag{4}$$



Taking the Model to the Data

- For a random sample of N workers we observe $\{\tilde{t}_u(i), w(i)\}_{i=1}^N$
- Job offers/termination arrive at random times with density between offers given by

$$\begin{aligned} q_u(t_u) &= \lambda \exp(-\lambda t_u), \lambda > 0 \\ q_e(t_e) &= \eta \exp(-\eta t_e), \eta > 0 \end{aligned}$$

- Reservation wage (4) in continuous time

$$w^\star = b + \frac{\lambda}{\rho + \eta} \int_{w^\star} (w - w^\star) dF(w)$$

- It is easy to show that $\partial w^\star / \partial \eta < 0$

Steady-state Proportions

- The probability that a randomly-sampled (at a point in time) individual is unemployed is

$$p(u) = \frac{\mathbb{E}(t_u)}{\mathbb{E}(t_e) + \mathbb{E}(t_u)} = \frac{[\lambda \tilde{F}(w^\star)]^{-1}}{\eta^{-1} + [\lambda \tilde{F}(w^\star)]^{-1}}$$
$$= \frac{\eta}{\eta + \lambda \tilde{F}(w^\star)}$$

- And conversely

$$p(e) = 1 - p(u) = \frac{\lambda \tilde{F}(w^\star)}{\eta + \lambda \tilde{F}(w^\star)}$$

Offered and Accepted Wages

- The optimal decision rule generates an endogenous truncation from below in the accepted wage distribution

$$g(w) = \frac{f(w)}{\tilde{F}(w^\star)}, w \geq w^\star$$

- This density is well defined, in the sense of integrating to 1 and being non-negative for all $w \geq w^\star$

Likelihood Contributions

- The likelihood of finding a randomly sampled agent in an ongoing unemployment spell is

$$L(t_u, u) = \tilde{f}_u(t_u)p(u) = \lambda \tilde{F}(w^\star) \exp[-\lambda \tilde{F}(w^\star)t_u] \times \frac{\eta}{\eta + \lambda \tilde{F}(w^\star)}$$

- The likelihood of finding an individual employed and earning a wage w is

$$L(w, e) = \frac{f(w)}{\tilde{F}(w^\star)} \times \frac{\lambda \tilde{F}(w^\star)}{\eta + \lambda \tilde{F}(w^\star)} = \frac{\lambda f(w)}{\eta + \lambda \tilde{F}(w^\star)}$$

Likelihood Function

- The likelihood function for a random sample of N individuals is then

$$L(w_1, .., w_{N_e}, t_1, .., t_{N_u}) = \prod_{i \in e} \left[\frac{\lambda f(w(i))}{\eta + \lambda \tilde{F}(w^\star)} \right] \times \prod_{i \in u} \left[\frac{\eta \lambda \tilde{F}(w^\star) \exp[-\lambda \tilde{F}(w^\star) t_u(i)]}{\eta + \lambda \tilde{F}(w^\star)} \right]$$

- And the associated log-likelihood is

$$\ln L = -N \ln[\eta + \lambda \tilde{F}(w^\star)] + N \ln \lambda + \sum_{i \in e} \ln[f(w(i))] +$$
$$+ N_u \ln[\tilde{F}(w^\star)] + N_u \ln(\eta) - \lambda \tilde{F}(w^\star) \sum_{i \in u} t_u(i)$$

Identification

- The primitive parameters that explicitly enter in $\ln L$ are λ, η and F
- The other parameters b and ρ only enter through $w^\star$, which is a function of all the structural parameters
- Parameter $w^\star$ is part of the support of F (distribution of wage offers is identified only conditional on $w \geq w^\star$)
  - This feature generates a non-standard likelihood function

Flinn and Heckman (1982)

1. Estimate the reservation wage as the minimum accepted wage: $\hat{w}^\star = \min(w_1, ..., w_{N_e})$
   - Order statistics are super-consistent (i.e. converge at rate N)
2. Maximize log likelihood with respect to λ, η and μ conditional on $\hat{w}^\star$
   - $F(w)$ needs to be recoverable: $F(w|w \geq \hat{w}^\star) = \frac{F(w)-F(w^\star)}{\bar{F}(w^\star)}, \forall w \geq \hat{w}^\star$
3. Plug estimated parameters into equation for $w^\star$ and solve for either b or ρ
   - Usually fix ρ and recover the value of b

Bobba, Flabbi and Levy (IER, 2022)

- An equilibrium search model where:
  1. Search frictions generate mobility between formal and informal jobs
  2. Match-specific productivity and bargaining generate overlapping wage distributions
  3. Both ingredients generates a mix of formal and informal jobs in equilibrium
- One important long-term "cost of informality": Under-investment in education
  - Same features that create informality may also distort returns to schooling
  - Hold-up ex-ante investments in human capital

Environment

- Timing
  1. Schooling decision
  2. Searching status decision
  3. Labor market dynamics
- Labor Market States
  1. Unemployed
  2. Self-employed
  3. Informal Employee
  4. Formal Employee

Schooling Decision

- Irrevocable decision about schooling level h
- Discrete choice: $h \in \{0, 1\}$
- Individual-specific heterogeneity
  - costs $\kappa \sim T(\kappa)$
  - opportunity cost - PDV of participating in LMK as $h = 0$
- Schooling decision has reservation value rule: only agents with $\kappa < \kappa^{\star}(y)$ will acquire $h = 1$
- All labor market parameters are allowed to be schooling-specific

Searching-status Decision

- Irrevocable decision about searching as:
  - Self-employed
  - Unemployed
- Individuals search for an employee job in both states but receive offers at different rates: $\gamma_h < \lambda_h$
- Self-employment income $y \sim R(y|h)$
- Searching status decision has (again) reservation rule property: only agents with $y \geq y^\star(h)$ search for an employee job while also working as self-employed

Labor Market Dynamics

State	PDV	Shock	Flow Utility
Workers:			
Unemployed	$U(h)$	λ_h	$\xi_h + \beta_{0,h} B_0$
Self-Employed	$S(y,h)$	γ_h	$y + \beta_{0,h} B_0$
Informal Employee	$E_0[w,y,h]$	η_h, χ_h	$w_0(x;y,h) + \beta_{0,h} B_0$
Formal Employee	$E_1[w,y,h]$	η_h, χ_h	$w_1(x;y,h) + \beta_{1,h} B_1[w_1(x;y,h)]$
Firms:			
Vacancy	$V[h]$	ζ_h	ν_h
Filled Informal Job	$F_0[x,y,h]$	η_h, χ_h	$x - w_0(x;y,h)$
Filled Formal Job	$F_1[x,y,h]$	η_h, χ_h	$x - (1+t) w_1(x;y,h)$

- Match-specific heterogeneity: $x \sim G(x|h)$
- One-shot penalty for firms hiring illegally: $c_h w_0(x;y,h)$
- Matching function determines $\{\lambda_h, \gamma_h, \zeta_h\}$: $m_h = (u_h + \psi_h s_h)^{\iota_h} (v_h)^{1-\iota_h}$

Labor Market Institutions and Wage Determination

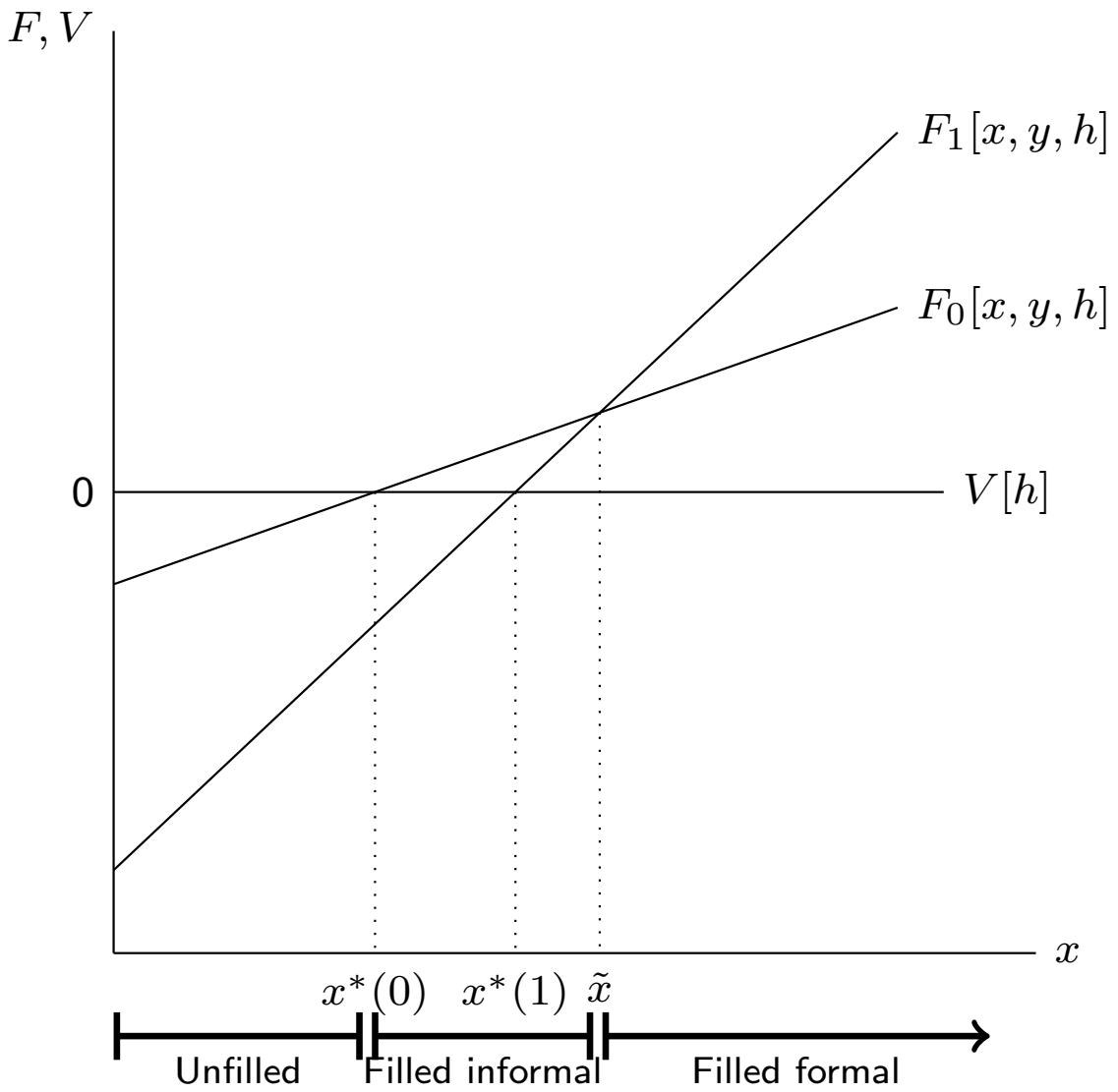
- Non-wage workers' flow value:
  - formal employee $= \beta_{1,h} B_1[w_1(x; y, h)] = \beta_{1,h}[\tau t w_1(x; y, h) + b_1]$
  - informal employee $= \beta_{0,h} B_0$
  - Notice: the endogenous b_1 introduces redistribution within and between schooling levels.

- Wage are determined by bargaining, conditioning on formality status endogenously posted by firms. Wage schedules (under free-entry of firms) are:

$$
\begin{aligned}
w_0(x; y, h) &= \frac{\alpha_h}{1 + \chi_h c_h} x + (1 - \alpha_h)[\rho Q(y, h) - \beta_{0,h} B_0] \\
w_1(x; y, h) &= \frac{\alpha_h}{1 + t} x + \frac{(1 - \alpha_h)}{1 + \beta_{1,h} \tau t}[\rho Q(y, h) - \beta_{1,h} b_1]
\end{aligned}
$$

where: $Q(y, h) \equiv \max\{S(y, h), U(h)\}$

Equilibrium Representation



Empirical Implications

- Main stylized facts of informal labor markets are replicated in equilibrium:
  1. A mixture of formal and informal jobs is realized
  2. Formal employees have on average higher wages than informal employees. But their accepted wage distributions overlap
  3. Informal employees and self-employed have markedly different labor market dynamics
  4. Some firms hire formal or informal workers at different points in time just as workers transit over time between different formality status

Data Sources

1. Mexico's Labor Force Survey (ENOE): Year 2005
   - Nonagricultural, full-time, male, private-sector, secondary-school workers between the ages of 25 and 55 who reside in urban areas
   - $w \equiv$ Hourly wages as employee, main job after labor contributions
   - $y \equiv$ Hourly labor income as self-employed, without paid employees
   - $f = 1$ if employee is contributing to the social-security fund; $= 0$ otherwise
   - $h = 1$ if Upper secondary completed $= 0$ if Lower secondary completed
2. Aggregate labor shares for Mexico in 2005
   - Total compensations per employee as percentage of GDP
3. Vacancy rates for 2005
   - Good coverage of vacancy posting in urban areas
   - Detailed information on the schooling level required for the job

Identification: Informality Parameters

- Different transition rates out of formal jobs and informal jobs identify χ_h
- The identification of β_1 and c_h is derived from the location and extent of the *overlap* between formal and informal accepted wage distributions
  - While movement in β_1 and c_h can achieve the same extent of the overlap, they do so by moving its location in different directions
  - This generates different shape in the accepted wage distribution of formal and informal employees
- Repeating the same argument over the y distribution generates
  - The large overlap observed in the data
  - Useful variation to separately identify the parameters

Identification: Informality Parameters (cont'd)

- The identification of β_0 requires the use of additional information
  - We exploit the roll-out of the *Seguro Popular* (SP) program in 2005-2006
  - In terms of our model, it can be seen as an exogenous increase in B_0
- Variation in B_0 across individuals identify β_0
  - As long as this variation is not correlated with changes in the primitive parameters of the model
  - Labor market outcomes pre-policy seem balanced

Identification: Self-employment and Schooling Parameters

- $R(y|h)$: Identified by observed self-employment earnings, once we assume a recoverable primitive distribution
  - We assume lognormal with parameters $\{\mu_{y,h}, \sigma_{y,h}\}$
- $T(\kappa)$: The threshold crossing decision rule allows for the identification of one parameter from the proportions of individuals in the two schooling levels

$$\frac{1}{n}\sum_{i=1}^{n} h_i = \int_y T(\kappa^*(y))dR(y|0)$$

  - We assume a negative exponential with parameters δ

Identification: Matching and Demand Side Parameters

- $\{\psi_h, \iota_h\}$: define tightness tightness $\omega_h \equiv \frac{v_h}{u_h + \psi_h s_h}$ so that:

$$\begin{aligned} \psi_h &= \frac{\gamma_h}{\lambda_h} \\ \iota_h &= \frac{\ln \omega_h - \ln \lambda_h}{\ln \omega_h} \end{aligned}$$

- With the matching function identified, we can compute the demand side parameters
  - $\zeta_h = \omega_h^{-\iota_h}$
  - ν_h: Use firm's value function and impose free entry

Identification: Bargaining Parameters

- Rewrite the wage schedules as

$$
\begin{aligned}
w_0(x;y,h) &= \alpha_h \left\{ \frac{x}{1+\chi_h c_h} - [\rho Q(y,h) - \beta_{0,h} B_0] \right\} + [\rho Q(y,h) - \beta_{0,h} B_0] \\
w_1(x;y,h) &= \alpha_h \left\{ \frac{x}{1+t} - \frac{[\rho Q(y,h) - \beta_{1,h} b_1]}{(1+\beta_{1,h}\tau t)} \right\} + \frac{[\rho Q(y,h) - \beta_{1,h} b_1]}{(1+\beta_{1,h}\tau t)}
\end{aligned}
$$

- α_h is governing the portion of the surplus appropriated by the worker through the wage
- Labor shares are the ratio between the aggregate value of worker's wages $w_f(x;y,h)$ and the aggregate value of production x
- They provide sufficient information to identify α_h

Identification: Unobserved Ability Types

- Type is known to the individual but unobserved in the data. We denote each type with k and its proportion in the population with π_k.

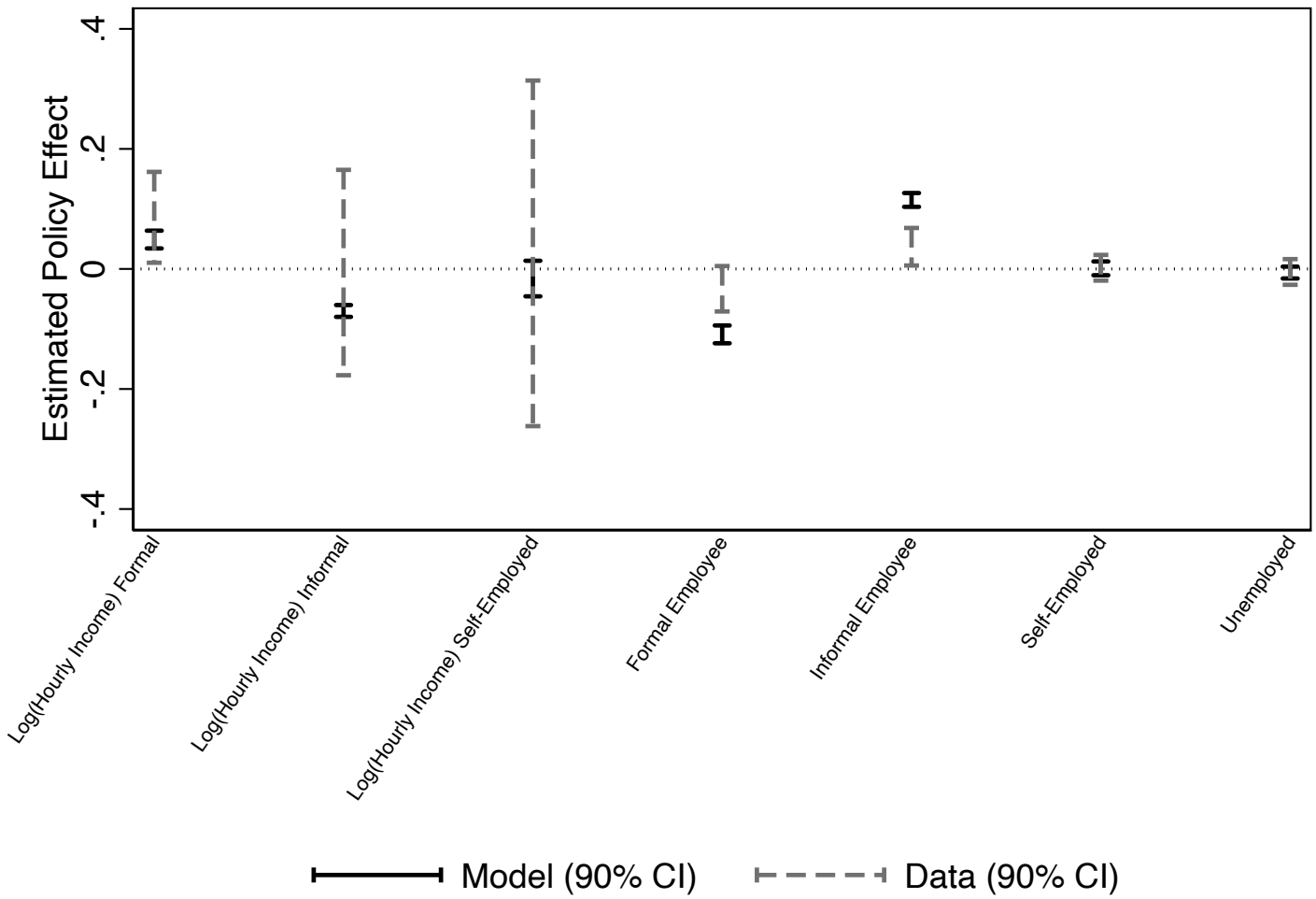
$$x|k = a_k^G x$$
$$y|k = a_k^R y$$
$$\kappa|k = a_k^T \kappa$$

- Duration dependence in unemployment identifies these parameters
  - Hazard rates at three and six months for both schooling levels
- Assume: $K = 2$
  - type $k = 1$ normalized to $a_1^T = a_1^R = a_1^G = 1$
  - type $k = 2$ exhibiting $a_2^T < 1; a_2^R > 1; a_2^G > 1$

Estimation

- For each schooling and treatment group, we match the following moments
  - Proportions of individuals in each labor market state
  - Accepted wage distributions of formal and informal employees
    - Mean and SD: overall and by quintiles
    - Overlap, as measured by proportion of formal employees for each quintile of the informal accepted wage distribution
  - Accepted earnings distributions of self-employed
    - Mean and SD
  - Transitions between LMK states (yearly)
  - Hazard rates out of unemployment (at 3 and 6 months)
  - Labor Shares
- Back-out demand-side parameters using vacancy rates

Out-of-Sample Model Validation



Returns to Schooling

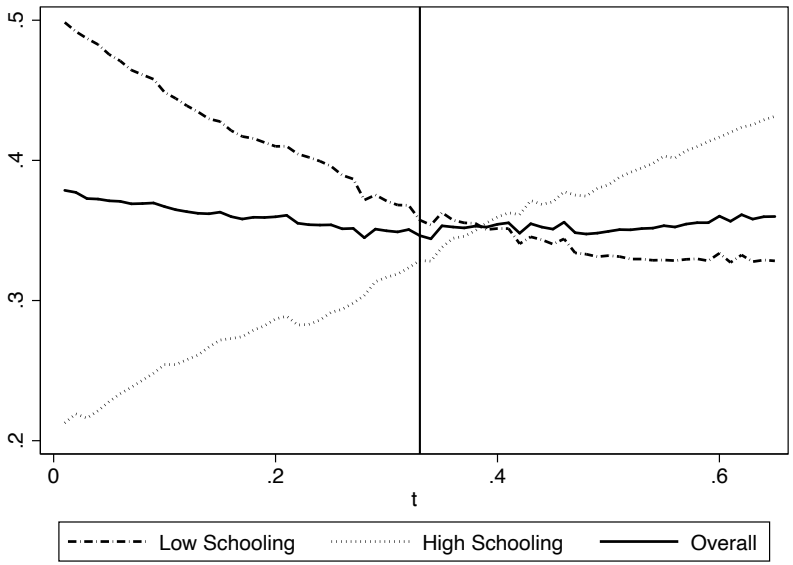
Ability:	Low $k=1$	High $k=2$
PDV of Labor Market Search:		
$\int_y Q(y,h)dR(y\|h)$	0.309	0.278
Average Accepted Wages:		
F: $E_h\left[w_1 \mid \tilde{x}(y,h) \leq x\right]$	0.479	0.435
I: $E_h\left[w_0 \mid x_0^*(y,h) \leq x < \tilde{x}(y,h)\right]$	0.281	0.296
Average Offered Wages:		
F: $E_h\left[w_1 \mid y < y^*(h)\right]$	0.213	0.210
F: $E_h\left[w_1 \mid y \geq y^*(h)\right]$	0.213	0.204
I: $E_h\left[w_0 \mid y < y^*(h)\right]$	0.133	0.134
I: $E_h\left[w_0 \mid y \geq y^*(h)\right]$	0.142	0.136

Counterfactual: The Equilibrium Effects of Informality

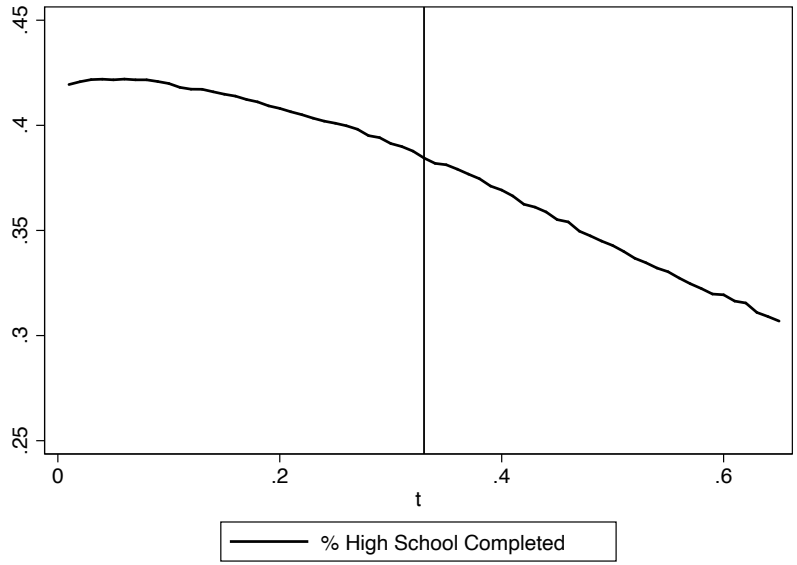
Model:	Firms can only offer a formal contract			
Specifications:	Baseline Model	Exogenous Schooling	Exogenous Contact Rates	Hosios-like Condition ($\alpha = \iota$)
Flow Welfare:				
Total	-0.0596	-0.0750	-0.0020	0.0478
Workers	-0.0460	-0.0599	0.0166	0.0570
Firms	-0.2821	-0.3219	-0.3055	-0.1589
Labor Market Proportions:				
Unemployed	0.0213	0.0636	0.0019	-0.0459
Self-employed	0.3353	0.3526	0.3625	0.2329
Formal Employees	0.0275	-0.0146	-0.0376	0.0076
Schooling Outcomes:				
% HS Completed	0.1029	–	0.0781	0.1501
% High Ability in HS	0.0538	–	0.0569	0.0628

NOTE: Relative changes wrt the benchmark model. Hosios increases α from 0.56 to 0.73.

Counterfactual: Changes in Payroll Tax Rate (t)



(a) Informality



(b) Schooling

Main Take-ways from Estimated Model

1. Returns to schooling are substantial
2. Informality is welfare improving but:
   - Significantly more so for firms than workers
   - Reduces human capital accumulation (hold-up problem is exacerbated by informality)
3. Payroll tax rate has a non-intuitive impact on equilibrium outcomes
   - Informality rate not a good indicator for policy
   - Redistributive components within the formal system are often ignored in the policy debate

Wrapping up on Job Search Models and Diff-in-Diffs

- Relevant institutional features are included in the model in a tractable way
- But these extra parameters are hard to separately identify using standard labor market survey data
- The staggered roll-out of the welfare policy provides additional variation outside the model environment
  - Identify the (average) valuation of non-contributory benefits
  - Validate the model on a different time period by simulating one-step ahead