Selection and Heterogeneity in the Returns to Migration*

Eduardo Cenci†, Marieke Kleemans‡, and Emilia Tjernström§

September 7, 2023

Abstract

There is considerable debate on the returns to rural-urban migration in developing countries, and magnitudes differ sharply depending on the method used. We aim to reconcile these divergent estimates by explicitly accounting for the role of heterogeneity in the returns to migration. We use detailed longitudinal data from four developing countries—Indonesia, South Africa, China, and Tanzania—where we observe the location choices and labor market outcomes of tens of thousands of adults over multiple periods. We model self-selection into migration in a multi-period Roy model that incorporates worker heterogeneity in both absolute and comparative advantage. We then estimate a correlated random coefficient model that considers both types of heterogeneity. This model lets us extrapolate the returns identified from switcher sub-populations to non-switchers—a group of particular interest to policymakers deciding whether to encourage migration as a development strategy. Our results reveal considerable heterogeneity in the returns to migration and show a clear pattern in the relationship between absolute and comparative advantage across countries: those with the lowest productivity in rural areas stand most to gain from migrating. This suggests that migration is a pro-poor strategy but that barriers to migration may prevent workers from realizing their potential. As such, individuals appear to be inefficiently sorted across space; therefore, encouraging migration could lead to large returns.

Keywords: rural-urban migration, sectoral mobility, heterogeneity, self-selection

JEL codes: J24, J43, O13, O15, R23

---

* Corresponding author: Marieke Kleemans (kleemans@illinois.edu). We are grateful to David Buller for helping prepare the data. This project greatly benefited from conversations with David Albouy, Dan Bernhardt, Joshua Deutschmann, Paul Castañeda Dower, Andy Garin, Douglas Gollin, Robert Jensen, Thomas Lemieux, Nicholas Y. Li, Ana Paula Melo, Edward Miguel, Mushfiq Mobarak, Suresh Naidu, Laura Schechter, Jeff Smith, Sergio Urzua, and seminar and conference participants at the University of Illinois, Urbana-Champaign, the MWIEDC 2023, and the CDES Sustainable Development Conference 2023. All errors are our own.

† University of Nebraska at Omaha

‡ University of Illinois, Urbana-Champaign

§ Monash University

Introduction

As economies undergo structural transformation, labor tends to migrate out of rural areas into higher-productivity sectors in cities. While most areas of the world have undergone rapid urbanization in the past 50 years, we still observe striking income and consumption gaps between rural and urban areas (Lagakos, 2020). Across the developing world, the average urban resident earns 2 to 3 times more than their rural counterpart.[1] Earnings, consumption, and productivity differentials stubbornly persist, even after researchers control for cost-of-living differences, educational attainment, and a host of other observable characteristics (Young, 2013; Gollin, Lagakos and Waugh, 2014; Herrendorf and Schoellman, 2018).

An ongoing debate in both academic and policy circles centers on why these residual gaps remain. One interpretation is that labor is misallocated across space, perhaps partially due to policies. This would imply that policy changes that reduce labor allocation frictions could substantially increase productivity and welfare. Another explanation is that workers sort across sectors based on characteristics that are known to them but not observed in the data. From a policy perspective, this latter explanation makes wage gaps much less exploitable.

A key input into these debates and the accompanying policy decisions is reliable estimates of the returns to migration. In other words, we would like to have estimates of the causal effect of migration on consumption or earnings. Cross-sectional wage gaps, such as those in the structural transformation literature do not reflect the returns to migration because migrants differ from non-migrants in many ways, some of which are unobservable to the econometrician. Two recent studies aim to address the issue of unobserved heterogeneity by using panel data methods that account for time-invariant individual-level characteristics. Alvarez (2020) and Hamory, Kleemans, Li and Miguel (2021) estimate the earnings gaps in data from Brazil, Indonesia, and Kenya. Hamory *et al.* (2021) find that

[1] Restuccia, Yang and Zhu (2008), Caselli (2005), and Lagakos and Waugh (2013) document that urban and non-agricultural workers earn more than rural agricultural workers in almost all countries but that the gap is particularly large in lower-income countries, even though rural areas in these countries hosts a larger share of the workforce.

including individual-level fixed effects reduces the earnings gaps to between 1 and 24 percent—a much lower gap than the 200–300 percent seen in cross-sectional studies. Leveraging an experiment in Bangladesh, in which people randomly received a financial incentive to migrate to urban areas in the agricultural off-season, Bryan, Chowdhury and Mobarak (2014) obtain a causal estimate of the returns to migration and find an average return of 30 percent for those who moved in response to the experimental incentive.

We augment the findings of the recent literature with two key contributions. First, we relax the implicit assumption in Hamory *et al.* (2021) that unobserved skills have the same returns in both sectors.[2] We follow the many occupational and sectoral choice models that build on Roy's seminal 1951 paper and allow workers to have sector-specific skills that are rewarded differently in the rural and urban labor markets. Empirically, this suggests that we should use a random-coefficient model in order to allow the slope coefficients—including the returns to migration—to vary at the individual level. Further, we acknowledge that workers' (unobserved) comparative advantage may play a role in determining both earnings and sectoral choice. Once we allow for this correlation, it puts our empirical model in the class of *correlated* random coefficient models (CRC).

Second, building on a group random coefficient model, we show that an additional linearity assumption enables us to extrapolate returns to sub-populations whose returns are unidentified in standard models. Specifically, we can impose, and test the appropriateness of, a restriction on response heterogeneity that assumes returns to migration are linear in workers' comparative advantage. This contribution relates to a vast literature on microeconomic policy evaluation that examines models with heterogeneous effects and endogenous regressors. In such models, including CRC models, it is generally challenging to extrapolate from local treatment effects, defined over some range of the regressors, to a policy-relevant parameter (Imbens, 2007). Borrowing a thought experiment from a related paper by Lagakos, Marshall, Mobarak, Vernot and Waugh (2020), consider a policymaker deciding whether to allocate government resources towards reducing the costs of internal migration. Typically, she would be most interested in knowing the expected effect

[2] To be clear, this assumption is common to all fixed effects panel data models, not just Hamory et al. (2021). In addition, because the standard fixed effects estimator depends on a conditional independence assumption, it is biased in the presence of treatment effect heterogeneity that depends on unobservable fixed effects.

for a randomly drawn person from the population or the overall average effect of the policy. In general, if the true causal returns to migration are heterogeneous, then neither fixed-effects approaches nor experimental studies recover this policy parameter. Our approach takes advantage of the fact that our data contain different subpopulations, defined in terms of their complete migration histories. These subpopulations—which we refer to as "types"—include workers who move between the urban and rural sectors, who we call switchers. Similarly, we can refer to workers who remain in the location where we first observe them as stayers.

We start with an unrestricted random coefficient model that nests the restricted model as a special case, as in Tjernström, Ghanem, Cabanillas, Lybbert, Michler and Michuda (2023).[3] This group random-coefficient model (GrRC) is essentially a fixed-effects model that allows for subpopulation-specific marginal effects. By assuming that returns to migration are linear in workers' comparative advantage, we obtain a set of testable extrapolation restrictions. If we do not reject the linearity in comparative advantage restriction (LCA), then we are able to extrapolate returns to non-switcher subpopulations.

The theoretical framework we develop in this paper is closely related to the two-period Roy (1951) models in Lagakos and Waugh (2013) and Lagakos *et al.* (2020). In the first, workers' subsistence constraints and their sorting across sectors explain productivity gaps between the agriculture and non-agriculture sectors in different countries. In the second, workers with different migration costs and returns sort themselves across multiple regions generating different aggregate estimates for rural-urban consumption gaps. We build on these models to develop a multi-period migration model that allows for back-and-forth migration between rural and urban areas, thus matching migration trajectories observed in our data and the often-temporary nature of migration. We maintain the core assumption in Lagakos and Waugh (2013) and Lagakos *et al.* (2020) —that workers are endowed with location-specific skills—but we modify assumptions on the role of preferences and prices in those models so that the resulting migration decision rule in our setting maps onto a correlated random coefficient model in the likes of Lemieux (1998) and

[3] In our context, the time homogeneity assumption means that, conditional on the worker's time-invariant characteristics and complete migration history, the distribution of period-specific productivity shocks does not vary over time.

Suri (2011). As a result, we can estimate heterogeneous returns to rural-urban migration using the GRC framework discussed above.

We estimate our model using detailed longitudinal data from four developing countries: Indonesia, South Africa, China, and Tanzania. These datasets have between 15,600 and 49,400 individuals each, for whom we observe location choices and labor market outcomes for three to five periods spanning several years.The model allows us to estimate returns to moving between rural and urban locations. We can estimate returns on consumption as our main welfare measure, and present results on income in the appendix.

Our first set of results, obtained with pooled OLS and panel regressions with individual fixed effects, corroborate findings in the literature. We find large average rural-urban gaps in consumption ranging from 39 log points in Indonesia to 66 log points in Tanzania. These gaps decrease with the inclusion of controls (to 21–59 log points) and individual fixed effects (to 5–16 log points). Because including individual fixed effects effectively restricts the identifying variation to a sample of switchers, we also restrict our estimations to that sample and find similar estimates (8–17 log points with controls). This supports our claim that returns to switchers and non-switchers are different.

Then, we turn to the estimation of our correlated random coefficients model and find considerable heterogeneity in the returns to location choice for different groups of switchers. Importantly, we find a consistent negative relationship between comparative and absolute advantage across countries. This provides evidence that rural-urban migration is a “pro-poor” technology. The types with the lowest base returns to rural locations get the highest incremental returns when migrating to urban areas. When using the linear extrapolation of the restricted GRC model that we show is supported by the data, we find significantly higher returns for non-switchers, especially those always residing in rural areas. This suggests that there is inefficiency and misallocation of skills across space in these economies and, therefore, opportunities for policy interventions.

Our paper provides a way to reconcile divergent estimates in the literature by explicitly accounting for the role of heterogeneity in the returns to rural-urban migration. Different empirical methods rely on specific subpopulations for identification, and if some of those groups have higher returns than others, estimates are bound to differ. If average income or consumption gaps are partly due to efficient sorting—as posited by Lagakos and

Waugh (2013), Young (2013), Hamory *et al.* (2021), and others—then existing earnings gaps would be poor predictors of the returns to migration for rural residents weighing the migration decision. Observational returns—such as those recorded by Alvarez (2020) and Hamory *et al.* (2021)—apply only to those who voluntarily move (or switch labor markets). Without further assumptions, it is hard to argue that these returns should necessarily extend to those who do not voluntarily move. In fact, returns to those who stayed behind may be smaller precisely because any higher returns have already been "arbitraged away" by the early movers.

Stayers (i.e., non-switchers) could also have high expected returns if external constraints or costs keep them from migrating in the status quo—an argument made by Lagakos, Mobarak and Waugh (2023). Experimental studies, such as Bryan *et al.* (2014), estimate the returns for individuals who are induced to move by financial incentives to migrate. Again, we have little basis for extrapolating such estimates to the broader population that did not respond to the incentives.[4] While these separate estimates are all informative, each approach provides but a piece of the puzzle of the returns to internal migration.

Our work contributes to two strands of the literature. First, we reconcile diverging estimates of returns to rural-urban migration and occupational mobility from agriculture to non-agriculture sectors by explicitly accounting for heterogeneity in the returns of subpopulations defined by their history of location or occupation choices. We show that the returns for most of the switcher types are close to small and homogenous estimates found in the literature (Alvarez, 2020; Hamory *et al.*, 2021), whereas, in some specifications, the returns for non-switchers are close to large estimates found in the cross-sectional comparisons of the structural transformation literature (Lagakos and Waugh, 2013; Young, 2013; Gollin *et al.*, 2014).[5] Second, we adapt existing generalized Roy models and econometric methods previously used by studies in labor (Lemieux, 1998) and technology

[4] Experimental studies and others that rely on instrumental variables approach to estimate local effects, which are not necessarily valid for the entire subpopulation of non-switchers.

[5] Lagakos, Mobarak and Waugh (2023) also reconcile estimates, but their focus is on reconciling early experimental evidence of high returns to temporary rural-to-urban migration with later experiments that showed much smaller estimates.

adoption (Suri, 2011) to the study of heterogeneity in the returns to rural-urban migration. Our empirical strategy builds on recent work in the econometrics literature (Tjernström *et al.*, 2023; Verdier, 2020) to estimate returns to switcher subpopulations in a flexible manner while also assessing the validity of the identifying restrictions necessary to estimate returns to non-switchers.

Our model can be used in different settings that involve binary choices, self-selection, and potential heterogeneity in returns, and our empirical strategy can be used by studies that aim to reconcile estimates from analysis with cross-sectional, experimental, and longitudinal designs. In migration studies, in particular, our model can be extended to accommodate temporary and permanent costs that affect a worker's migration decisions.

2. Model and Identification

We consider an economy with two labor markets, rural (R) and urban (U), into which individuals (i) sort themselves based on their skills and preferences.[6] As is common in the literature (Bazzi, Gaduh, Rothenberg and Wong, 2016; Alvarez, 2020; Lagakos and Waugh, 2013; Pulido and Święcki, 2021), we adapt the classic Roy (1951) model to motivate our empirical strategy. In line with this literature, we allow for workers' unobservable comparative advantage to play a role in their choice to migrate, that is, to switch between labor markets (or sectors).

Different labor markets in the economy require different skills from the workers they employed in them and reward these skills differently. Workers in the economy have different endowments, summarized by a vector of market-specific skills $a_i = \{a_i^R, a_i^U\}$. These skills determine a worker's productivity when carrying out market-specific tasks. Each person in the economy receives a draw from a joint distribution of skills, $G(a^R, a^U)$, which we assume has support on the real line and finite first two central moments.

[6] Throughout this section, we motivate our model using migration (labor market) decisions. The implications of this model extend naturally to discussing sectoral choices. In this case, the binary choice dimension changes from the labor market $l = \{\text{Rural, Urban}\}$ to the employment sector $s = \{\text{Rural, Urban}\}$. Our empirical analysis investigates the returns to this alternative choice dimension as well whenever information on the sector of employment is available in the data.

Consumption, utility, and choices

We assume that the log consumption of individual i in period t is determined by her unobserved ability, a_i, her labor market choice in that period, D_{it}, her vector of observed characteristics, X_{it}, and an error term, ε_{it}. Assuming that the last two components are additive, and omitting the vector of observed characteristics for now, we can write individual log consumption, for $i = 1, \dots, N$ and $t = 1, \dots, T$ as:

$$y_{it} = f(a_i, D_{it}) + \varepsilon_{it}. \tag{1}$$

Let $D_{it} = 1$ for individuals working in the urban labor market and 0 otherwise. Then, the relationship between unobserved ability and migration choice given by f takes only two possible values. The first, $f(a_i, 0)$, is the individual-specific log consumption when choosing the rural market. The second, $f(a_i, 1)$, is her log consumption when choosing the urban market. By denoting the first value by $\mu_i \equiv f(a_i, 0)$ and their difference by $\Delta_i \equiv f(a_i, 1) - f(a_i, 0)$, we can write an individual's log consumption as an intercept-and-slope format:

$$y_{it} = \mu_i + \Delta_i D_{it} + \varepsilon_{it}. \tag{2}$$

In this scenario, and ignoring the error term ε_{it} for now, an individual chooses the urban market whenever the premium associated with such choice is positive:

$$D_{it} = 1 \Leftrightarrow (y_{it}|a_i, D_{it} = 1) - (y_{it}|a_i, D_{it} = 0) = \Delta_i > 0. \tag{3}$$

By assuming that unobserved ability is time-invariant, we assume time-homogeneity in the individual-specific premium, Δ_i. In this context, an individual would only switch markets after the first period if (*i*) there is a change in her observed characteristics that are rewarded differently in each labor market (e.g., schooling), or (*ii*) there are fluctuations in market-specific factors that affect all individuals equally (e.g., firms' productivity).[7]

We assume that individuals make their migration decisions based not on income or consumption alone but on overall utility, and this utility is subjected to idiosyncratic shocks that vary by market and time.[8] Though income and consumption surely are important determinants of an individual's utility, they're not the only ones. Individuals have preferences over non-monetary aspects such as proximity to family, local amenities, and other utility-impacting factors that vary by market and time. We allow such factors to affect one's migration decisions by introducing a simple utility shock. With this modification, individuals in our model may switch between labor markets even when their returns are time-invariant, and there are no differences in the rewards to observed characteristics or market-specific factors that affect all individuals equally.

Specifically, we let ν_{it}^{l} denote the individual i's idiosyncratic utility shock when she is in market $l = \{\text{Rural, Urban}\}$, and $V_{it}^{l} = y_{it} + \nu_{it}^{l}$ denote her utility in period t. At the end of each period, the individual realizes her consumption, observes her utility shocks, and forms expectations on the shocks for the next period. She then decides to switch to (or remain in) the urban market at the beginning of period t if her expected utility there is higher. That is, she chooses $D_{it} = 1$ if

$$\mathrm{E}\left(V_{it}^{U}\right) > \mathrm{E}(V_{it}^{R}) \Leftrightarrow \Delta_i + \mathrm{E}\left(\nu_{it}^{U} - \nu_{it}^{R}\right) > 0. \tag{4}$$

In what follows, we assume that individuals make their migration choices in each period according to the decision rule in equation (4). Implicit in this equation, is the assumption that the individual knows her ability and the returns to such ability in both markets ($\mu_i + \Delta_i$). Moreover, in a model extended with observed characteristics and (market-specific) period fixed effects, δ_t^{l}, the individual also knows the returns in each market to her characteristics ($X_{it}'(\gamma^{U} - \gamma^{R})$), and any difference in period fixed effects

[7] In the second case, the order (in terms of ability) of the individuals crossing the decision threshold would always be the same: a positive fluctuation that moves someone with Δ_i over the threshold will necessarily move all other individuals i' with $\Delta_{i'} > \Delta_i$.

[8] Like we do with the alternative dimension choice of sectoral mobility, we can also investigate returns to income in our empirical analysis. We show the results for this alternative dependent variable in the appendix.

($\delta_t^U - \delta_t^R$). Finally, equation (4) implies that, for the individual, $\mathrm{E}(\varepsilon_{it}|a_i, D_{it} = 1) - \mathrm{E}(\varepsilon_{it}|a_i, D_{it} = 0) = 0$, that is, the expected difference in idiosyncratic consumption shocks is zero and, therefore, do not affect their migration decision in period t.[9]

Heterogeneous returns for different trajectory types

Settings like ours, with a binary choice and multiple periods, allow for grouping individuals into different types based on their full history of migration choices, i.e., their trajectories of urban employment.[10] For example, with two periods, we can group individuals into four types: those never observed in the urban labor market, those who start in the rural market but join the urban in the second period, those who start in the urban market but leave for the rural in the second period, and those always observed in the urban labor market. Since an individual's trajectory in the two-period case is given by $D_i = \{D_{i1}, D_{i2}\}$ and our convention is to use $D_{it} = 1$ for urban and 0 otherwise, the four trajectory types listed above can be referred to more concisely as types {0,0}, {0,1}, {1,0} and {1,1}. In the general case with T periods, an individual's trajectory $D_i = \{D_{i1}, \dots, D_{iT}\}$ will assign her to one of 2^T possible types. The first and last of these types are non-switchers: individuals who never migrate from the rural to the urban labor market or vice-versa, those for whom D_{it} is always equal to zero or always equal to one. Everyone else, those for whom $D_{it} \neq D_{it'}$ for at least some $t \neq t'$, is a switcher.

Trajectory types represent a relevant dimension of heterogeneity in the returns to rural-urban migration. Traditional panel estimations with individual fixed effects estimate a unique average return that is a weighted average of the returns for all switcher types. The non-switchers provide no identifying variation. Such averaging hides heterogeneity. For example, the urban premium for the {0,1} type can be larger, in magnitude, than the rural penalty for the {1,0} type. For example, individuals moving from the rural to the urban

[9] The difference in expected utility shocks, on the other hand, may be different from zero and will thus affect the individual migration decision as posed in equation (4). In particular, given that the other determinant of the migration decision, the individual-specific increment, Δ_i, is time-invariant, migration beyond the first period is mostly determined by the utility shocks (the other possible drivers are observed characteristics and period effects that would move only marginal individuals).

[10] We use a broad definition of the term employment in this paper, which includes cases of individuals who engage in informal work, self-employment or home production while living in an urban location.

market may gain 20% more on average, whereas those moving in the opposite direction may lose 30% on average. In other words, the changes in consumption associated with a change of labor market for these two switcher types can have opposite directions, which is intuitive, but their magnitudes may not be symmetric.[11] More importantly, traditional panel estimations say nothing about the returns for the non-switcher types. We cannot identify the returns for non-switchers unless we assume these returns are equal to some weighted average of the returns for switchers, which is unlikely to be true in many settings.

Heterogeneity in the returns for different trajectory types does not preclude heterogeneity in other dimensions such as gender, or age group. In fact, all these different heterogeneities and their intersections can be investigated in empirical studies. Nonetheless, focusing on trajectories as the main dimension of heterogeneity in the returns to rural-urban migration has advantages. First, trajectories can be related to individual unobserved ability and their efficient allocation. For example, if sorting is efficient, individuals with the lowest returns to working in the urban market will seldom be observed in that market and the opposite will be true for those with high returns. Also, individuals moving back and forth between the rural and urban markets will be those for whom the returns to both markets are similar. On the other hand, if sorting is inefficient, individuals with high returns do not migrate and market failures such as information barriers or missing credit or insurance markets can be to blame and target with interventions. In fact, trajectory groups—especially when split into two bigger groups, switchers and non-switchers—overlap with groups of interest in public policy interventions that can specifically target marginal or infra-marginal movers.

The full trajectory of urban employment across all T periods in the data, albeit unknown to the individual at every period t except the last, is known to the econometrician who can leverage this information for identification purposes. In fact, this is what we do in our empirical strategy, following other papers in the correlated random coefficients literature like Lemieux (1998) and Suri (2011).[12]

---

[11] Pulido and Święcki (2021) show evidence of such an asymmetry in the returns to sectoral mobility in Indonesia (see their Table 4).

[12] Specifically, we assume strict exogeneity in the error term, $\mathrm{E}(\varepsilon_{it}|a_i, D_i) = 0$, which conditions its expectation on the entire trajectory of migration choices of an individual (past, present, and future). In a

We can express the expected value of log consumption, $\mathrm{E}(y_{it})$, as a combination of different group averages. Let $d \in D = \{0,1\}^T$ denote any possible realized trajectory in T periods. Since the set of all possible trajectories, D, finite, under the strict exogeneity assumption, $\mathrm{E}(\varepsilon_{it}|a_i, D_i) = 0$, we can integrate log consumption across all the 2^T possible trajectory groups in the data obtaining

$$\mathrm{E}(y_{it}|D_i = d) = \mu_d + \Delta_d D_{it}, \tag{5}$$

where $\mu_d = \mathrm{E}(\mu_i|D_i = d)$ is the group-average return to choosing the rural labor market, and $\Delta_d = \mathrm{E}(\Delta_i|D_i = d)$ is the group-average urban premium, i.e., the incremental return to choosing the urban labor market.

Equation (5) can be readily estimated in our data using a regression specification that contains a different dummy variable for each trajectory type, interactions of the urban dummy and the dummies for all switcher types, and additive control variables like observed individual characteristics and period fixed effects. This estimation identifies heterogeneous returns to choosing the urban labor market for all trajectory types except two: the non-switchers. For those who never choose urban, $d = \{0,\ldots,0\}$, we can only identify μ. And for those who always choose urban, $d = \{1,\ldots,1\}$, we can only identify a combined parameter $\kappa = \mu + \Delta$. To identify Δ for the first type, and to separately identify μ and Δ for the last, we must impose additional restrictions to our model.

Decomposing unobserved ability into absolute and comparative advantage

We assume unobserved ability can be decomposed into three parts or sets of skills: one that is equally useful (and rewarded) in both markets, τ_i, one that is more useful in the rural labor market, θ_i^R, and one that is more useful in the urban labor market, θ_i^U. The difference between the θ_i^U and θ_i^R skill sets can be interpreted as the individual's comparative advantage in the urban market, and the τ_i skill set as her absolute advantage. By assuming

model with observed characteristics, X_{it}, and period fixed effects, δ_t, the strict exogeneity assumption becomes $\mathrm{E}(\varepsilon_{it}|a_i, X_i, \delta_t, D_i) = 0$.

that unobserved ability is time-invariant, we are assuming that an individual is born with such a tripartite skill set.[13]

Next, like Lemieux (1998) and Suri (2011), we impose a specific structure to the relationship between these skill sets, redefining their relationship as linear projections:

$$\theta_i^R = b_R\left(\theta_i^U - \theta_i^R\right) + \tau_i \text{ and } \theta_i^U = b_U\left(\theta_i^U - \theta_i^R\right) + \tau_i, \tag{6}$$

where by defining $\sigma_U^2 = \mathrm{Var}\left(\theta_i^U\right)$, $\sigma_R^2 = \mathrm{Var}(\theta_i^R)$, and $\sigma_{UR} = \mathrm{Cov}\left(\theta_i^U, \theta_i^R\right)$ we have $b_U = (\sigma_U^2 - \sigma_{UR})/(\sigma_U^2 + \sigma_R^2 - 2\sigma_{UR})$, $b_R = (\sigma_{UR} - \sigma_R^2)/(\sigma_U^2 + \sigma_R^2 - 2\sigma_{UR})$.

Within this structure, the term τ_i does not vary with the individual's labor market and is orthogonal to the difference $\theta_i^U - \theta_i^R$ by construction. Next, defining $\theta_i = b_R\left(\theta_i^U - \theta_i^R\right)$ and $\phi \equiv (b_U - b_R)/b_R$ we can rewrite both market-specific skill sets as functions of the individual's comparative advantage θ_i and the parameter ϕ:

$$\theta_i^R = \theta_i + \tau_i \text{ and } \theta_i^U = (1 + \phi)\theta_i + \tau_i. \tag{7}$$

Finally, we assume that absolute advantage is mean independent of the labor market choice, $\mathrm{E}(\tau_i|D_i) = \mathrm{E}(\tau_i)$, we normalize the comparative advantage to be mean zero, $\mathrm{E}(\theta_i) = 0$, and we let the return to urban employment have a component that is common to all individuals, β.[14] With these additional assumptions and normalization (i.e., modeling restrictions), we obtain the restricted version of equation (2):

$$y_{it} = \theta_i + \tau_i + (\beta + \phi\theta_i)D_{it} + \varepsilon_{it}, \tag{8}$$

where $\mu_i = \theta_i + \tau_i$ and $\Delta_i = \beta + \phi\theta_i$.

---

[13] The assumption that individuals are endowed with a vector of market-specific skills or abilities is common in the literature on rural/urban productivity gap (or the agricultural/nonagricultural gap, in some cases). See, for example, Lagakos and Waugh (2013), Lagakos *et al.* (2020), and Pulido and Święcki (2021).

[14] These are the same assumptions made by Suri (2011), which in turn follows Lemieux (1993, 1998) and Carneiro, Hansen and Heckman (2003).

Interpreting the linear extrapolation that identifies returns to non-switchers

Equation (8) says that an individual's return to choosing the rural market, μ_i, is determined by her unobserved ability in that market (from equation (7), $\theta_i^R = \theta_i + \tau_i$), whereas her incremental return to choosing urban, Δ_i, is determined by an urban premium that is common to all individuals, β, and the parameter ϕ. This parameter directly relates to the individual's comparative advantage in the urban labor market. In fact, equation (7) implies that the ratio of an individual's market-specific skills net of her absolute advantage, i.e., her net comparative advantage in the urban market is equal to $1 + \phi$, which in turn represents a relationship between the dispersion of these skills and their correlation in the population:

$$\frac{\theta_i^U - \tau_i}{\theta_i^R - \tau_i} = 1 + \phi = \frac{\mathrm{Var}(\theta_i^U) - \mathrm{Cov}(\theta_i^U, \theta_i^R)}{\mathrm{Cov}(\theta_i^U, \theta_i^R) - \mathrm{Var}(\theta_i^R)}. \tag{9}$$

Equation (9) has two important takeaways. The first is the interpretation of comparative advantages in terms of the variance and covariance of market-specific skills in the population, as one would expect in any generalized Roy (1951) model. The second takeaway is a direct result from the linear projections used in equation (6), which yields a relationship between the individual's pair of market-specific skills that is also linear. Such a relationship is key to the identification of the returns to non-switcher types in our model, which rely on a linear extrapolation from the returns to switchers (usefully illustrated by Verdier, 2020). Tjernström *et al.* (2022) call this the Linearity in Comparative Advantage (LCA) assumption and show that for any two trajectory types $d \neq d'$ the following equality holds:[15]

$$\phi = \frac{\Delta_d - \Delta_{d'}}{\mu_d - \mu_{d'}}, \qquad \mu_d \neq \mu_{d'}. \tag{10}$$

A particularly useful way to interpret the linear extrapolation in equation (10) is to think of urban employment as a consumption-enhancing technology and the corresponding labor market choice as technology adoption. Then, a positive ϕ means the technology gives

[15] See Proposition 1 in Tjernström *et al.* (2022).

higher returns for those who start ahead (it is "pro-rich"): for any two trajectory types, the type with the highest average consumption in the rural market will also have the highest average incremental return to migrating to the urban market. Conversely, a negative ϕ means the technology is "pro-poor:" it is the type with the lowest average consumption in the rural market that will benefit the most from migrating to the urban market. Notice that a positive ϕ implies a positive correlation between the unobserved market-specific skills in the population, but the converse is not true.[16]

The restricted model in equation (8) can be estimated via GMM using the flexible Group Correlated Coefficient (GRC) approach proposed by Tjernström *et al.* (2023).[17] Under strict exogeneity of the error term and the LCA assumptions, the GRC identifies returns to non-switchers. It also provides a framework in which we can test key identifying restrictions and calculate weak-identification robust inference confidence intervals for the parameters of interest.

Figure 1 illustrates the linear extrapolation determined by equation (10) in a two-period panel, in which the restricted GRC model is exactly identified. The figure shows estimates for the average consumption level in the rural market, μ, and the urban consumption premium, Δ, for the two switcher types $d = \{0,1\}$ and $d = \{1,0\}$ on the (μ, Δ)-

[16] From equation (9), we have

$$\phi = \frac{\mathrm{Var}(\theta_i^U - \theta_i^R)}{\mathrm{Cov}(\theta_i^U, \theta_i^R) - \mathrm{Var}(\theta_i^R)} > 0 \Leftrightarrow \mathrm{Cov}(\theta_i^U, \theta_i^R) > \mathrm{Var}(\theta_i^R) \Leftrightarrow \rho > \frac{\sigma_R}{\sigma_U}, \qquad \rho = \frac{\mathrm{Cov}(\theta_i^U, \theta_i^R)}{\sigma_U \sigma_R}.$$

A positive ϕ requires that (*i*) $\rho > 0$ since $\sigma_R/\sigma_U > 0$, and (*ii*) $\sigma_U \geq \sigma_R$ since $|\rho| \leq 1$. That is, we must have a positive correlation between the two market-specific skill sets, and a larger dispersion of urban-specific skills than rural-specific skills. The condition $\phi < 0$, on the other hand, can occur with $\rho = 0, < 0$ or > 0 and $\sigma_R/\sigma_U < 1$ or > 1. In other words, observing $\phi < 0$ does not allow us to infer the direction of the correlation between market-specific skills, ρ.

[17] The GMM estimation uses the following equation (vector of covariates and period fixed effects not shown):

$$y_{it} = \sum_{d \in D \setminus \{1,\ldots,1\}} \mu_d 1\{D_i = d\} + \Delta_b D_{it} 1\{D_i = b\} + \sum_{d \in D_S \setminus b} \phi(\mu_d - \mu_b) D_{it} 1\{D_i = d\}$$
$$+ \left(\mu_{\{1,\ldots,1\}} + \phi\left(\mu_{\{1,\ldots,1\}} - \mu_b\right)\right) h_{it} 1\{h_i = \{1, \ldots, 1\}\} + \varepsilon_{it}$$

where D is the set of all trajectory types, D_S is the subset of switcher types, $\{1, \ldots, 1\}$ is the trajectory of the "always urban" type, and b is the "base" switcher trajectory. Since the base trajectory can be any among the switcher types, we chose, among those with more than five individuals, the one with the most precise Δ estimate in the unrestricted GRC estimation used to populate initial values for the GMM estimation.

plane. The vertical solid line denotes the average level in the rural market for the non-switcher type $d = \{0,0\}$, the parameter $\mu_{\{0,0\}}$, whereas the 45-degree dashed line denotes the combination of average rural level and urban premium for the non-switcher type $d = \{1,1\}$, i.e., the combined parameter $\kappa = \mu_{\{1,1\}} + \Delta_{\{1,1\}}$. All these parameters can be estimated using the unrestricted version of the GRC (equation (5)). The restricted GRC imposes a common intercept ϕ for the line that connects the (μ, Δ)-estimates for all types, both switchers and non-switchers. From the switchers, we obtain the blue extrapolation line. And from the intersection of this line with the vertical and 45-degree lines, we obtain estimates for the non-switcher types (in blue).

Figure 2 illustrates the linear extrapolation determined by equation (10) in a three-period panel, in which the restricted GRC model is overidentified (the illustration of cases with $T > 3$ periods would be similar but more "crowded"). The figure mimics the ones we use later in the paper to present results, so it is worth explaining its structure and conventions in more detail here.

The size of the gray diamonds is proportional to the number of individuals in each switcher trajectory relative to the total number of switchers in the data. Because the model is overidentified, the estimation of the slope of the extrapolation line, ϕ, considers the precision of the (μ, Δ) estimates (confidence intervals not shown) and the number of individuals in each switcher trajectory. As such, the ϕ estimate resembles a weighted average of all possible different slopes. For illustration purposes, we place the extrapolation line intercepting the (μ, Δ) estimate of the switcher trajectory with the most precise Δ estimate among those with more than five individuals in the sample.

A note on balance

The discussion of our model and the trajectory types, as well as our estimation equations, implicitly assume perfect balance: individuals are observed in all periods, leaving no gaps in their full location or employment histories. Sadly, this is not always true in the data. Many individuals are missing in one or more periods. Using their missing period to form new trajectories is impractical for several reasons, chiefly because the "always missing" type is not observed and because the number of possible trajectories would rapidly grow.[18] On the other hand, unbalanced individuals could provide useful variation to identify coefficients on controls such as demographic characteristics and period fixed effects. For this reason, we adjusted our regression specifications to use unbalanced panels.

The adjustment is simple. In all specifications, we include a dummy variable equal to one if the individual is not part of the perfectly balanced panel, that is, if the individual is missing from the data for one or more periods. We also include an interaction between the unbalanced dummy and the choice variable (the urban or the non-agriculture dummies). These two variables capture the base consumption level and the urban increment for unbalanced individuals. Only the variation of the balanced individuals is left to identify returns to the different trajectory types, including the non-switchers. In the appendix, we show results using just the perfect balanced panel data.

3. Data

We use longitudinal data from four developing countries—Indonesia, South Africa, China, and Tanzania— to understand selection and heterogeneity in the returns to migration and test the model's predictions described in the previous section. For each country, we draw from data collected in household surveys designed to collect information on the living standards of people in settings where informal employment, home production, and rural work are prevalent. Our data include detailed information on tens of thousands of individuals across multiple geographies and time periods providing rich information on their demographic characteristics, comprehensive measures of income and consumptions, place of residence (rural or urban), and sector of employment (agriculture or non-agriculture).[19]

Despite spanning multiple years, these surveys have relatively low attrition rates and represent a significant portion of the population in each country. For Indonesia, we use data

[18] If missing is an accepted status, the indicator variable assumes three values: $D_{it} \in \{0,1,\text{missing}\}$ and the number of possible trajectories goes from 2^T to 3^T.

[19] In this draft we use data on sector of employment for Indonesia and South Africa only. Future drafts will include these data also for Tanzania and China.

from all five waves of the Indonesia Family Life Survey (henceforth, IFLS), which were collected between 1993 and 2015 and is representative of about 83% of the Indonesian population. The IFLS has low attrition rates, with re-contact rates of over 90% between any two consecutive survey waves and 87% of households from the first wave were contacted in all five waves (Strauss et al., 2016).

For South Africa, we use data from five waves of the South African National Income Dynamics Study (NIDS) that were collected between 2008 and 2017. Of the individuals from the first round of the survey, 73% was re-contacted in the fifth round of the survey (Lagakos et al., 2020).[20] We use Chinese data from four waves of the China Family Panel Study (CFPS), which were collected biannually from 2010 to 2016. The survey is representative of 95% of the Chinese population, and re-contact are never below 85% in the first three waves (Lagakos et al., 2020). Finally, we use three waves of the Tanzania National Panel Survey (NPS), which were collected between 2008 and 2013. NPS also has very low attrition, with re-contact rates of 96.5% from the first to second wave and 95.2% from the first to third wave (Lagakos et al., 2020).[21] Nonetheless, we observe significant fluctuations in the number of observations in each sample over the periods included in our study and notice that only a subset of individuals in each sample are observed in all periods (perfectly balanced).[22] We acknowledge that entering or dropping out of a sample can be correlated with aging, employment, and migration trajectories and that the motivation of our model assumes individuals are observed in all periods. For that reason, we use both the full sample and a subsample of perfectly balanced observations for each country in our study. We use the full, unbalanced samples when presenting summary statistics, performing our main empirical analysis, and discussing results in the text. We then repeat it all for the balanced

[20] If individuals from a household were temporarily away or unwell, enumerators made an effort to get a proxy from that household to answer questions on behalf of the missing individual.

[21] In addition to having information on the sector of employment and current location, the IFLS also has information on an individual's birth location (if rural or urban), allowing us to perform additional exercises that support our identification assumptions. Refer to Kleemans and Magruder (2018) and Kleemans (2023) for more details on the IFLS data.

[22] Indonesia, in particular, shows a large drop in observations between the balanced and unbalanced panels. Even with good rates of follow-up, the IFLS survey loses individuals along the way because it covers a 23 year period. An individual in the perfectly balanced panel must have been 16 or older in the first survey and followed up on every subsequent survey. On top of that, she must have reported location and consumption in all survey waves, as well as employment and income if we use the sector of employment.

samples, presenting the corresponding tables and figures in the appendix and discussing them briefly in the text, pointing out any noticeable differences between the results obtained from the two types of samples.

Table 1 provides an overview of our data sources, including the years when each survey round was collected, and the number of individuals and individual-year pairs in their full and balanced samples. In each case, we keep only individuals aged 16 and above with non-missing information on urban/rural status and total consumption. When looking at the sector of employment or using total income as the dependent variable, we adjust the sample restrictions accordingly, keeping only those with non-missing information on employment and income. The majority of individuals kept in our samples have non-missing information on demographic characteristics used as controls in our empirical analysis (sex, education, and household size), so there is not much variation in the samples used in regressions. We cleaned and organized the data for Indonesia and South Africa ourselves. For China and Tanzania, we used data already organized by Lagakos, Marshall, Mobarak, Vernot, and Waugh (2020), available in their replication package.

Table 1 also highlights the proportion of individuals that switch, at least once, between the rural and urban locations and between the agriculture and non-agriculture sectors. The proportion of rural/urban switchers in the unbalanced samples varies from 7% (South Africa and China) to 16% in Indonesia, and the proportion of agriculture/non-agriculture switchers is 7% in South Africa and 11% in Indonesia. These proportions are higher in the balanced panel, ranging from 8% to 38% (rural/urban) and 16% to 29% (agriculture/non-agriculture). Because only switchers contribute to the identifying variation in panel regressions with individual fixed effects, these figures show that up to 93% of the individuals in our sample – the non-switchers – are left out of the analysis if we rely on this approach. In other words, by imposing additional restrictions in our model, we are able to estimate the returns to migration for a very substantial share of individuals in our sample, whom are particularly relevant to policy makers and in discussions about misallocation. Moreover, we are able to allow their returns to be different from the average returns for switchers. The traditional methods must either ignore these subpopulations or assume their returns are the same as the returns for switchers.

Choice dimension and outcome of interest

Our model and empirical strategies can be used to study either heterogeneity in the returns to rural-urban migration or to sectoral mobility between agriculture and non-agriculture. In other words, the dimension of the binary choice the individuals in our model and data face can be the rural/urban location or the agriculture/non-agriculture employment.[23] Similarly, regarding the outcome of interest, the dependent variable in our regressions, we can study returns in consumption or income. We focus our analysis and lead the presentation and discussion of our results, focusing on the returns in consumption to rural-urban migration.

We focus on rural/urban location as the choice dimension for two reasons. First, this dimension is strongly connected to an important and truly exogenous event for all individuals: their birth. Location of birth correlates with the location where we observe individuals in the first period of the data. In Indonesia, where the data informs the location of birth, we find that those born in urban locations are 40–70% more likely to be first observed in the urban labor market (depending on the specification used, results not shown). Birth, like the draw of unobserved market-specific skills from a bivariate distribution in our model, is akin to a random lottery. Thus, the exogeneity of birth, though not necessary for identification in our model, helps make sense of our assumptions and interpretations, namely, that individuals may be first observed in a market where their skills are not being efficiently allocated and have high potential returns to migrating. Second, we have the matter of data availability. Information on place of residence (rural or urban) is available for all countries in our data, whereas information on the sector of employment (agriculture or non-agriculture) is available for Indonesia and South Africa only. Moreover, in the South African data, information on the sector of employment is missing for a significant part of the observations, while the same is not true for rural/urban locations.

We focus on consumption as the main outcome of interest also for two reasons. First, consumption is a relevant measure of well-being and also serves as a broader measure of

[23] We assign to the urban location all individuals in the sample that report living in a city or town rather than a village and to the non-agriculture sector all those whose primary employment is in non-agriculture.

income that include sources important in developing countries like in-kind transfers and home production. We refrain from using wages as the outcome of interest because computing wages requires information on hours worked, which is less reliable for non-formal employment. In other words, using wages can introduce confounders related to labor supply (hours worked), which might not be equally well measured in the two labor markets. Second, in many datasets, information on income is missing for a significant part of the observations, while the same is not true for consumption. Nonetheless, we investigate the returns to income as well for the sub-samples for which the information is available and show the results of the corresponding analysis in tables and figures in the appendix. Consumption in our data is constructed using detailed expenditure data, including the estimated value of home production consumed. Income is the sum of earnings from formal and informal employment and self-employment.

Summary statistics

We close this section by presenting summary statistics for location, employment and switcher statuses, outcomes of interest, and demographic characteristics for all countries in our data. We show the difference in means across the rural/urban and (for Indonesia and South Africa) agriculture/non-agriculture dimensions in Tables 2 and 3, respectively.

For the countries where such information is available, we notice overlaps between location and employment shown by the share of observations in agriculture or non-agriculture across the rural and urban samples (Table 2) and the share in rural or urban locations across the agriculture and non-agriculture samples (Table 3). These overlaps are strong between urban location and non-agriculture employment but much less between rural location and agriculture. In Indonesia, for example, Table 2 shows that 89% of the observations in urban locations work in non-agriculture, but only 51% of observations in rural locations work in agriculture. Looking from a different angle in Table 3, we notice that 60% of the observations working in non-agriculture in Indonesia live in urban locations, whereas 84% of those working in agriculture live in rural locations. In South Africa, the proportion of individuals working in non-agriculture is higher across both locations, but there is no perfect overlap between location and employment either.

Next, we look at differences in switcher status. These are quite small across the rural/urban dimension except for Tanzania, where the share of switchers is twice as high in the urban subsample (16% vs. 8% in the rural subsample). Differences in switcher status across the agriculture/non-agriculture dimension are more pronounced. The agriculture subsamples have a much higher share of switchers (13p.p. more in Indonesia and 30 p.p. more in South Africa), but they are also smaller as a proportion of the total sample: the agriculture subsample represents 33% of the total number of observations in Indonesia and only 14% in South Africa (bottom row of Table 3). The rural/urban split shown at the bottom of Table 2, for that matter, is much more balanced, close to 50/50 in most cases.

The middle panel in Tables 2 and 3 shows differences in means for the two outcomes of interest (consumption and income, in logs) and the share of observations for whom income is zero or missing. Across the board, we observe, not surprisingly, that both consumption and income are higher in urban than rural locations and in non-agriculture than non-agriculture sectors. Also, the proportion of observations with zero or missing income is smaller. These numbers are in line with the notion that individuals in urban locations are more likely to report positive income and to have higher levels of income and consumption than those in rural locations. The same goes for the sector of employment. Non-agriculture is associated with higher income and consumption.

Finally, the bottom panel in both tables shows differences in demographic characteristics. Most differences are statistically significant (p-values shown in brackets) but not always substantial. Gender differences, for example, are quite small. So are the differences in household size. We do notice that the urban subsamples in each country tend to be younger and more educated. Table 2 shows that urban observations have around two years of education more than the rural ones and are 1–3 years younger in Indonesia and Tanzania. In Table 3, we see that differences in age and education across sectors of employment are larger. The non-agriculture subsample in Indonesia is 3.5 years younger and has 5.5 more years of education. In South Africa, non-agriculture observations are three years younger and have 3.6 additional years of education on average.

In the appendix, we show a table of differences in means by location of birth (if rural or urban) for Indonesia only (Table A1). Of particular note on that table is the fact that, for the subsample born in an urban location, the share of observations currently living in an urban location is 47p.p. higher, and the share employed in non-agriculture is 29p.p. higher on average. These numbers corroborate our argument made earlier in this section that current urban location is strongly correlated with the "lottery" of birth. We also show analogous versions of Tables 2 and 3 in the appendix using the balanced samples of each country (Tables A2 and A3). The takeaways one can take from those tables are similar to the ones discussed here: non-perfect overlap between location and sector of employment and higher income, consumption, and education levels in urban (non-agriculture) sub-samples on average.

4. Results

OLS estimations with homogenous returns

We first show results of OLS regressions that show average consumption gaps across locations and how these change as we progressively add controls and change samples. We show results of the returns to urban location for all countries, and in the Appendix we investigate returns to non-agriculture employment for Indonesia and South Africa. In the appendix, we also repeat results uses a balanced panel and using log income instead of log consumption as the outcome of interest.[24]

The first specification we use includes only an urban dummy. Its coefficient is the difference in average log consumption of individuals living in urban or rural locations (and thus working in urban or rural markets) and reflects the raw productivity gap that motivates much of the literature on rural-urban migration and structural transformation. In the second specification, we include controls for individual characteristics and period fixed effects. Beyond being usual practice, including these covariates is justified by the differences in age education and education between the rural and urban subsamples seen in our summary statistics. Similarly, the fact that our datasets comprise multiple periods and cover a large span of time justifies the inclusion of period fixed effects. The coefficient of interest from this specification reflects the average difference in log consumption

[24] The sample is smaller when studying income instead of consumption gaps, because the former is conditional on working.

controlling for individual characteristics and shocks that affect all individuals equally between each survey period in each country.

The third specification adds individual fixed effects to our set of regression controls and shows the average rural-urban log consumption gap controlling for any time-invariant individual characteristics, including unobserved ability. Unfortunately, this specification drops all individuals observed only once in the data from the analysis. More importantly, it identifies the coefficient of interest only out of the variation between switchers, the individuals who switch between rural and urban locations at least once in the data.

As pointed out in our Data section, the proportion of switchers in our samples can be rather small, ranging from 7% to 16%, depending on the country and choice dimension used. That is, between 84% and 93% of the individuals in the sample do not contribute to the identification of the coefficient on the average urban premium when we add individual fixed effects. We attempt to tease out how much of the change in coefficients resulting from including individual fixed effects is due to the incidental sample selection and restriction of the identifying variation and how much is due to actually controlling for time-invariant characteristics. We do so by repeating all regressions for a subsample that includes only switchers.

The results for OLS regressions of log consumption on an urban indicator and various controls are shown in Table 4. Columns one to three show the results for the full sample, whereas columns four to six show the results for the subsample restricted to switchers. Specifications are consistent across panels in which we show results for all four countries in our data.[25] Each panel reports the number of observations and the number of unique individuals, which corresponds to the number of clusters used to calculate standard errors. For conciseness, we do not show coefficients on covariates, period fixed effects, or the intercept, nor do we show the R^2 for each regression.

Table 4 shows that the raw rural-urban consumption gaps for all countries in our data are large: 39 log points [lp] in Indonesia, 56lp in South Africa, 44lp in China, and 65lp

[25] Many observations in our China and Tanzania datasets miss information on one or more covariates. This makes the number of observations in their corresponding regressions drop substantially when covariates are included in the second and third specifications.

in Tanzania. This corresponds to average differences ranging from 48% to 93%.[26] We see a drop in the magnitude of all coefficients in column two, which shows the results of specifications controlling for individual characteristics and period fixed effects, and a more substantial drop in column three when individual fixed effects enter the specifications. Coefficients in column three range from 5lp in Indonesia to 16lp in South Africa. Such a drop is in line with the evidence in the literature and with the notion that unobserved ability plays an important role in the selection and returns to living in an urban location.

Columns four to six show results for regressions in which the same specifications used for the full sample were used for the sample restricted to switchers. A comparison of the results for both samples suggests that much of the drop in coefficients observed with the inclusion of individual fixed effects comes from incidental restrictions in the sample and identifying variation. The reduction in coefficients from columns five to six, where the sample and identifying variation are held constant because we have only switchers in the sample, is much smaller than the reduction from columns two to three. This observation corroborates the notion that switchers and non-switchers may be inherently different. In particular, they may have substantially different returns to migrating to an urban market.

Table 5 shows the results for OLS regressions of log consumption on an indicator of non-agriculture employment, and various controls are shown in Table 4. The structure used is the same as the previous one; only the number of countries is different since information on the sector of employment is available only for Indonesia and South Africa.[27]

The takeaways from the results in Table 5 are similar to those in Table 4. Estimates of the raw consumption gaps between the agriculture and non-agriculture sectors are large: 46lp in Indonesia and 59lp in South Africa. And the magnitude of these coefficients drops substantially as we move across specifications as before. When both controls and individual fixed effects are included, the magnitude drop to 7lp in Indonesia. In South Africa, it is not

[26] Most coefficients in Table 4 and other results tables in our paper are large in magnitude and statistically significant at the 1% level or lower. Therefore, we follow Pulido and Święcki (2021) and discuss coefficients using log points without mentioning their statistical significance and only occasionally calculating their corresponding percentage differences.

[27] Throughout this section, we first show results for regressions that use urban location as the choice dimension and then non-agriculture employment. We use the same organization and visual structure when presenting results for both choices.

statistically different than zero. Also, as before, the comparison of estimates across the full and switchers samples suggest that most of this drop comes from incidental restrictions in the sample and identifying variation caused by the inclusion of individual fixed effects rather than the fact that these fixed effects control for time-invariant unobserved ability.

In the appendix, we show similar versions of Tables 4 and 5 for a perfectly balanced panel (Tables A4 and A5) and with log income as the dependent variable (Tables A6 and A7). The pattern of the results and the takeaways are very similar to the ones discussed here for results using unbalanced panels and log consumption as the dependent variable.

Heterogeneity in the returns for switchers

Following the same country and choice dimension order, we show the results from the unrestricted version of the GRC model. We estimate regression equations analogous to equation (5) in the Model and Identification section. In the most basic version, we use a fully saturated specification in which we include dummy variables for each trajectory type and for the group of individuals not in the perfectly balanced panel (the unbalanced individuals). We also include the interaction between these group dummies and the binary choice variable (urban location or non-agriculture employment). By adding a constant to the regression, we normalize results such that the base consumption level for the first trajectory type—the "never" type, whose full location or employment history is $\{0, \ldots, 0\}$—is zero. We then expand the basic specification by adding controls. The results shown in this subsection come from specifications with observable individual characteristics and period fixed effects as controls (analogous to the specification in column two in Tables 4 and 5).

In Figure 3, we show first, in Panel A, the estimates of $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$, the base consumption level in rural locations. Then, in Panel B, we show the estimates of $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$, the increment in consumption received over the base level when an individual switches to urban locations. Estimates of the non-switcher trajectories ($\mu_{\{0,\ldots,0\}}$ and $\kappa = \mu_{\{1,\ldots,1\}} + \Delta_{\{1,\ldots,1\}}$), the group of unbalanced individuals, and controls are not shown. To improve visualization, we also omit from the figures the estimates of trajectories whose number of individuals is too small and therefore produce too-wide confidence intervals. The switcher trajectories selected to appear in each figure are those whose number of individuals is at least 0.5% of all balanced individuals. In Figure 4, we show the same set of estimates using the sector of employment as the choice dimension. As with OLS regressions and every other results table and figure in this section, we show similar versions in the appendix using a perfectly balanced panel (Figures A1 and A2) and using log income as the dependent variable (Figures A3 and A4).

The purpose of this exercise is twofold. First, it helps us inspect the degree of heterogeneity in the base consumption levels (the μ estimates). From equation (10), we know that the difference between these estimates for any two given switcher trajectories enters is the denominator of the expression (or moment) that estimates the extrapolation slope, ϕ. Being a denominator, this difference cannot be equal to zero. Though an overidentified GMM estimation can handle a few cases of equal μs, we cannot have all of them equal or even too close to each other, lest our estimation will have weak identification problems.[28] To complement visual inspection, we test statistically whether all estimates are equal and report the corresponding F-statistics and p-values.

Figures 3 and 4 show that, across countries and choices, μ estimates are indeed not all equal to each other. Many trajectories have μs that are similar in magnitude, but not all. Visually, we note at least a few very different estimates (some on opposite sides of the zero line). More importantly, the statistics reject the hypothesis that all estimates are equal with a high degree of certainty.

The second purpose of this empirical exercise is to provide evidence of heterogeneity in the returns to urban location or non-agriculture employment across the different switcher trajectories (the Δ estimates). Though not as important as the difference between μs is for identification, the difference between Δs supports our claim that choice histories, or trajectories, represent a relevant dimension of heterogeneity. Moreover, the difference in Δs is the numerator of the expression for ϕ in equation (10). Thus, no difference means a flat slope that, when extrapolated to non-switchers, will assign them the average return observed for all switchers. That is, no heterogeneity in the returns for switchers implies no heterogeneity in the returns for non-switchers too. This would be the

[28] In fact, in a few cases, our GMM procedure does not achieve convergence. Likely, because too many $of\ the\ \mu$ estimates are very similar in magnitude. We do not show the results for such cases in our tables and figures.

case implicitly assumed in panel specifications with individual fixed effects. If that is the case, there is no gain in imposing additional restrictions in our model to identify returns to non-switchers.

Again, the results in Figures 3 and 4 support heterogeneity. Estimates of the returns to urban location and non-agriculture employment show they are often very different from each other (again, sometimes on opposite sides of the zero line). Heterogeneity in returns is more pronounced in the returns to non-agriculture employment than in the returns to urban location. Also, again, the visual evidence is corroborated by the statistical tests: the hypothesis that all Δ estimates are equal is rejected in all cases with $p < 0.01$. differences for all countries.

Returns for non-switchers

Encouraged by the evidence of heterogeneity in the returns to urban location and non-agriculture employment obtained from plotting and testing results from the unrestricted version of the GRC model, we now turn to the results from its restricted version. We estimate, via GMM, regression equations analogous to equation (8) in the Model and Identification section. Again, we adjust the regression equation to include the group of unbalanced individuals and expand it, adding individual controls and period fixed effects. We show results from two specifications, one without controls and one that controls for observable individual characteristics and period fixed effects (thee are analogous to specifications in columns one and two in Tables 4 and 5, respectively).

Tables 6 and 7 show the results for the restricted GRC estimations for all countries using the location of residence and sector of employment as the choice variable, respectively. For brevity, we show only the Δ estimates for the two non-switcher types and for the slope of the extrapolation line, ϕ. The "Never" type corresponds to the $\{0, \dots, 0\}$ trajectory (never in urban or never in non-agriculture), and the "Always" type to the $\{1, \dots, 1\}$ trajectory (always in urban or always in non-agriculture). The estimates of μ and Δ for switcher types, $\mu_{\{0,\dots,0\}}$, and $\mu_{\{1,\dots,1\}}$ are not shown, nor are those covariates and fixed effects. We report J-statistics and p-values from the overidentification test at the bottom of the tables. In the appendix, we show analogous versions of these tables using a perfectly balanced panel (Tables A8 and A9) and using log income as the dependent variable (Figures A10 and A11).

We focus our discussion on the results from specifications with the full set of controls (the second column under each country). Results are quite varied. In Table 6, for example, we observe mostly negative estimates of the extrapolation slopes, ϕ (the exception is South Africa, column four). This means rural-urban migration is a "pro-poor" technology in our setting. Those trajectories with the lowest base consumption level in rural locations are the ones that can benefit the most from moving to an urban location. In Table 7, the ϕ estimates suggest that non-agriculture employment is a pro-poor technology in Indonesia but not in South Africa. In South Africa, both urban location and non-agriculture employment seem to be "pro-rich." Those ahead in rural areas or in the agriculture sector are the ones with higher returns to switching.

In some cases, the magnitude of the extrapolation slope is quite small, yielding an extrapolation line that is mostly flat and, therefore, identifies returns for non-switchers that are close to the average return for the switcher groups. This "average return" is roughly equivalent to the estimate we get from OLS regressions with individual fixed effects, so it is useful to compare the estimates of returns for non-switchers in Table 6 and the estimates in column three of Table 4, Panels A to D (we do a similar comparison between the Δ estimates in Table 7 and the coefficient on the Non-Agriculture dummy in column three of Table 5, Panels A and B).

In Indonesia, for example, the ϕ estimate in column two is –0.08 and not statistically different from zero. Accordingly, the estimated returns for the Never and Always types (8lp and 7lp) are only slightly larger than the average return estimated for all switcher individuals in column three in Panel A of Table 4 (5lp). In Tanzania, on the other hand, the ϕ estimate is –0.84 (column eight), and the return to urban location for the Never type is 33lp, three times the average return for all switcher individuals (11lp, in column three in Panel D of Table 4). The point estimate of the return for the Always type in Tanzania is even larger in magnitude (–150lp), but very imprecise and thus not statistically significant. We verify a similar pattern for the returns to non-agriculture in Indonesia. The extrapolation slope is negative (–0.97), and the Δ estimate for the Never type (20lp) is much higher than the average return in column three, Panel A in Table 5 (7lp), while the Δ estimate for the Always type is not statistically significant.

The overidentification tests in Table 6 reject the null hypothesis (all moment restrictions are equal) in all cases except for Indonesia. In Table 7, however, overidentification is not rejected at the 10% level in the specifications shown in columns two and four. This suggests that switcher trajectories are "misaligned," generating several possible different slopes. To verify this possibility, we plot the estimates from the restricted and unrestricted versions of the GRC estimates in the same fashion as Figures 1 and 2 in the Model and Identification section. The results from this exercise are shown in Figures 5 and 6 below (Figures A5 to A8 are the corresponding appendix). These figures shed light on the heterogeneity in the returns to switchers and non-switchers, provide a clear visualization of the extrapolation line generated by the data, and allow us to visually assess the validity of the overidentifying restrictions in our model.

As with the previous figures, we only show results from the regression specifications with covariates and period fixed effects. The gray diamonds represent point estimates from the unrestricted GRC: the μs and Δs for all switcher types, μ for the non-switcher type $\{0, \ldots, 0\}$, and $\kappa = \mu + \Delta$ for the non-switcher type $\{1, \ldots, 1\}$. Their sizes are proportional to the number of individuals in each trajectory relative to the total number of switchers. In blue, we show the slope of the extrapolation line, ϕ, and the estimates identified by it in the restricted GRC: Δ for the non-switcher type $\{0, \ldots, 0\}$ (hollow circle), and μ and Δ for the non-switcher type $\{1, \ldots, 1\}$ (filled circle). Circle sizes are fixed and have no connection to the number of individuals in the non-switching trajectories. Confidence intervals are not shown. We place the extrapolation line intercepting the (μ, Δ) estimate of the switcher trajectory with the most precise Δ estimate among those with more than five individuals in the sample. To improve visualization, we omit the (μ, Δ) estimate for the non-switcher type $\{1, \ldots, 1\}$ when it would extend the scale of the graph considerably (that is the case for Tanzania in Figure 5 and Indonesia in Figure 6).

Discussion of results

Overall, our results support the argument that choice histories or trajectories represent a relevant dimension of heterogeneity. First, OLS results show that a substantial part of the difference in coefficients from regressions with and without individual fixed effects is due to the inherent changes in the sample and identifying variation caused by the inclusion of these fixed effects. Putting it another way, the subsample composed solely of switchers produces regression coefficients that differ significantly from the regression coefficients produced by the full sample suggesting switchers and non-switchers have different returns to rural-urban migration and sectoral mobility. Second, results from unrestricted GRC estimations show that neither the base level consumption levels (μs) nor the urban or non-agriculture premia (Δs) of the switcher types are all statistically equal. At least some trajectories have very different estimates, sometimes in opposite directions. Third, results from restricted GRC estimations shown in tables and figures show that the extrapolation line determined by the differences in μs and Δs of switcher types has a slope that is significantly different from zero and extrapolates estimated returns for non-switchers that can be up to three times larger than the average return estimate for switchers.

Our results also show that urban location and non-agriculture employment can be seen as pro-poor technologies in some contexts and as pro-rich in others (the ϕ estimate can be negative or positive). In the first set of cases, individuals in rural locations and those employed in agriculture have lower consumption, on average, than the individuals in urban locations and those employed in non-agriculture. Yet, they have the most to gain from moving to an urban location or switching to the non-agriculture sector. This can be particularly true for the group of never-movers. Our results show that, in many settings, those never observed in urban locations or the non-agriculture sector have a large potential gain from moving. So why do they not move? Market failures such as information barriers or missing credit or insurance markets are likely causes. In such cases, sorting is inefficient: individuals are misallocating their skills across labor markets and sectors in the economy. There is space for policy interventions and potentially large welfare gains. In a smaller set of cases, we observe positive ϕ estimates (urban location and non-agriculture employment are pro-rich technologies), and individuals always observed in urban locations or in non-agriculture are the ones who, indeed, had the most to benefit from such location and employment choices. This can represent a scenario where sorting based on unobserved skills produced an efficient outcome. Perhaps not surprisingly, this is observed more often in our data in South Africa, a country where large shares of the population already live in urban locations and work in non-agriculture.

5. Conclusion

Existing estimates of productivity gaps suggest that rural-urban migration can increase labor productivity, income, and welfare in developing countries. However, these estimates say little about the potential returns for groups of individuals who are particularly relevant for the design and evaluation of possible policy interventions: the non-switchers, the individuals who are always (or never) observed in rural areas or the agricultural sector and for whom returns may be considerably different from those of switchers.

We model self-selection into rural-urban migration in a multi-period Roy model that incorporates worker heterogeneity in both absolute and comparative advantage. We then estimate a correlated random coefficient model that considers both types of heterogeneity. We also draw on recent developments in the literature on non-parametric panel data identification and employ a group-random coefficient estimator that explicitly tests the parametric assumptions that identify the returns to non-switchers.

Using rich longitudinal data from four developing countries (Indonesia, South Africa, China, and Tanzania) in which we observe location choices for tens of thousands of individuals over multiple periods, we estimate returns to urban location and non-agriculture employment for switchers and non-switchers. We show results from OLS regressions with and without individual fixed effects and from the restricted and unrestricted versions of the GRC model. Finally, we combine these results in graphs that highlight heterogeneity in returns for switcher types and illustrate how the extrapolation line in our model identifies returns for non-switchers.

We find considerable heterogeneity in returns to rural-urban migration, including differences between the returns for switchers and non-switchers and between the different trajectory types within these two groups. Importantly, our results suggest that rural-urban migration can be seen as a "pro-poor" technology that disproportionately benefits those with low returns to rural locations . Furthermore, we find significantly higher returns for non-switchers, especially those in rural areas. This is indicative of inefficiency and misallocation of skills across locations. Such cases represent opportunities for policy interventions that encourage and facilitate rural-urban migration, especially for rural populations.

References

Alvarez, J. A. (2020). The agricultural wage gap: Evidence from Brazilian micro-data. *American Economic Journal: Macroeconomics*, 12(1), 153–73.

Bazzi, S., Gaduh, A., Rothenberg, A. D. and Wong, M. (2016). Skill transferability, migration, and development: Evidence from population resettlement in Indonesia. *American Economic Review*, 106(9), 2658–98.

Bryan, G., Chowdhury, S. and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5), 1671–1748.

Carneiro, P., Hansen, K. T. and Heckman, J. J. (2003). Lawrence R. Klein Lecture: Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44(2), 361–422.

Gollin, D., Lagakos, D. and Waugh, M. E. (2014). Agricultural productivity differences across countries. *American Economic Review*, 104(5), 165–70.

Hamory, J., Kleemans, M., Li, N. Y. and Miguel, E. (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, 19(3), 1522–55.

Imbens, G. W. (2007). Nonadditive models with endogenous regressors. In R. Blundell, W. Newey and T. Persson (eds.), *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, *Econometric Society Monographs*, vol. 3, Cambridge University Press, p. 17–46.

Kleemans, M. (2022). Migration choice under risk and liquidity constraints. *Working Paper*.

— and Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613), 2032– 2065.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives*, 34(3), 174–92.

—, Marshall, S., Mobarak, A. M., Vernot, C. and Waugh, M. E. (2020). Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics*, 113, 138–154.

—, Mobarak, A. M. and Waugh, M. E. (2023). The welfare effects of encouraging rural-urban migration. *Econometrica*, 91(3), 803-837.

— and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2), 948–80.

Lemieux, T. (1993). Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection. *Working paper*.

— (1998). Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection. *Journal of Labor Economics*, 16(2), 261–291.

Pulido, J. and Święcki, T. (2021). Barriers to mobility or sorting? Sources and aggregate implications of income gaps across sectors in Indonesia. *Working paper*.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135–146.

Strauss, J., Witoelar, F. and Sikoki, B. (2016). *The fifth wave of the Indonesia Family Life Survey: Overview and field report*, Vol. 1. RAND Santa Monica, CA.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica*, 79(1), 159–209.

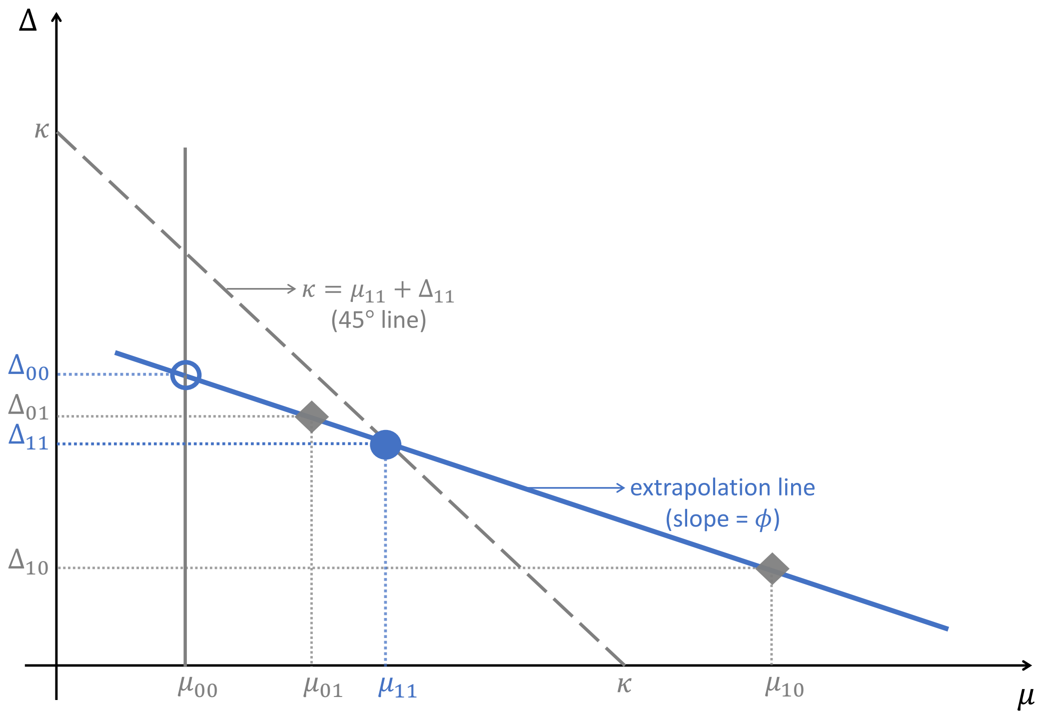
Tjernström, E., Ghanem, D., Cabanillas, O. B., Lybbert, T., Michler, J. and Michuda, A. (2022). A group random coefficient approach to modeling heterogeneity in technology adoption. *Working paper*.

Verdier, V. (2020). Average treatment effects for stayers with correlated random coefficient models of panel data. *Journal of Applied Econometrics*, 35(7), 917–939.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics*, 128(4), 1727–1785, publisher: MIT Press.

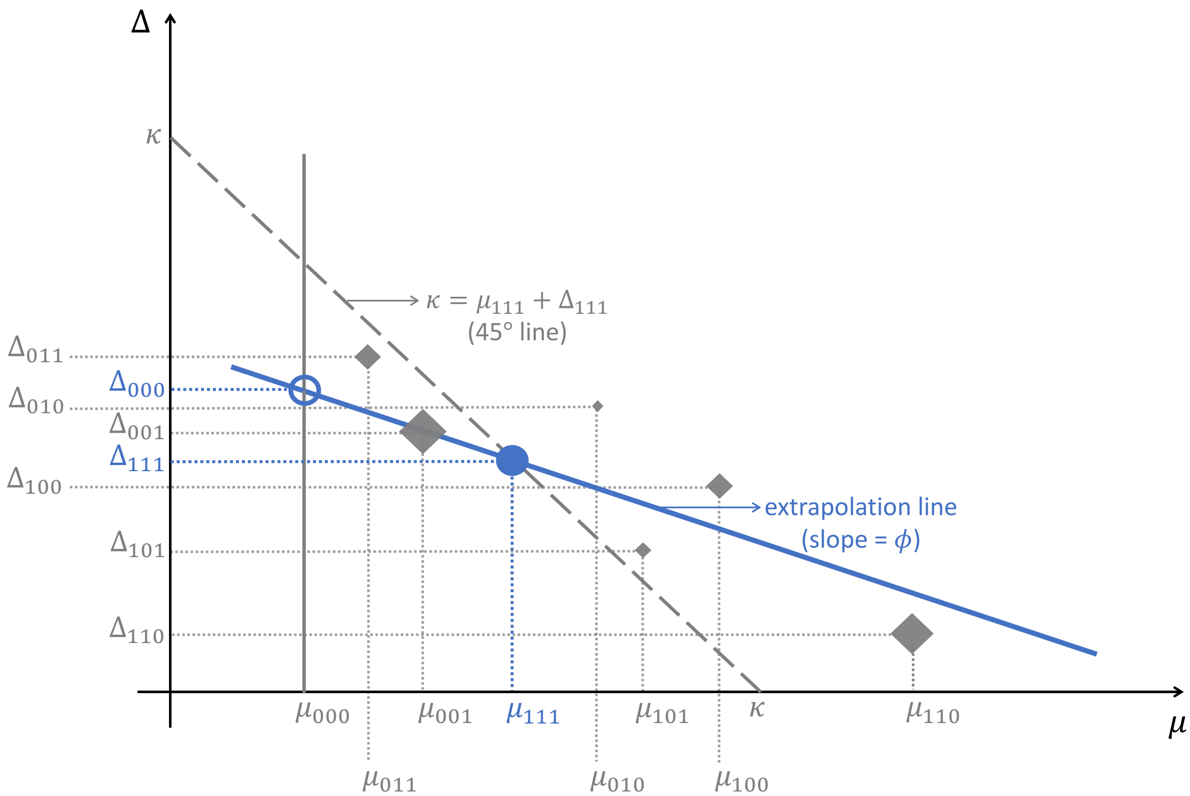
Tables and Figures

Figure 1: The extrapolation line with two periods (exactly identified)



Notes: The parameter $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$ in the x-axis denotes the base consumption level in the rural labor market, and the parameter $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$ in the y-axis denotes the increment in consumption received in the urban market over the base level. The subscripts in each estimate denote the type $d \in \{0,1\}^2$. For conciseness, we omit curly brackets and commas in the subscripts using 00 instead of {0,0}, 01 instead of {0,1}, and so on. In gray, we show estimates from the unrestricted version of the GRC: the μs and Δs for the two switcher types {0,1} and {1,0}, μ for the non-switcher type {0,0}, and $\kappa = \mu + \Delta$ for the non-switcher type {1,1}. And in blue, we show the extrapolation line and estimates identified by it in the restricted version of the GRC: Δ for the non-switcher type {0,0}, and μ and Δ for the non-switcher type {1,1}.

Figure 2: The extrapolation line with three periods (overidentified)



Notes: The parameter $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$ in the x-axis denotes the base consumption level in the rural labor market, and the parameter $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$ in the y-axis denotes the increment in consumption received in the urban market over the base level. The subscripts in each estimate denote the type $d \in \{0,1\}^3$. For conciseness, we omit curly brackets and commas in the subscripts using 000 instead of {0,0,0}, 001 instead of {0,0,1}, and so on. In gray, we show estimates from the unrestricted version of the GRC: the μ s and Δ s for the six switcher types {0,0,1}, {0,1,0}, {0,1,1}, {1,0,0}, {1,0,1} and {1,1,0}, μ for the non-switcher type {0,0,0}, and $\kappa = \mu + \Delta$ for the non-switcher type {1,1,1}. And in blue, we show the extrapolation line and estimates identified by it in the restricted version of the GRC: Δ for the non-switcher type {0,0,0}, and μ and Δ for the non-switcher type {1,1,1}.

Table 1: Overview of data sources and samples

Country	Indonesia	South Africa	China	Tanzania
Data source	Indonesia Family Life Survey	National Income Dynamics Study	China Family Panel Study	National Panel Survey
Number of waves	5	5	4	3
Years included	1993, 1997/98, 2000, 2007/08, 2014/15	2008, 2010/11, 2012, 2014/15, 2017	2010, 2012, 2014, 2016	2008/09, 2010/11, 2012/13
Full sample				
Observations	77,744	104,008	129,466	34,527
Individuals	34,399	38,427	49,398	15,667
Switchers, Urban	16%	7%	7%	8%
Switchers, Non-Ag.	11%	7%		
Balanced panel				
Observations	9,430	43,575	57,876	23,526
Individuals	1,886	8,715	14,469	7,842
Switchers, Urban	38%	13%	8%	15%
Switchers, Non-Ag.	29%	16%		

Table 2: Summary statistics and differences by Rural/Urban location

	Indonesia			South Africa			China			Tanzania		
	Urban	Rural	Difference	Urban	Rural	Difference	Urban	Rural	Difference	Urban	Rural	Difference
Urban Switcher	0.24	0.21	0.03	0.10	0.12	-0.02	0.09	0.08	0.00	0.16	0.08	0.08
	(0.43)	(0.41)	[0.00]	(0.30)	(0.32)	[0.00]	(0.28)	(0.28)	[0.01]	(0.37)	(0.27)	[0.00]
Non-Agricultural	0.89	0.49	0.39	0.94	0.74	0.21						
	(0.32)	(0.50)	[0.00]	(0.23)	(0.44)	[0.00]						
Log Consumption	12.29	11.90	0.39	8.25	7.69	0.56	10.71	10.27	0.44	15.35	14.69	0.66
	(0.77)	(0.77)	[0.00]	(1.02)	(0.80)	[0.00]	(0.90)	(0.91)	[0.00]	(0.79)	(0.74)	[0.00]
Log Income	15.15	14.67	0.48	7.94	7.40	0.54	9.26	8.36	0.90	14.41	13.22	1.19
	(1.08)	(1.14)	[0.00]	(1.19)	(1.10)	[0.00]	(1.63)	(2.06)	[0.00]	(1.85)	(1.80)	[0.00]
No Income	0.09	0.19	-0.10	0.55	0.74	-0.19	0.49	0.59	-0.10	0.50	0.66	-0.16
	(0.29)	(0.39)	[0.00]	(0.50)	(0.44)	[0.00]	(0.50)	(0.49)	[0.00]	(0.50)	(0.47)	[0.00]
Female	0.44	0.44	-0.00	0.56	0.59	-0.04	0.51	0.50	0.01	0.54	0.52	0.02
	(0.50)	(0.50)	[0.72]	(0.50)	(0.49)	[0.00]	(0.50)	(0.50)	[0.00]	(0.50)	(0.50)	[0.01]
Age (years)	38.00	39.35	-1.35	37.80	38.23	-0.43	45.67	45.44	0.23	33.52	36.30	-2.77
	(13.09)	(14.35)	[0.00]	(16.50)	(18.63)	[0.00]	(17.09)	(17.06)	[0.02]	(15.18)	(17.76)	[0.00]
Education (years)	9.32	6.59	2.72	9.58	7.70	1.88	8.80	6.14	2.66	8.53	6.99	1.53
	(4.37)	(4.38)	[0.00]	(3.62)	(4.33)	[0.00]	(4.68)	(4.54)	[0.00]	(2.83)	(2.33)	[0.00]
Household Size	4.91	4.70	0.21	4.63	6.18	-1.55	3.95	4.63	-0.68	5.85	6.77	-0.92
	(2.35)	(2.03)	[0.00]	(2.81)	(3.81)	[0.00]	(1.78)	(2.00)	[0.00]	(3.20)	(4.48)	[0.00]
Observations	35,132	42,612	77,744	53,110	50,898	104,008	59,640	69,826	129,466	12,314	22,213	34,527
Share	45%	55%		51%	49%		46%	54%		36%	64%	

Notes: Standard errors are in parentheses, and the p-value of the difference in means is in square brackets.

Table 3: Summary statistics and differences by Agriculture/Non-Agriculture employment

	Indonesia			South Africa		
	Non-Ag.	Agric.	Difference	Non-Ag.	Agric.	Difference
Non-Ag. Switcher	0.12	0.26	-0.13	0.07	0.37	-0.30
	(0.33)	(0.44)	[0.00]	(0.26)	(0.48)	[0.00]
Urban	0.60	0.16	0.44	0.68	0.26	0.41
	(0.49)	(0.36)	[0.00]	(0.47)	(0.44)	[0.00]
Log Consumption	12.23	11.77	0.46	8.27	7.68	0.59
	(0.78)	(0.74)	[0.00]	(1.02)	(0.85)	[0.00]
Log Income	15.09	14.42	0.67	7.77	7.33	0.43
	(1.08)	(1.15)	[0.00]	(1.22)	(0.85)	[0.00]
No Income	0.09	0.27	-0.19	0.01	0.27	-0.27
	(0.28)	(0.44)	[0.00]	(0.09)	(0.45)	[0.00]
Female	0.45	0.42	0.03	0.51	0.43	0.08
	(0.50)	(0.49)	[0.00]	(0.50)	(0.50)	[0.00]
Age (years)	36.93	42.45	-5.52	37.93	40.94	-3.01
	(12.63)	(15.29)	[0.00]	(11.80)	(15.17)	[0.00]
Education (years)	9.02	5.38	3.64	10.21	6.61	3.60
	(4.41)	(3.91)	[0.00]	(3.57)	(4.20)	[0.00]
Household Size	4.84	4.70	0.14	4.49	4.97	-0.48
	(2.25)	(2.05)	[0.00]	(3.07)	(3.46)	[0.00]
Observations	52,229	25,515	77,744	30,049	4,709	34,758
Share	67%	33%		86%	14%	

Notes: Standard errors are in parentheses, and the p-value of the difference in means is in square brackets

Table 4: OLS estimates of the returns to Urban location on Log Consumption

Dependent Variable:	Log Consumption					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Urban	0.392	0.214	0.054	0.330	0.108	0.065
	(0.007)***	(0.006)***	(0.009)***	(0.011)***	(0.010)***	(0.009)***
Observations	77,744	77,744	77,744	17,343	17,343	17,343
Individuals	34,399	34,399	34,399	5,435	5,435	5,435
	Panel B: South Africa					
Urban	0.564	0.445	0.162	0.271	0.174	0.160
	(0.008)***	(0.007)***	(0.014)***	(0.017)***	(0.017)***	(0.016)***
Observations	104,008	103,568	103,593	11,217	11,183	11,183
Individuals	38,427	38,261	38,286	2,807	2,807	2,807
	Panel C: China					
Urban	0.440	0.360	0.144	0.390	0.079	0.096
	(0.007)***	(0.007)***	(0.015)***	(0.016)***	(0.019)***	(0.018)***
Observations	129,466	113,543	113,543	10,939	9,379	9,379
Individuals	49,398	40,340	40,340	3,559	3,047	3,047
	Panel D: Tanzania					
Urban	0.655	0.592	0.110	0.135	0.124	0.118
	(0.012)***	(0.010)***	(0.016)***	(0.019)***	(0.017)***	(0.016)***
Observations	34,527	26,095	26,095	3,778	3,062	3,062
Individuals	15,667	12,685	12,685	1,320	1,172	1,172
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

Notes: The dependent variable is the log of total consumption. Urban is an indicator equal to one when the individual reports living in a city or town rather than a village. The Switchers sample is restricted to individuals who switch between rural/urban locations at least once. Covariates: female, age (years), age squared, education (years of schooling), education squared, and household size (members). Period refers to the survey round or wave. Coefficients on covariates and period fixed effects not shown. Robust standard errors clustered at the individual level are in parentheses. Stars denote: * p<0.10; ** p<0.05; *** p<0.01.

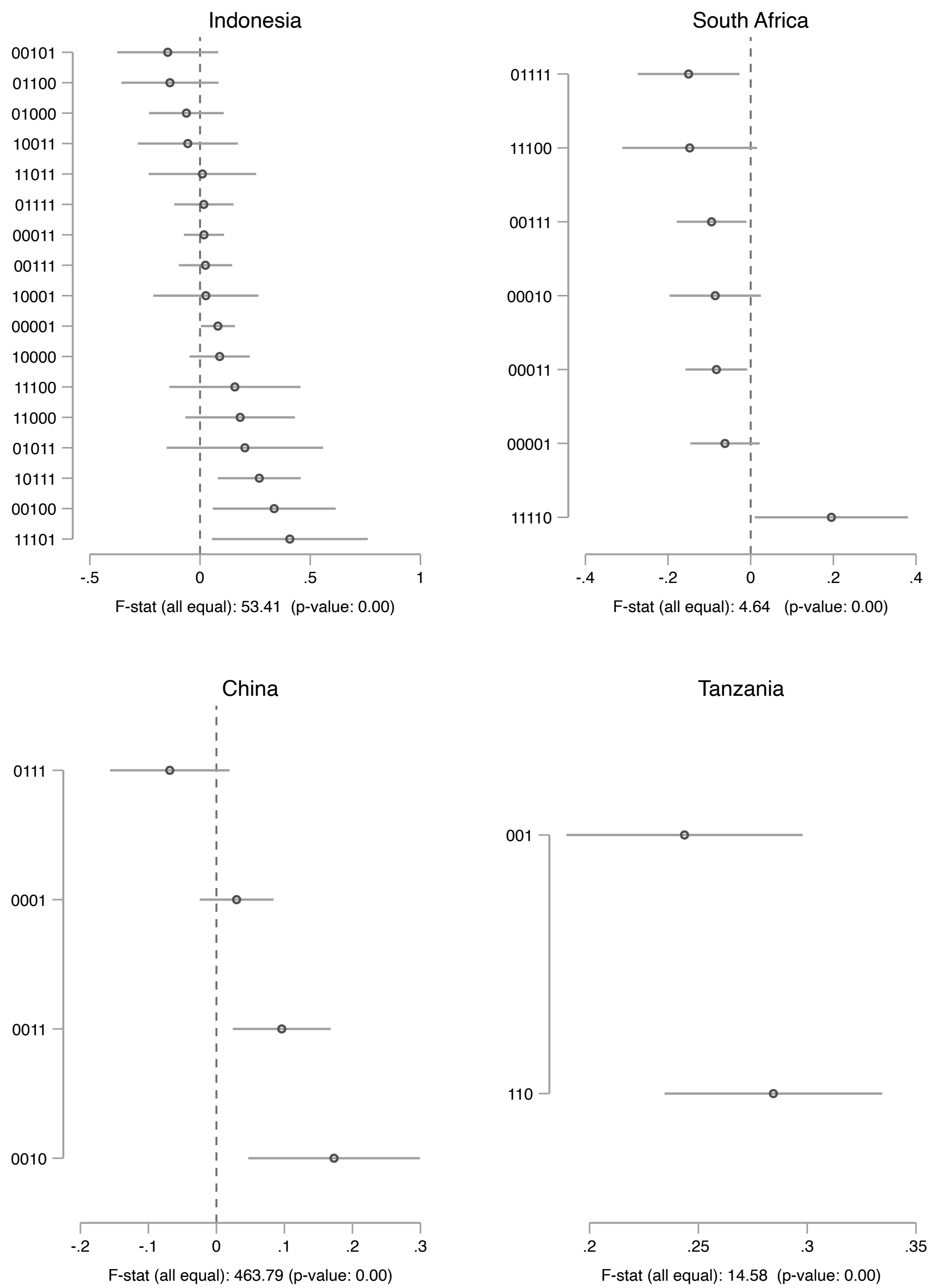
Table 5: OLS estimates of the returns to Non-Agriculture employment on Log Consumption

Dependent Variable:	Log Consumption					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Non-Agriculture	0.462	0.249	0.067	0.168	0.084	0.063
	(0.007)***	(0.006)***	(0.010)***	(0.014)***	(0.011)***	(0.010)***
Observations	77,744	77,744	77,744	12,895	12,895	12,895
Individuals	34,399	34,399	34,399	3,835	3,835	3,835
	Panel B: South Africa					
Non-Agriculture	0.591	0.155	-0.025	0.044	-0.048	-0.026
	(0.016)***	(0.014)***	(0.019)	(0.025)*	(0.021)**	(0.019)
Observations	34,758	34,640	34,645	3,912	3,908	3,908
Individuals	18,129	18,070	18,075	1,279	1,279	1,279
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

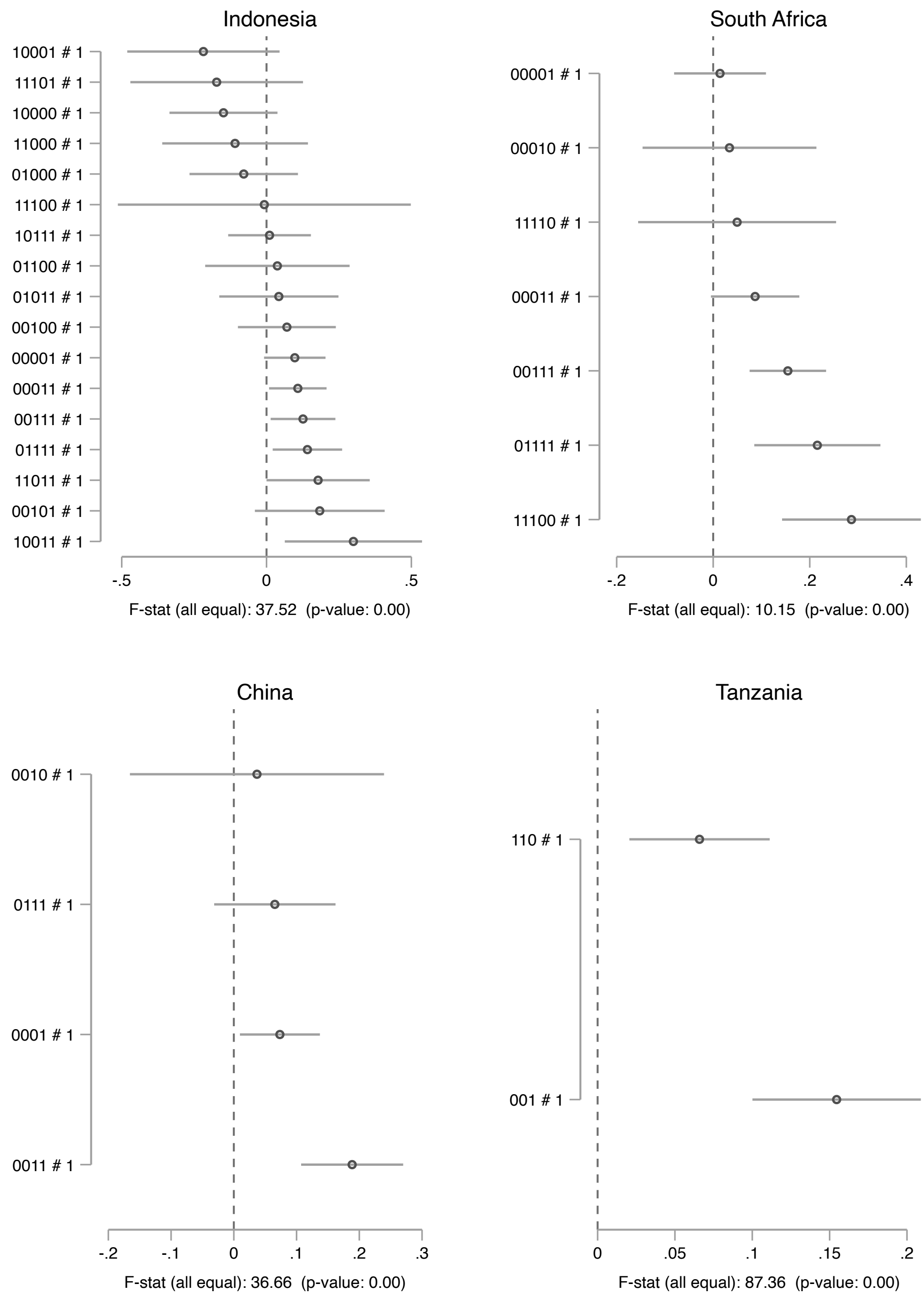
Notes: The dependent variable is the log of total consumption. Non-Agriculture is an indicator equal to one when the individual's primary employment is in non-agriculture. The Switchers sample is restricted to individuals who switch between agriculture/non-agriculture employment at least once. Covariates: female, age (years), age squared, education (years of schooling), education squared, and household size (members). Period refers to the survey round or wave. Coefficients on covariates and period fixed effects not shown. Robust standard errors clustered at the individual level are in parentheses. Stars denote: * p<0.10; ** p<0.05; *** p<0.01.

Figure 3: Unrestricted GRC estimates of the returns to Urban location on Log Consumption, Selected switcher types

Panel A: Base level in the rural market, $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$, *when* $\mu_{\{0,\ldots,0\}} = 0$



Panel B: Urban premium, $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$



Notes: The sample consists of switchers of trajectories that include at least 0.5% of individual in the balanced panel. The dependent variable is the log of total consumption. Dots show point estimates of μ and Δ obtained for selected switcher types from the unrestricted GRC model estimated with covariates and period fixed effects. Bars show 95% confidence intervals. Labels on the y-axis represent trajectory types to which individuals are assigned based on their full location history. For example, an individual whose history is {Rural, Rural, Urban} is in trajectory type {0,0,1}. The "# 1" in the labels on panel B means estimates come from interactions of the trajectory and choice indicators when the second is equal to one. We omit curly brackets and commas in labels using 000 instead of {0,0,0}, 001 instead of {0,0,1}, and so on. Point estimates for μ are normalized so that the estimate for the {0, … ,0} type is zero. For each country, we test whether all point estimates are equal (joint difference equal to zero) and show the F-statistic and p-value from this test below the estimates.

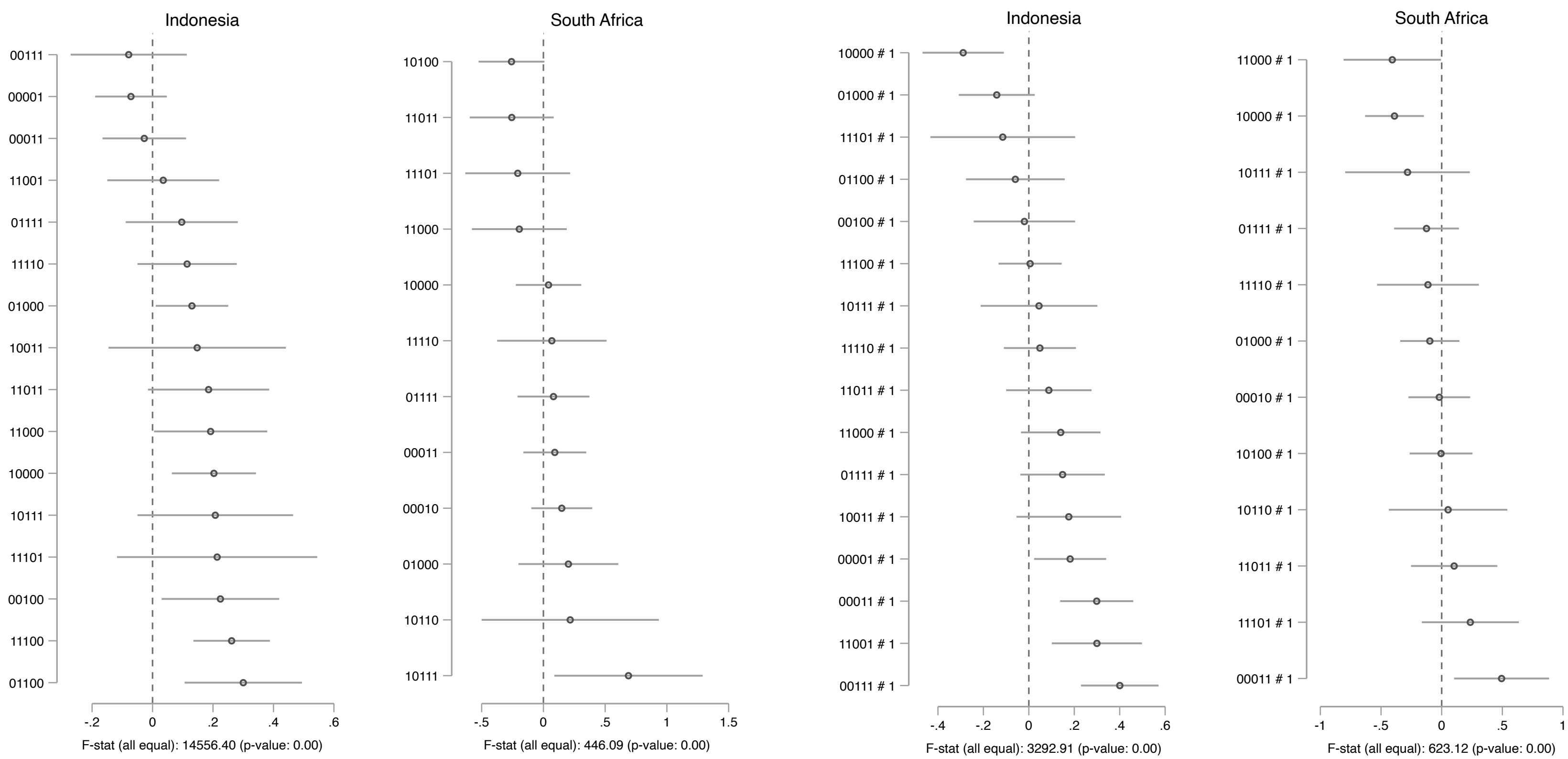
Figure 4: Unrestricted GRC estimates of the returns to Non-Agriculture employment on Log Consumption, Selected switcher types

Panel A: Base level in agriculture

$\mu = \mathrm{E}(y_{it}|D_{it} = 0),\ when\ \mu_{\{0,\ldots,0\}} = 0$

Panel B: Non-Agriculture premium

$\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$



Notes: The dependent variable is the log of total consumption. Dots show point estimates of μ and Δ obtained for selected switcher types from the unrestricted GRC model estimated with covariates and period fixed effects. Bars show 95% confidence intervals. Selected switchers are trajectories whose number of individuals is at least 0.5% of all balanced individuals. Labels on the y-axis represent trajectory types to which individuals are assigned based on their full employment history. For example, an individual whose history is {Agriculture, Agriculture, Non-Agriculture} is in trajectory type {0,0,1}. The "# 1" in the labels on panel B means estimates come from interactions of the trajectory and choice indicators when the second is equal to one. We omit curly brackets and commas in labels using 000 instead of {0,0,0}, 001 instead of {0,0,1}, and so on. Point estimates for μ are normalized so that the estimate for the {0, … ,0} type is zero. For each country, we test whether all point estimates are equal (joint difference equal to zero) and show the F-statistic and p-value from this test below the estimates.

Table 6: Restricted GRC estimates of the returns to Urban location on Log Consumption

Dependent Variable:	Log Consumption							
Country:	Indonesia		South Africa		China		Tanzania	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Never	0.324	0.082	0.109	0.142	0.419	0.095	0.528	0.332
	(0.087)***	(0.021)***	(0.035)***	(0.034)***	(0.022)***	(0.022)***	(0.041)***	(0.055)***
Δ Always	0.412	0.068	1.015	0.228	-0.781	0.050	3.026	-1.503
	(0.046)***	(0.028)**	(0.092)***	(0.050)***	(10.608)	(0.045)	(1.773)*	(2.222)
ϕ (extrapolation slope)	-3.222	-0.077	-1.766	0.525	-0.992	-0.145	-1.117	-0.838
	(0.469)***	(0.155)	(0.151)***	(0.299)*	(0.066)***	(0.140)	(0.077)***	(0.188)***
J-stat (overidentification)	54.268	29.034	51.380	38.034	82.035	18.586	7.756	7.186
p-value	0.001	0.310	0.000	0.018	0.000	0.017	0.051	0.066
Observations	77,744	77,744	104,008	103,568	129,466	113,543	34,527	26,095
Individuals	34,399	34,399	38,427	38,261	49,398	40,340	15,667	12,685
Covariates		Y		Y		Y		Y
Time FE		Y		Y		Y		Y

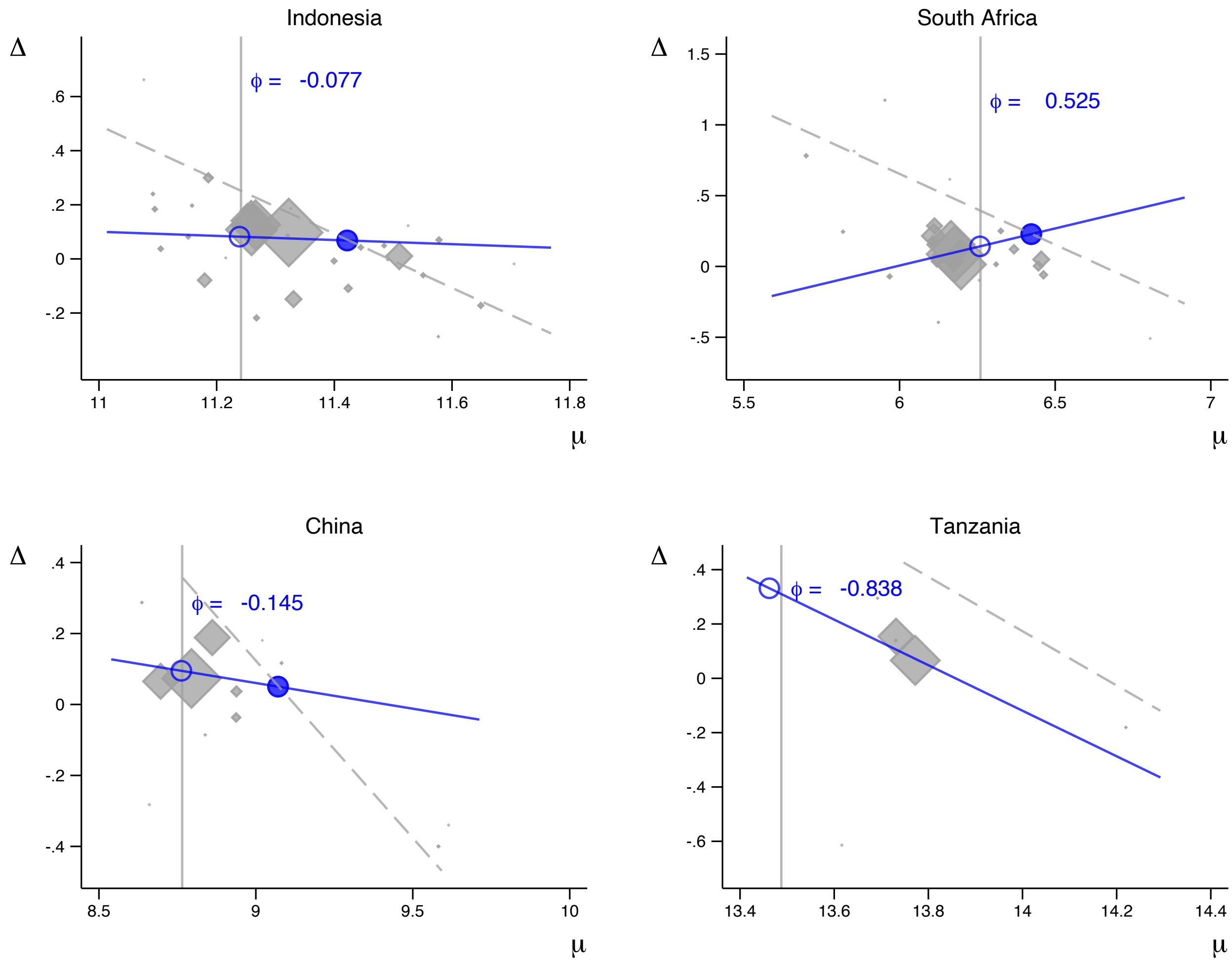
Notes: The dependent variable is the log of total consumption. The choice variable is the location of residence (an indicator equal to one if urban and 0 if rural). Individuals are assigned to trajectory types based on their full location history. Thus, the "Never" type corresponds to the {0, … ,0} trajectory (never urban), and the "Always" type to the {1, … ,1} trajectory (always urban). Estimates of μ and Δ for switcher types, $\mu_{\{0,\dots,0\}}$, and $\mu_{\{1,\dots,1\}}$ not shown. Covariates: female, age (years), age squared, education (years of schooling), education squared, and household size (members). Period refers to the survey round or wave. Coefficients on covariates and period fixed effects not shown. J-statistics and its p-value from the overidentification test are shown at the bottom of the table. Robust standard errors clustered at the individual level are in parentheses. Stars denote: * p<0.10; ** p<0.05; *** p<0.01.

Table 7: Restricted GRC estimates of the returns to Non-Agriculture employment on Log Consumption

Dependent Variable:	Log Consumption			
Country:	Indonesia		South Africa	
	(1)	(2)	(3)	(4)
Δ Never	0.444	0.200	-0.536	-0.065
	(0.062)***	(0.034)***	(0.222)**	(0.037)*
Δ Always	0.582	-3.308	0.378	-0.015
	(0.069)***	(15.023)	(0.088)***	(0.044)
ϕ (extrapolation slope)	-1.918	-0.967	1.528	0.202
	(0.140)***	(0.139)***	(0.629)**	(0.142)
J-stat (overidentification)	40.037	30.682	34.145	27.817
p-value	0.011	0.103	0.035	0.145
Observations	77,744	77,744	34,758	34,640
Individuals	34,399	34,399	18,129	18,070
Covariates		Y		Y
Time FE		Y		Y

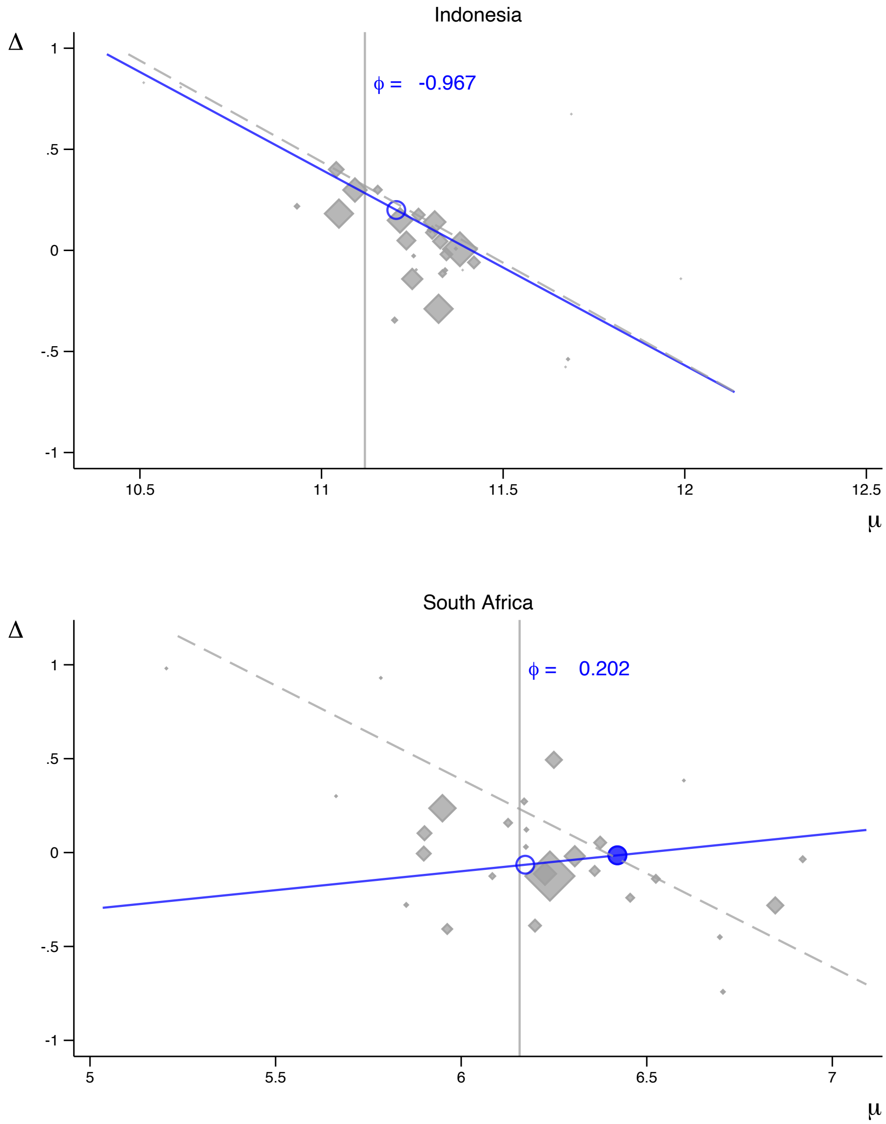
Notes: The dependent variable is the log of total consumption. The choice variable is the sector of employment (an indicator equal to one if non-agriculture and 0 if agriculture). Individuals are assigned to trajectory types based on their full employment history. Thus, the "Never" type corresponds to the {0, … ,0} trajectory (never in non-agriculture), and the "Always" type to the {1, … ,1} trajectory (always in non-agriculture). Estimates of μ and Δ for switcher types, $\mu_{\{0,\dots,0\}}$, and $\mu_{\{1,\dots,1\}}$ not shown. Covariates: female, age (years), age squared, education (years of schooling), education squared, and household size (members). Period refers to the survey round or wave. Coefficients on covariates and period fixed effects not shown. J-statistics and its p-value from the overidentification test are shown at the bottom of the table. Robust standard errors clustered at the individual level are in parentheses. Stars denote: * p<0.10; ** p<0.05; *** p<0.01.

Figure 5: Extrapolation line and GRC estimates of the returns to Urban location on Log Consumption for switchers and non-switchers



Notes: The parameter $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$ in the x-axis denotes the base consumption level in the rural labor market, and the parameter $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$ in the y-axis denotes the increment in consumption received in the urban market over the base level. The dependent variable is the log of total consumption. Results come from estimating the restricted and unrestricted GRC models with covariates and period fixed effects in the regression specification. The gray diamonds represent point estimates from the unrestricted GRC: the μs and Δs for all switcher types, μ for the non-switcher type $\{0, \ldots, 0\}$, and $\kappa = \mu + \Delta$ for the non-switcher type $\{1, \ldots, 1\}$. Their sizes are proportional to the number of individuals in each trajectory relative to the total number of switchers. In blue, we show the slope of the extrapolation line, ϕ, and the estimates identified by it in the restricted GRC: Δ for the non-switcher type $\{0, \ldots, 0\}$ (hollow circle), and μ and Δ for the non-switcher type $\{1, \ldots, 1\}$ (filled circle). Circle sizes are fixed and have no connection to the number of individuals in the non-switching trajectories. Confidence intervals are not shown. We place the extrapolation line intercepting the (μ, Δ) estimate of the switcher trajectory with the most precise Δ estimate among those with more than five individuals in the sample. To improve visualization, we omit the (μ, Δ) estimate for the non-switcher type $\{1, \ldots, 1\}$ when it would extend the scale of the graph considerably. We omit all estimates from the restricted GRC model when its GMM estimation does not converge.

Figure 6: Extrapolation line and GRC estimates of the returns to Non-Agriculture employment on Log Consumption for switchers and non-switchers



Notes: The parameter $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$ in the x-axis denotes the base consumption level in agriculture, and the parameter $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$ in the y-axis denotes the increment in consumption received in non-agriculture over the base level. The dependent variable is the log of total consumption. Results come from estimating the restricted and unrestricted GRC models with covariates and period fixed effects in the regression specification. The gray diamonds represent point estimates from the unrestricted GRC: the μs and Δs for all switcher types, μ for the non-switcher type $\{0, \dots, 0\}$, and $\kappa = \mu + \Delta$ for the non-switcher type $\{1, \dots, 1\}$. Their sizes are proportional to the number of individuals in each trajectory relative to the total number of switchers. In blue, we show the slope of the extrapolation line, ϕ, and the estimates identified by it in the restricted GRC: Δ for the non-switcher type $\{0, \dots, 0\}$ (hollow circle), and μ and Δ for the non-switcher type $\{1, \dots, 1\}$ (filled circle). Circle sizes are fixed and have no connection to the number of individuals in the non-switching trajectories. Confidence intervals are not shown. We place the extrapolation line intercepting the (μ, Δ) estimate of the switcher trajectory with the most precise Δ estimate among those with more than five individuals in the sample. To improve visualization, we omit the (μ, Δ) estimate for the non-switcher type $\{1, \dots, 1\}$ when it would extend the scale of the graph considerably. We omit all estimates from the restricted GRC model when its GMM estimation does not converge.

Appendix

Table A1: Summary statistics and differences by Location of Birth (Rural/Urban), Indonesia

	All	Born Urban	Born Rural	Difference
Urban	0.45	0.79	0.33	0.47
	(0.50)	(0.41)	(0.47)	[0.00]
Urban Switcher	0.22	0.18	0.24	-0.06
	(0.42)	(0.38)	(0.43)	[0.00]
Non-Agricultural	0.67	0.88	0.59	0.29
	(0.47)	(0.32)	(0.49)	[0.00]
Log Consumption	12.08	12.29	12.00	0.30
	(0.79)	(0.77)	(0.79)	[0.00]
Log Income	14.90	15.20	14.79	0.41
	(1.14)	(1.08)	(1.14)	[0.00]
No Income	0.15	0.10	0.16	-0.06
	(0.35)	(0.31)	(0.37)	[0.00]
Female	0.44	0.44	0.44	-0.00
	(0.50)	(0.50)	(0.50)	[0.55]
Age (years)	38.74	36.01	39.76	-3.75
	(13.81)	(12.27)	(14.20)	[0.00]
Education (years)	7.83	9.86	7.07	2.78
	(4.58)	(4.30)	(4.45)	[0.00]
Household Size	4.79	5.01	4.71	0.29
	(2.19)	(2.34)	(2.12)	[0.00]
Observations	77,546	20,964	56,582	77,546
Share	100%	27%	73%	

Notes: Standard errors are in parentheses, and the p-value of the difference in means is in square brackets

Table A2: Summary statistics and differences by Rural/Urban location, Balanced panel

	Indonesia			South Africa			China			Tanzania		
	Urban	Rural	Difference	Urban	Rural	Difference	Urban	Rural	Difference	Urban	Rural	Difference
Urban Switcher	0.47	0.33	0.15	0.12	0.13	-0.01	0.09	0.08	0.01	0.21	0.11	0.11
	(0.50)	(0.47)	[0.00]	(0.33)	(0.34)	[0.00]	(0.28)	(0.27)	[0.00]	(0.41)	(0.31)	[0.00]
Non-Agricultural	0.86	0.44	0.42	0.94	0.73	0.21						
	(0.35)	(0.50)	[0.00]	(0.23)	(0.44)	[0.00]						
Log Consumption	12.23	11.85	0.38	8.10	7.63	0.47	10.63	10.21	0.42	15.29	14.61	0.68
	(0.77)	(0.80)	[0.00]	(0.95)	(0.79)	[0.00]	(0.89)	(0.89)	[0.00]	(0.79)	(0.71)	[0.00]
Log Income	15.27	14.80	0.47	7.80	7.34	0.46	9.13	8.11	1.03	14.55	13.29	1.26
	(1.01)	(1.11)	[0.00]	(1.16)	(1.11)	[0.00]	(1.58)	(2.05)	[0.00]	(1.83)	(1.81)	[0.00]
No Income	0.05	0.12	-0.07	0.51	0.70	-0.18	0.49	0.59	-0.10	0.48	0.64	-0.17
	(0.22)	(0.32)	[0.00]	(0.50)	(0.46)	[0.00]	(0.50)	(0.49)	[0.00]	(0.50)	(0.48)	[0.00]
Female	0.34	0.36	-0.02	0.60	0.65	-0.05	0.52	0.51	0.01	0.52	0.52	0.00
	(0.47)	(0.48)	[0.02]	(0.49)	(0.48)	[0.00]	(0.50)	(0.50)	[0.00]	(0.50)	(0.50)	[0.89]
Age (years)	44.70	43.55	1.15	40.85	42.91	-2.06	49.11	49.08	0.03	35.72	39.07	-3.35
	(11.47)	(11.75)	[0.00]	(15.52)	(17.67)	[0.00]	(14.92)	(14.45)	[0.82]	(15.24)	(17.16)	[0.00]
Education (years)	8.56	6.23	2.33	9.30	7.18	2.13	8.40	5.61	2.78	8.54	6.93	1.61
	(4.38)	(4.28)	[0.00]	(3.73)	(4.61)	[0.00]	(4.65)	(4.37)	[0.00]	(2.90)	(2.33)	[0.00]
Household Size	4.91	4.73	0.18	4.56	5.94	-1.38	3.78	4.40	-0.62	5.86	6.49	-0.63
	(2.16)	(1.90)	[0.00]	(2.73)	(3.63)	[0.00]	(1.68)	(1.88)	[0.00]	(3.22)	(4.43)	[0.00]
Observations	3,702	5,728	9,430	20,969	22,606	43,575	25,155	32,721	57,876	8,361	15,165	23,526
Share	39%	61%		48%	52%		43%	57%		36%	64%	

Notes: Standard errors are in parentheses, and the p-value of the difference in means is in square brackets

Table A3: Summary statistics and differences by Agriculture/Non-Agriculture employment, Balanced panel

	Indonesia			South Africa		
	Non-Ag.	Agric.	Difference	Non-Ag.	Agric.	Difference
Non-Ag. Switcher	0.23	0.37	-0.14	0.11	0.57	-0.47
	(0.42)	(0.48)	[0.00]	(0.31)	(0.50)	[0.00]
Urban	0.56	0.14	0.42	0.71	0.21	0.49
	(0.50)	(0.35)	[0.00]	(0.46)	(0.41)	[0.00]
Log Consumption	12.16	11.77	0.39	8.38	7.66	0.72
	(0.80)	(0.77)	[0.00]	(0.99)	(0.82)	[0.00]
Log Income	15.25	14.53	0.72	8.11	7.50	0.62
	(1.00)	(1.10)	[0.00]	(1.14)	(0.76)	[0.00]
No Income	0.04	0.17	-0.13	0.00	0.03	-0.02
	(0.20)	(0.38)	[0.00]	(0.04)	(0.16)	[0.00]
Female	0.36	0.34	0.02	0.55	0.28	0.27
	(0.48)	(0.47)	[0.02]	(0.50)	(0.45)	[0.00]
Age (years)	42.79	45.87	-3.07	41.58	41.03	0.55
	(10.92)	(12.48)	[0.00]	(9.33)	(10.95)	[0.27]
Education (years)	8.33	5.30	3.03	10.64	6.63	4.01
	(4.53)	(3.66)	[0.00]	(3.54)	(4.11)	[0.00]
Household Size	4.91	4.63	0.28	4.10	4.12	-0.02
	(2.05)	(1.93)	[0.00]	(2.55)	(2.98)	[0.91]
Observations	5,730	3,700	9,430	3,886	539	4,425
Share	61%	39%		88%	12%	

Notes: Standard errors are in parentheses, and the p-value of the difference in means is in square brackets

Table A4: OLS estimates of the returns to Urban location on Log Consumption, Balanced panel

Dependent Variable:	Log Consumption					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Urban	0.379	0.181	0.020	0.371	0.072	0.034
	(0.021)***	(0.019)***	(0.021)	(0.027)***	(0.024)***	(0.023)
Observations	9,430	9,430	9,430	3,615	3,615	3,615
Individuals	1,886	1,886	1,886	723	723	723
	Panel B: South Africa					
Urban	0.472	0.383	0.147	0.252	0.170	0.151
	(0.014)***	(0.012)***	(0.020)***	(0.024)***	(0.024)***	(0.023)***
Observations	43,575	43,457	43,457	5,600	5,578	5,578
Individuals	8,715	8,712	8,712	1,120	1,120	1,120
	Panel C: China					
Urban	0.424	0.330	0.095	0.361	0.034	0.044
	(0.011)***	(0.010)***	(0.021)***	(0.025)***	(0.028)	(0.026)*
Observations	57,876	56,754	56,754	4,860	4,637	4,637
Individuals	14,469	14,214	14,214	1,215	1,166	1,166
	Panel D: Tanzania					
Urban	0.684	0.580	0.100	0.127	0.117	0.108
	(0.015)***	(0.013)***	(0.016)***	(0.020)***	(0.018)***	(0.016)***
Observations	23,526	17,752	17,752	3,414	2,762	2,762
Individuals	7,842	6,494	6,494	1,138	1,007	1,007
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

Notes: See notes in Table 4.

Table A5: OLS estimates of the returns to Non-Agriculture employment on Log Consumption, Balanced panel

Dependent Variable:	Log Consumption					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Non-Agriculture	0.389	0.264	0.073	0.108	0.098	0.072
	(0.022)***	(0.019)***	(0.023)***	(0.034)***	(0.024)***	(0.023)***
Observations	9,430	9,430	9,430	2,700	2,700	2,700
Individuals	1,886	1,886	1,886	540	540	540
	Panel B: South Africa					
Non-Agriculture	0.719	0.218	-0.015	0.040	-0.017	-0.014
	(0.058)***	(0.044)***	(0.048)	(0.067)	(0.051)	(0.049)
Observations	4,425	4,413	4,413	725	724	724
Individuals	885	885	885	145	145	145
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

Notes: See notes in Table 5.

Table A6: OLS estimates of the returns to Urban location on Log Income

Dependent Variable:	Log Income					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Urban	0.497	0.267	0.040	0.210	0.122	0.057
	(0.011)***	(0.009)***	(0.014)***	(0.016)***	(0.015)***	(0.015)***
Observations	68,877	68,283	68,283	15,286	15,171	15,171
Individuals	31,310	31,108	31,108	4,791	4,790	4,790
	Panel B: South Africa					
Urban	0.538	0.258	0.069	0.292	0.057	0.027
	(0.017)***	(0.013)***	(0.030)**	(0.041)***	(0.035)	(0.032)
Observations	37,186	37,015	37,022	2,496	2,489	2,489
Individuals	18,560	18,471	18,478	782	782	782
	Panel C: China					
Urban	0.886	0.646	0.249	0.363	0.135	0.185
	(0.017)***	(0.016)***	(0.060)***	(0.048)***	(0.058)**	(0.056)***
Observations	64,614	54,603	54,603	3,422	2,789	2,789
Individuals	38,743	31,435	31,435	1,404	1,163	1,163
	Panel D: Tanzania					
Urban	1.185	0.870	0.033	0.231	0.102	0.060
	(0.036)***	(0.036)***	(0.091)	(0.087)***	(0.096)	(0.093)
Observations	13,995	11,618	11,618	1,279	1,092	1,092
Individuals	8,456	7,012	7,012	510	449	449
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

Notes: See notes in Table 4.

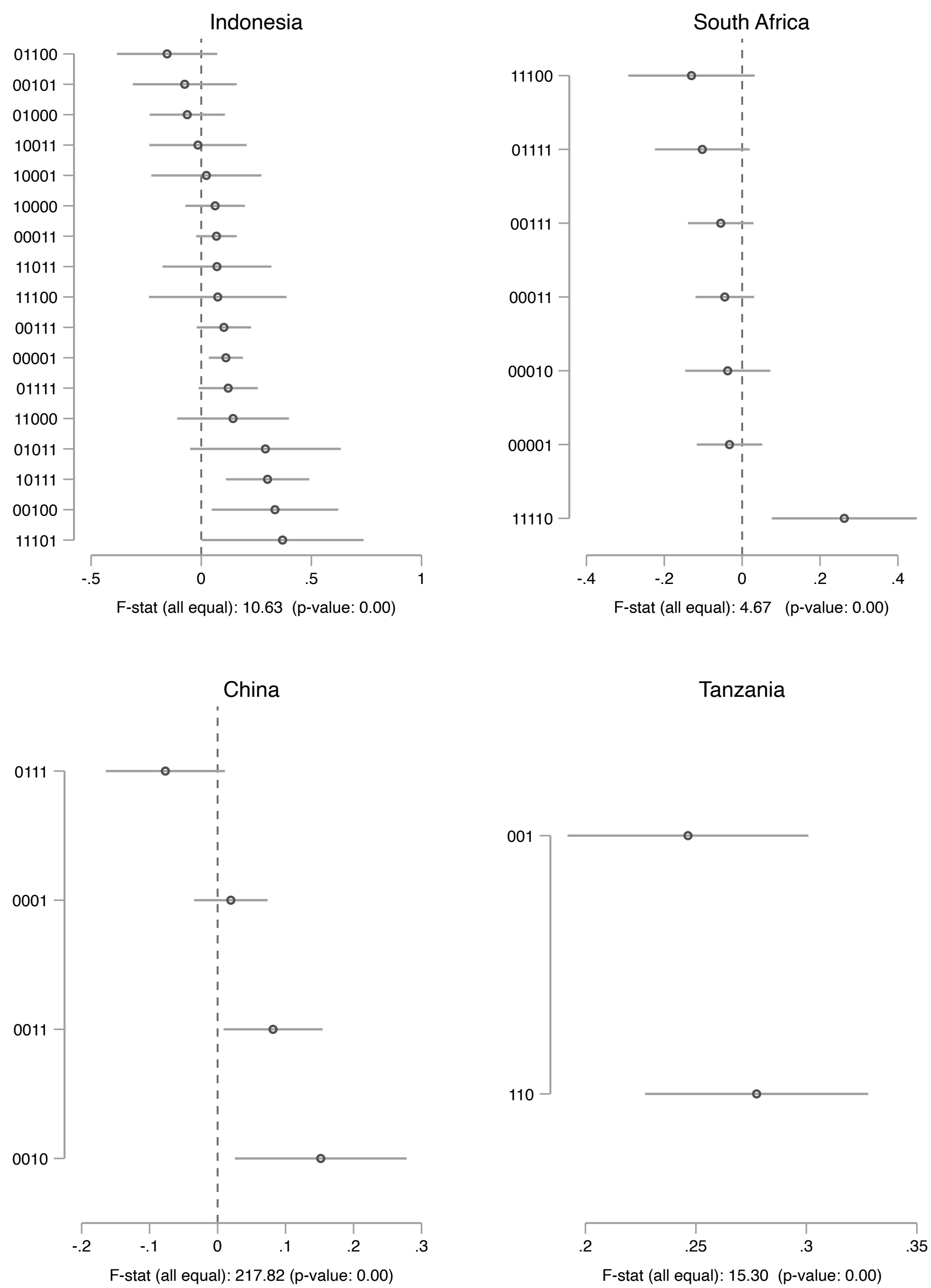
Table A7: OLS estimates of the returns to Non-Agriculture employment on Log Income

Dependent Variable:	Log Income					
Sample:	Full			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Indonesia					
Non-Agriculture	0.685	0.431	0.227	0.330	0.262	0.226
	(0.012)***	(0.012)***	(0.020)***	(0.022)***	(0.020)***	(0.020)***
Observations	68,877	68,283	68,283	10,169	10,096	10,096
Individuals	31,310	31,108	31,108	3,037	3,037	3,037
	Panel B: South Africa					
Non-Agriculture	0.432	-0.078	-0.032	-0.012	-0.058	-0.030
	(0.019)***	(0.018)***	(0.028)	(0.034)	(0.029)**	(0.028)
Observations	33,218	33,104	33,109	3,052	3,049	3,049
Individuals	17,139	17,083	17,088	965	965	965
Covariates		Y	Y		Y	Y
Period FE		Y	Y		Y	Y
Individual FE			Y			Y

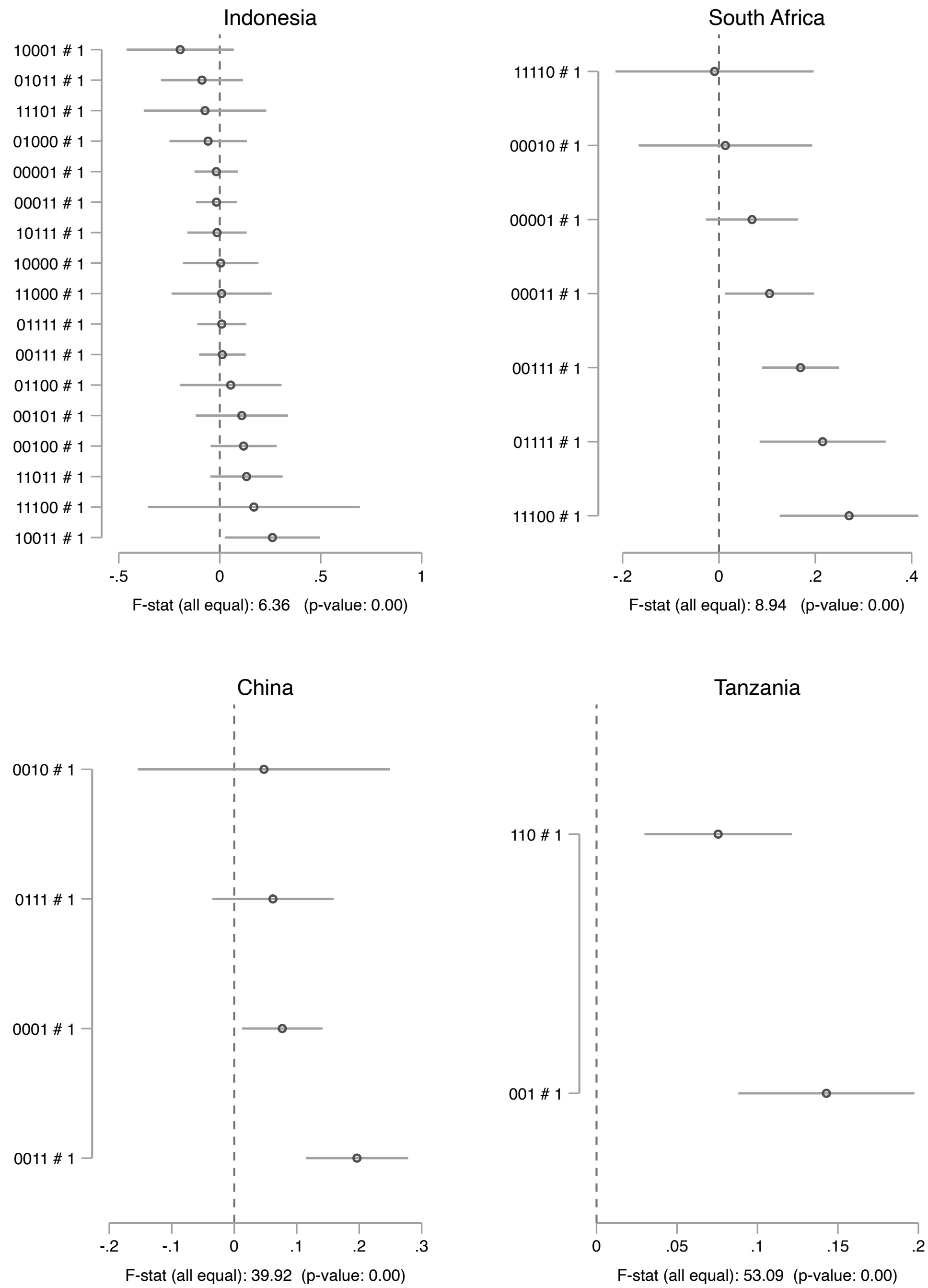
Notes: See notes in Table 5.

Figure A1: Unrestricted GRC estimates of the returns to Urban location on Log Consumption, Selected switcher types, Balanced panel

Panel A: Base level in the rural market, $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$, *when* $\mu_{\{0,\ldots,0\}} = 0$



Panel B: Urban premium, $\Delta = \mathrm{E}(y_{it}|D_{it}=1) - \mathrm{E}(y_{it}|D_{it}=0)$



Notes: See notes in Figure 3.

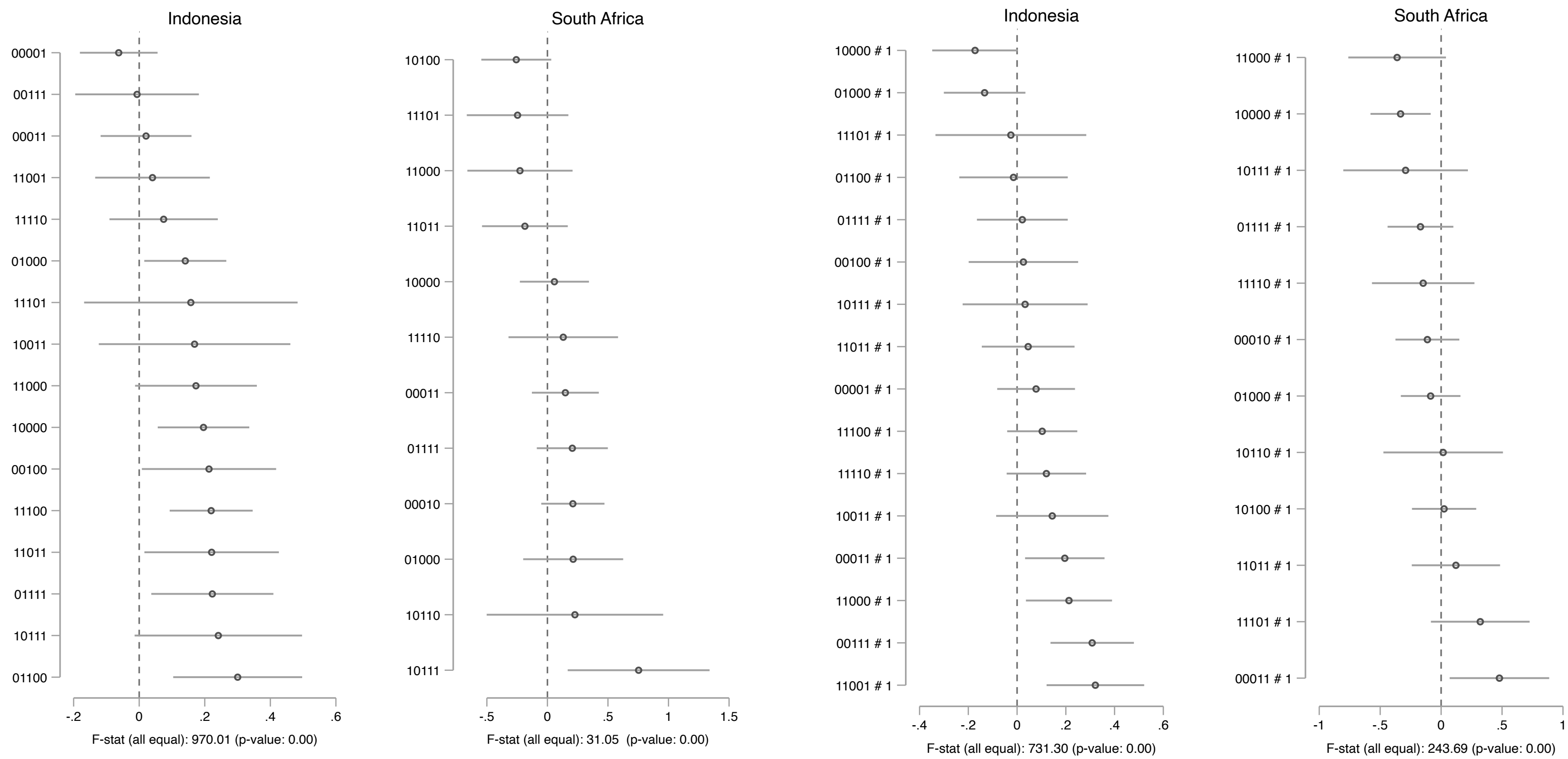
Figure A2: Unrestricted GRC estimates of the returns to Non-Agriculture employment on Log Consumption, Selected switcher types, Balanced panel

Panel A: Base level in agriculture

$\mu = \mathrm{E}(y_{it}|D_{it} = 0)$, *when* $\mu_{\{0,\dots,0\}} = 0$

Panel B: Non-Agriculture premium

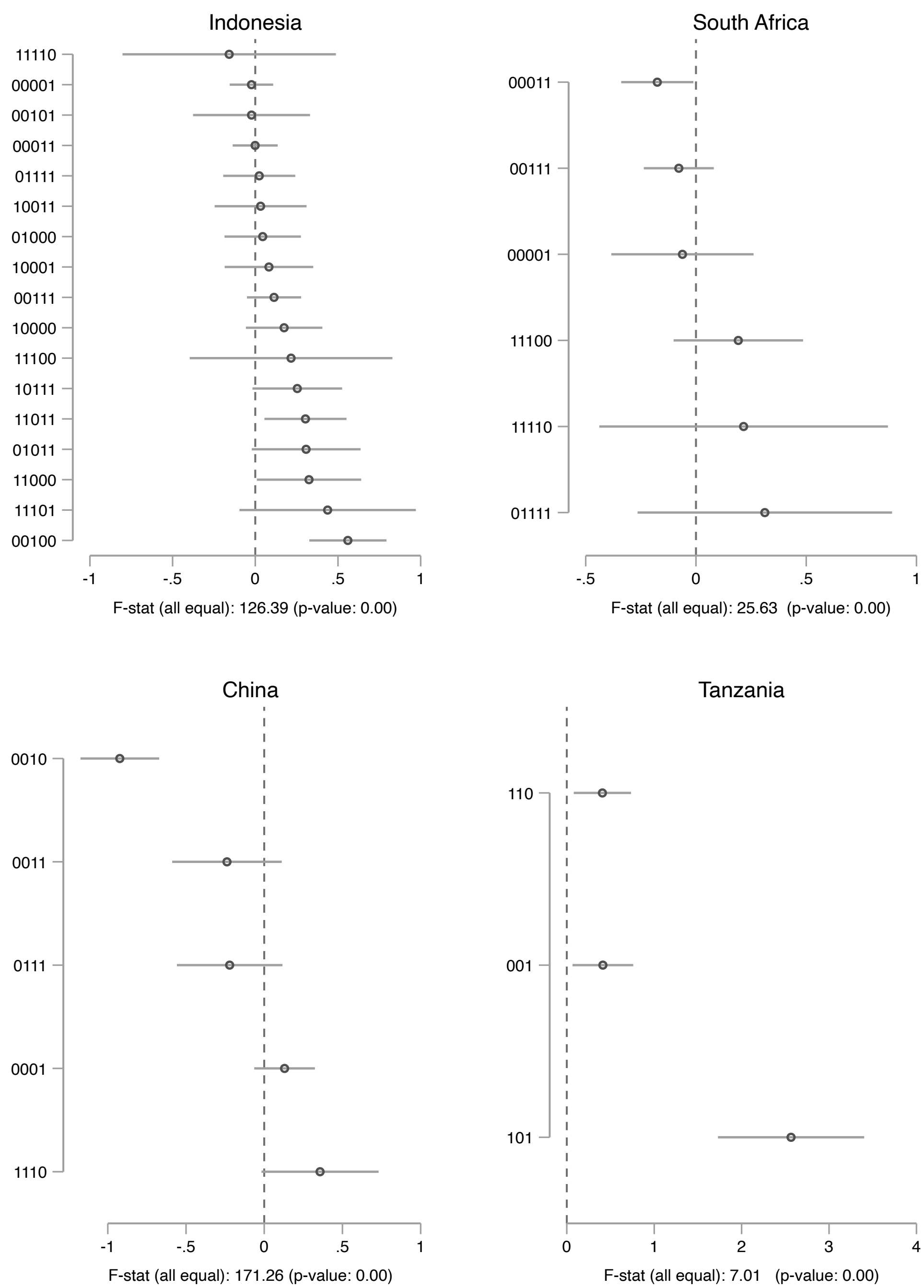
$\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$



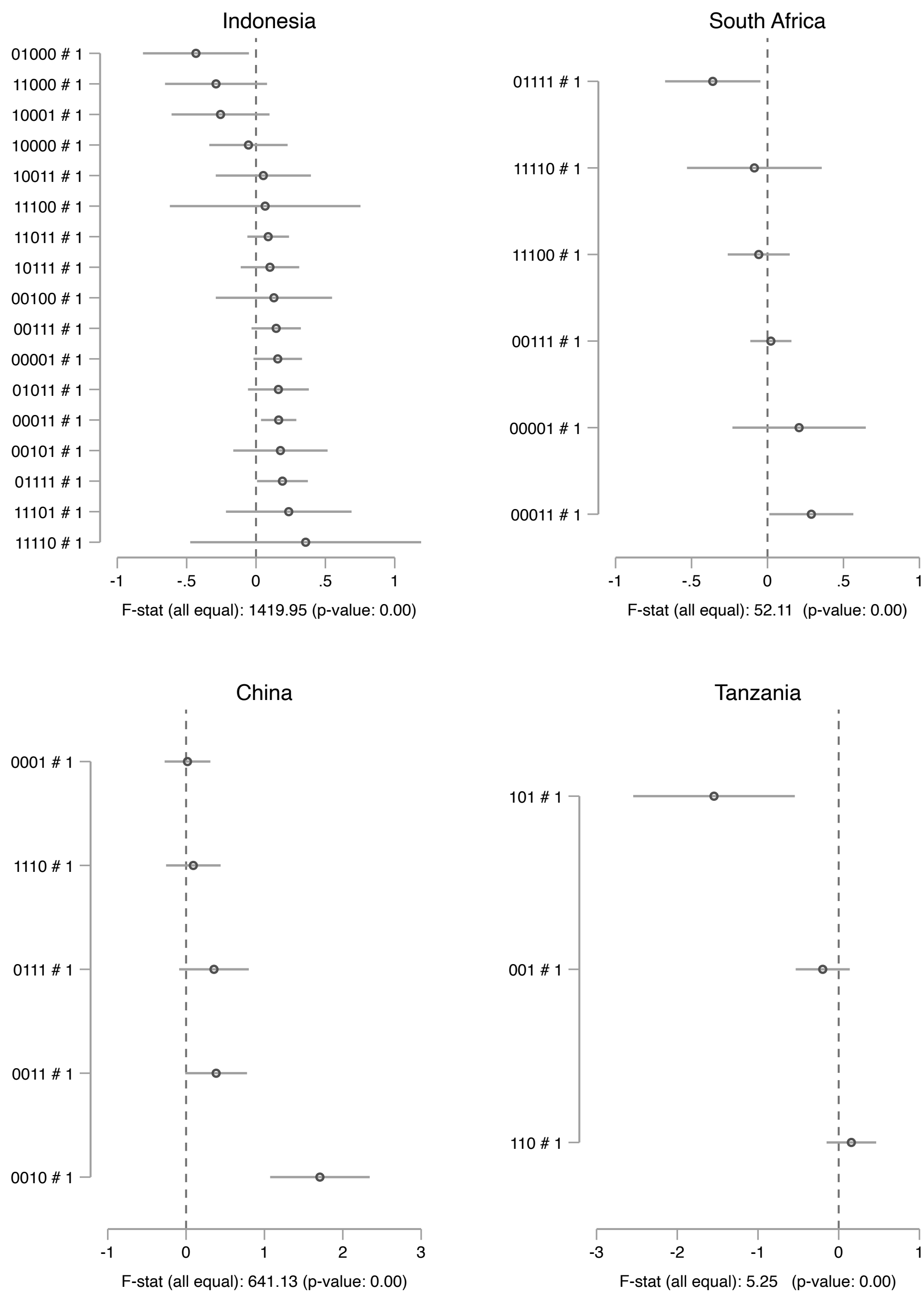
<u>Notes</u>: See notes in Figure 4.

Figure A3: Unrestricted GRC estimates of the returns to Urban location on Log Income, Selected switcher types

Panel A: Base level in the rural market, $\mu = \mathrm{E}(y_{it}|D_{it} = 0)$*, when* $\mu_{\{0,\ldots,0\}} = 0$



Panel B: Urban premium, $\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$



<u>Notes</u>: See notes in Figure 3.

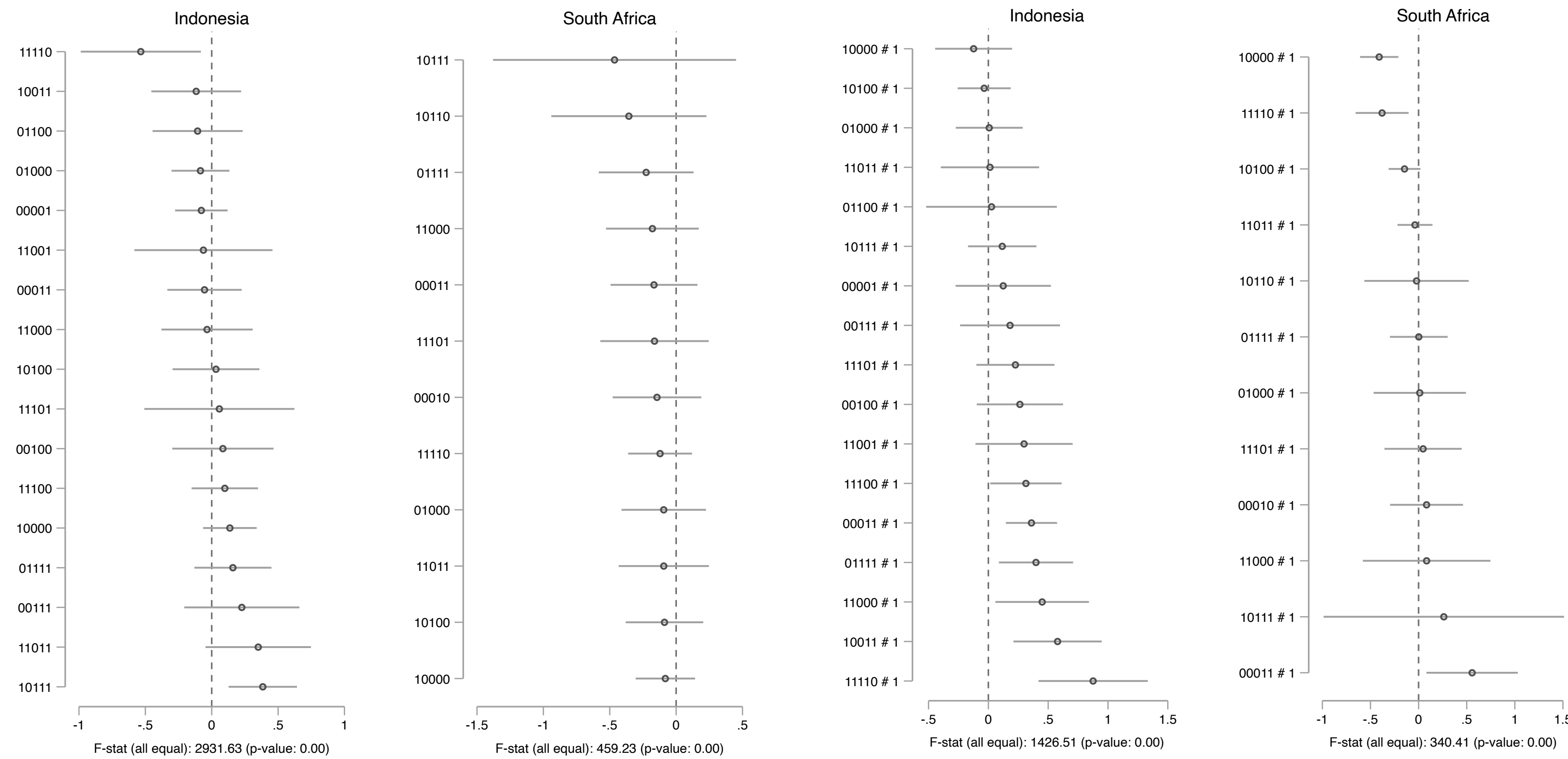
Figure A4: Unrestricted GRC estimates of the returns to Non-Agriculture employment on Log Income, Selected switcher types

Panel A: Base level in agriculture

$\mu = \mathrm{E}(y_{it}|D_{it} = 0)$, *when* $\mu_{\{0,\ldots,0\}} = 0$

Panel B: Non-Agriculture premium

$\Delta = \mathrm{E}(y_{it}|D_{it} = 1) - \mathrm{E}(y_{it}|D_{it} = 0)$



Notes: See notes in Figure 4.

Table A8: Restricted GRC estimates of the returns to Urban location on Log Consumption, Balanced panel

Dependent Variable:	Log Consumption							
Country:	Indonesia		South Africa		China		Tanzania	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Never	0.324	0.015	0.108	0.129	0.419	0.096	0.528	0.304
	(0.088)***	(0.028)	(0.035)***	(0.025)***	(0.022)***	(0.022)***	(0.041)***	(0.048)***
Δ Always	0.412	0.035	1.012	0.206	-1.537	0.050	3.026	-0.787
	(0.046)***	(0.029)	(0.091)***	(0.052)***	(27.667)	(0.047)	(1.773)*	(0.751)
ϕ (extrapolation slope)	-3.227	0.089	-1.772	0.375	-0.995	-0.149	-1.117	-0.742
	(0.471)***	(0.182)	(0.152)***	(0.285)	(0.066)***	(0.140)	(0.077)***	(0.162)***
J-stat (overidentification)	54.270	21.808	51.474	37.061	82.002	19.419	7.756	5.679
p-value	0.001	0.699	0.000	0.023	0.000	0.013	0.051	0.128
Observations	9,430	9,430	43,575	43,457	57,876	56,754	23,526	17,752
Individuals	1,886	1,886	8,715	8,712	14,469	14,214	7,842	6,494
Covariates		Y		Y		Y		Y
Time FE		Y		Y		Y		Y

Notes: See notes in Table 6.

Table A9: Restricted GRC estimates of the returns to Non-Agriculture employment on Log Consumption, Balanced panel

Dependent Variable:	Log Consumption			
Country:	Indonesia		South Africa	
	(1)	(2)	(3)	(4)
Δ Never	0.444	.	-0.538	-0.075
	(0.062)***	.	(0.223)**	(0.044)*
Δ Always	0.582	.	0.379	0.026
	(0.069)***	.	(0.089)***	(0.042)
ϕ (extrapolation slope)	-1.920	.	1.536	0.314
	(0.141)***	.	(0.634)**	(0.164)*
J-stat (overidentification)	40.037	.	34.147	29.973
p-value	0.011	.	0.035	0.093
Observations	9,430	.	4,425	4,413
Individuals	1,886	.	885	885
Covariates		Y		Y
Time FE		Y		Y

Notes: See notes in Table 7.

Table A10: Restricted GRC estimates of the returns to Urban location on Log Income

Dependent Variable:	Log Income							
Country:	Indonesia		South Africa		China		Tanzania	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Never	0.175	0.113	.	-0.021	0.569	0.248	-0.727	0.220
	(0.037)***	(0.037)***	.	(0.034)	(0.090)***	(0.086)***	(0.593)	(0.090)**
Δ Always	0.144	0.126	.	-0.001	0.555	0.406	0.438	-0.837
	(0.030)***	(0.028)***	.	(0.038)	(0.102)***	(0.419)	(0.144)***	(0.506)*
ϕ (extrapolation slope)	-0.064	0.048	.	0.086	0.097	-0.744	1.478	-0.592
	(0.080)	(0.150)	.	(0.101)	(0.166)	(0.267)***	(1.011)	(0.131)***
J-stat (overidentification)	28.766	23.577	.	16.943	7.208	4.227	8.370	7.167
p-value	0.274	0.544	.	0.202	0.125	0.238	0.039	0.067
Observations	68,877	68,283	.	37,015	64,614	54,603	13,995	11,618
Individuals	31,310	31,108	.	18,471	38,743	31,435	8,456	7,012
Covariates		Y		Y		Y		Y
Time FE		Y		Y		Y		Y

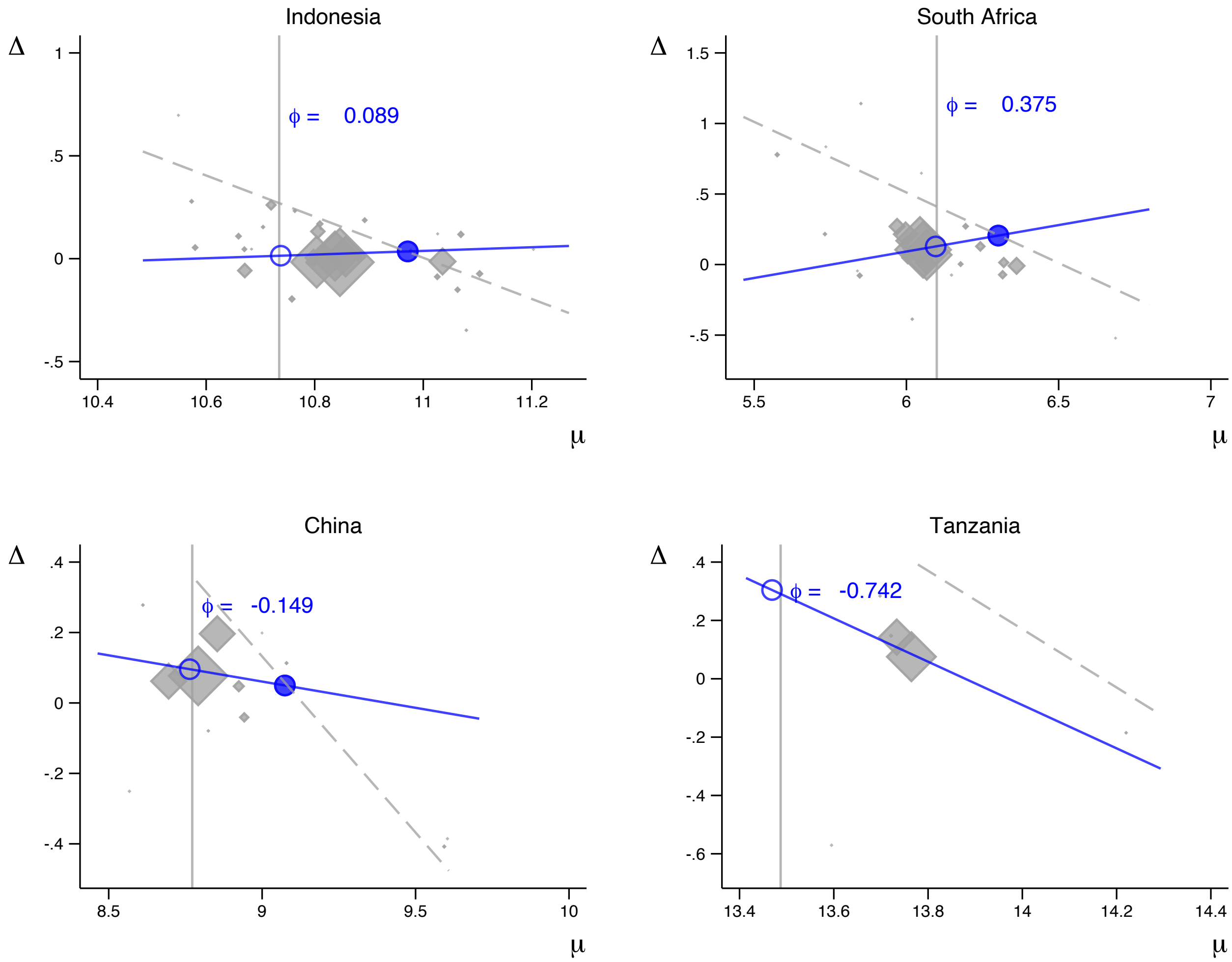
Notes: See notes in Table 6.

Table A11: Restricted GRC estimates of the returns to Non-Agriculture employment on Log Income

Dependent Variable:	Log Income			
Country:	Indonesia		South Africa	
	(1)	(2)	(3)	(4)
Δ Never	0.204	0.112	-0.276	-0.065
	(0.085)**	(0.113)	(0.261)	(0.055)
Δ Always	0.508	0.404	0.508	-0.036
	(0.074)***	(0.055)***	(0.105)***	(0.033)
ϕ (extrapolation slope)	0.975	1.581	3.490	0.626
	(0.490)**	(0.878)*	(1.780)**	(0.270)**
J-stat (overidentification)	28.221	25.807	44.845	19.158
p-value	0.168	0.260	0.001	0.512
Observations	68,877	68,283	33,218	33,104
Individuals	31,310	31,108	17,139	17,083
Covariates		Y		Y
Time FE		Y		Y

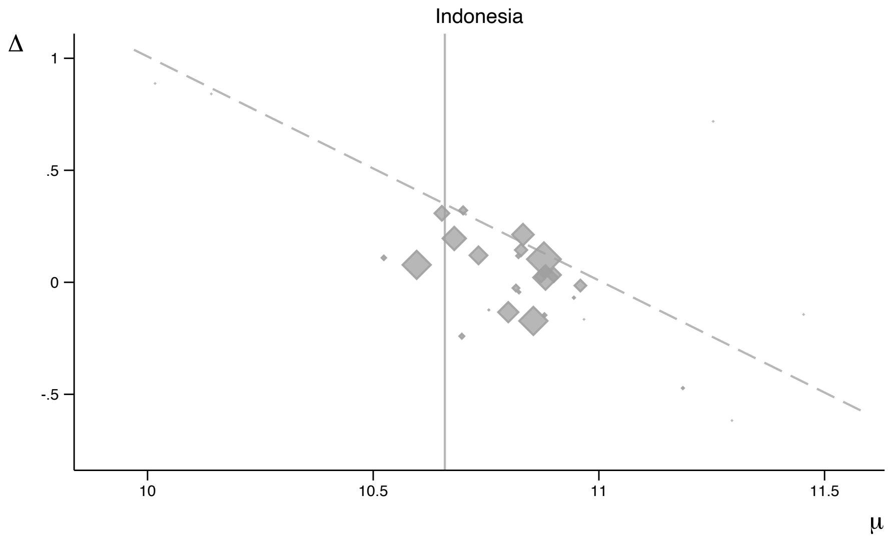
Notes: See notes in Table 7.

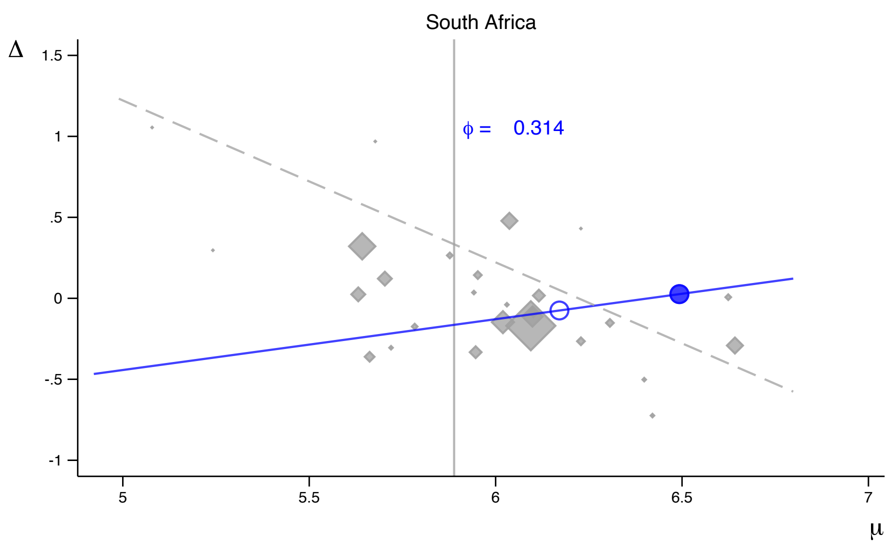
Figure A5: Extrapolation line and GRC estimates of the returns to Urban location on Log Consumption for switchers and non-switchers, Balanced panel



Notes: See notes in Figure 5.

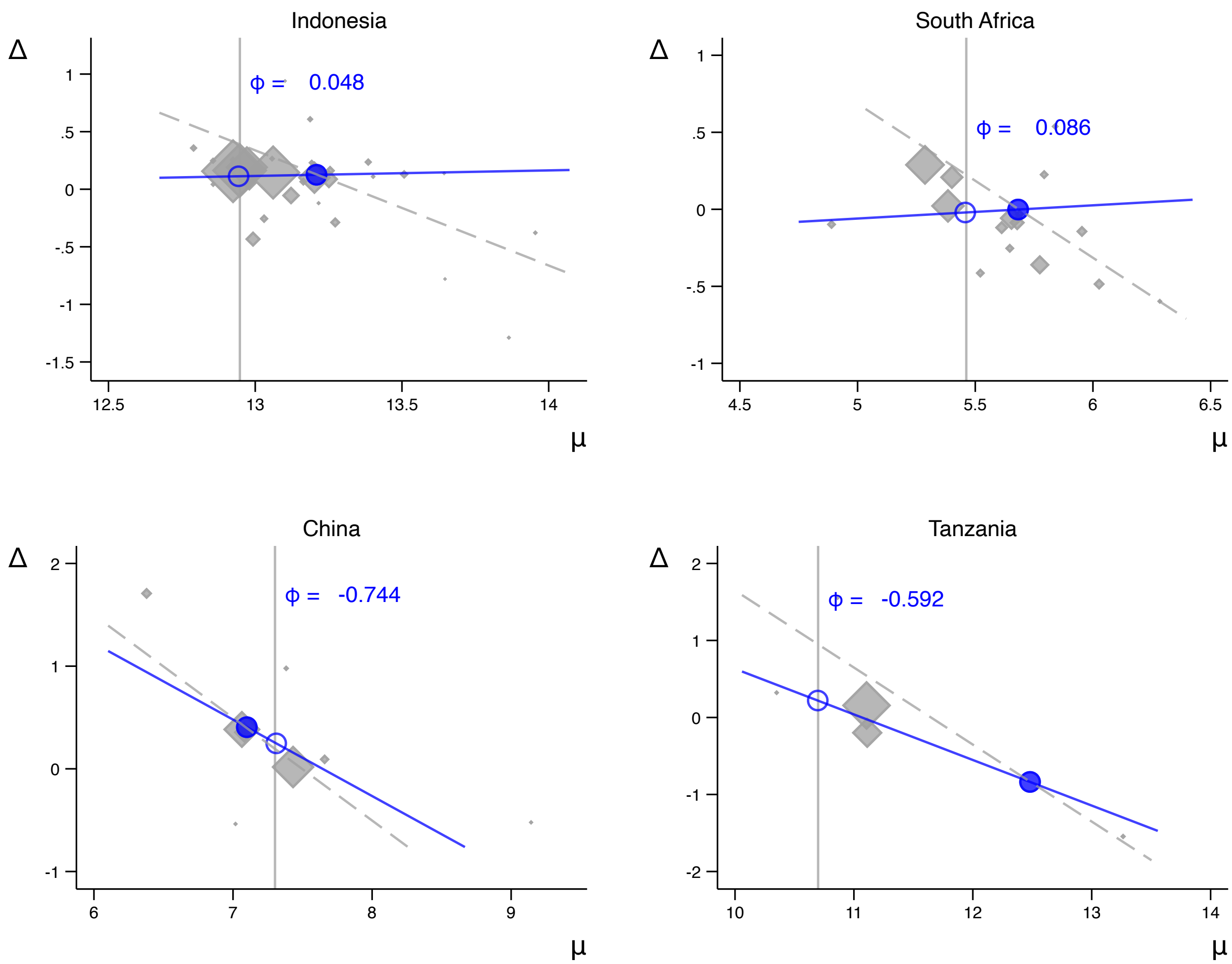
Figure A6: Extrapolation line and GRC estimates of the returns to Non-Agriculture employment on Log Consumption for switchers and non-switchers, Balanced panel





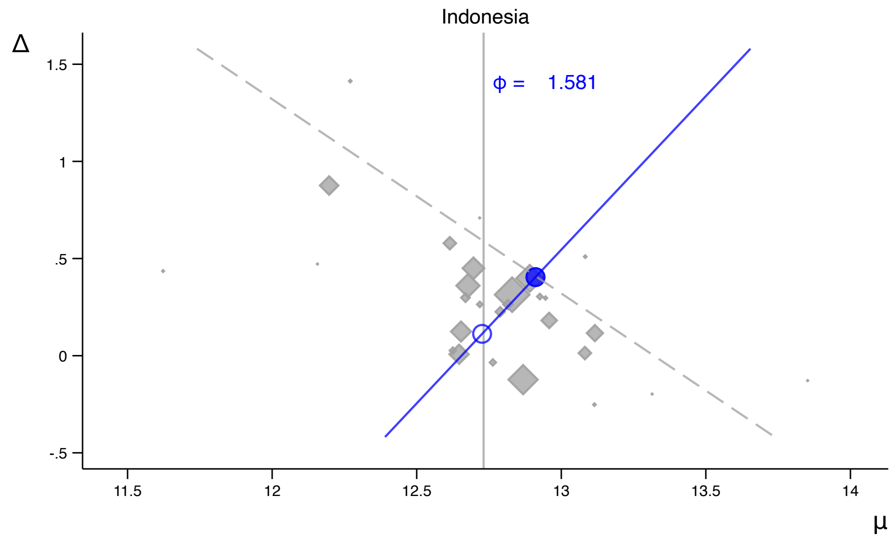
Notes: See notes in Figure 6.

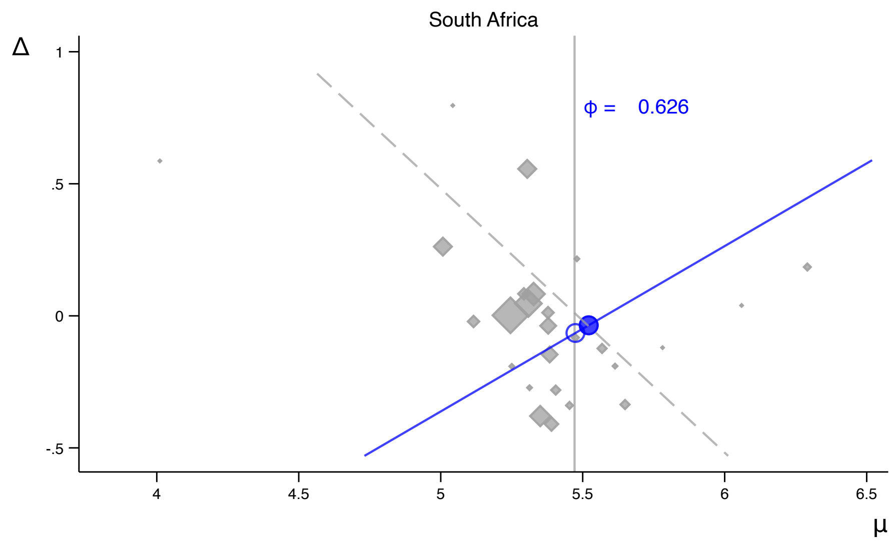
Figure A7: Extrapolation line and GRC estimates of the returns to Urban location on Log Income for switchers and non-switchers



Notes: See notes in Figure 5.

Figure A8: Extrapolation line and GRC estimates of the returns to Non-Agriculture employment on Log Income for switchers and non-switchers





Notes: See notes in Figure 6.