Learning, Selection, and the Misallocation of Households Across Sectors

Jean-Francois Gauthier* Teresa Molina †
Anant Nyshadham‡§

Abstract

We study the role of labor misallocation (i.e., suboptimal sorting of households across sectors) in explaining low productivity in developing countries. We estimate a generalized earnings equation with dynamic correlated random coefficients, allowing households to learn about their relative productivity across sectors. Estimates show that households select into non-farm enterprise on the basis of comparative advantage, but learn and converge slowly over time, with many households spending substantial amounts of time in a suboptimal sector. Roughly 35% of households are misallocated to start, earning nearly 50% less on average than they could have if they were properly sorted across sectors.

JEL Classification Codes: D24, J24, J31, J43, O14, O40

Keywords: sectoral choice, learning, sorting, comparative advantage, misallocation, dynamic correlated random coefficients

---

*Boston College; gauthija@bc.edu; https://sites.google.com/view/jean-francoisgauthier

†University of Hawaii at Manoa; tmolina@hawaii.edu; https://sites.google.com/site/teresasmolina/

‡University of Michigan & NBER; nyshadha@umich.edu; http://www.anantnyshadham.com/

§I would like to especially thank Mike Boozer, Penny Goldberg, and Chris Udry for their invaluable support in the development of the dissertation upon which this paper builds. I would also like to thank Ach Adhvaryu, Joe Altonji, Francesco Amodio, A.V. Chari, Andy Foster, Guido Imbens, Joe Kaboski, Fabian Lange, Costas Meghir, Ted Miguel, Tavneet Suri and participants in workshops and seminars at Yale, Stanford, UC Berkeley, Warwick, USC, UIUC, UC Riverside, Johns Hopkins, BREAD, and NBER for their helpful comments. All errors are our own.

1 Introduction

Productivity is much lower in developing countries than in developed countries (Bloom et al., 2010; Hall and Jones, 1999; Syverson, 2011). A large and growing literature has evaluated several hypothesized drivers of this gap, including managerial quality (Adhvaryu et al., 2019b; Bloom et al., 2013; Bloom and Van Reenen, 2007), trade relationships and costs (Adhvaryu et al., 2019a; Atkin and Donaldson, 2015; Atkin et al., 2017), and resource misallocation across sectors (Hsieh and Klenow, 2009). While this literature focuses mainly on non-agricultural sectors and larger formal firms, a parallel literature has documented that productivity gaps across developed and developing countries are particularly large in the agricultural sector (Gollin et al., 2014; Restuccia et al., 2008). Some studies in this latter literature have also hypothesized a role for misallocation of capital and land (Adamopoulos et al., 2017; Restuccia and Rogerson, 2013); while others have investigated the degree to which the self-selection of households across sectors might also contribute (Adamopoulos et al., 2017; Alvarez-Cuadrado et al., 2019; Lagakos and Waugh, 2013).

In this paper, we ask if the misallocation of households across sectors contributes to low productivity in developing countries across both agriculture and non-agricultural sectors. We hypothesize that imperfect information about relative productivity might lead developing country households to select suboptimally across sectors early on in their productive life cycles. Previous studies have modeled selection as a one-off sorting decision across sectors, limiting the ability to document sectoral sorting mistakes along households' productive life cycles. That is, these analyses can document sectoral sorting for a population at a given point in time, but cannot comment on whether this particular sorting decision is optimal for each household. To the degree that households converge to optimal sectoral choices over time as they learn about which sector best suits their skills, a dynamic approach is required to identify: i) for which sector each household ultimately appears best suited, ii) whether and for how long each household participates in an ill-matched sector, and iii) how much their earnings suffer along the way as a result.

We adapt the dynamic sectoral sorting framework in Gibbons et al. (2005) to the developing country household's decision to engage in non-farm enterprise. This model of selection in which households learn about their relative productivity across sectors yields a generalized earnings equation with dynamic correlated random coefficients (DCRC). We use an extension of projection-based panel methods (Chamberlain, 1982, 1984; Islam, 1995; Suri, 2011) to estimate the model on the national panel sample from the Indonesia Family Life Survey (IFLS) spanning more than two decades.[1] We analytically link the interpretation of our structural estimates to the seminal formulation of the Roy (1951) model in Borjas (1987), which allows us to use our estimates to characterize the nature of sorting in our context as either positive selection, negative selection, or sorting on comparative advantage.

Results show that households select into non-farm enterprise on the basis of comparative advantage, consistent with results from other recent studies (Adamopoulos et al., 2017; Lagakos and Waugh, 2013). We document substantial heterogeneity in the returns to engaging in non-farm enterprise. While the average annual return is roughly 1.5 million rupiah (100 USD), the expected return among households who actually switch in or stay in non-farm enterprise is 4 to 6 times as large and the returns for households who switch out or stay out are negative.

We also document substantial churning along the sectoral margin, an empirical regularity across most developing countries that only a few papers have studied (Adhvaryu et al., 2020; Adhvaryu and Nyshadham, 2017; Calderon et al., 2020). We find that this churning is at least in part a result of substantial learning and slow convergence such that many households spend substantial amounts of time in a sector which is suboptimal for them. At the start of the sample, roughly 35% of households are misallocated, and these households are earning almost 50% less on average than they could have if they were properly sorted across sectors. After 14 years, more than 25% of households (and not

[1]The fundamentals of this approach to panel data are reviewed in Crépon and Mairesse (2008). We discuss later when we develop the methodology how we draw from extensions developed in Islam (1995) to allow for dynamics and Suri (2011) to allow for selection on comparative advantage.

necessarily the same households) remain misallocated, sorting on persistently imprecise perceptions of relative productivity.

We recover structural estimates of both the latent relative ability across sectors and the household's evolving perceptions regarding it over time. We document that returns to non-farm enterprise are higher for households with members exhibiting higher cognitive ability, better physical health, lower risk aversion, as well as more open-mindedness, conscientiousness, and extraversion. However, the full set of observable covariates still only explain 2% of the variation in returns across sectors, consistent with the observed prevalence and persistence of suboptimal sorting decisions.

Our study contributes to two strands of the literature on the causes of low productivity in developing countries (Bloom et al., 2010; Hall and Jones, 1999; Syverson, 2011). Several papers have investigated the role of the misallocation of capital and other non-labor inputs (Adamopoulos et al., 2017; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). The misallocation of labor across sectors has also been hypothesized when documenting productivity gaps across sectors (Gollin et al., 2014). We expand on this work by quantifying the degree of labor misallocation and identifying learning as a cause– leveraging a long panel to document in which sector each household's earnings are maximized and how often they deviate from this optimal sector. In this sense our paper is closest to the recent work by Adamopoulos et al. (2017) showing in China that labor selection reinforces the negative productivity effects of land and capital misallocation across sectors. We complement this work by documenting that labor selection can be imperfect, leading to substantial and costly misallocation of labor as well.

In doing so, we also build on evidence of the sorting of households across sectors (Alvarez-Cuadrado et al., 2019; Lagakos and Waugh, 2013). We find strong evidence that households sort into non-farm enterprise on the basis of perceived comparative advantage, but extend the approaches in previous papers to assess whether a household's sorting decision is optimal in each period. Static approaches interpret realized sorting as revealed preference; whereas our dynamic correlated random coefficients model allows for households to have imperfect information and make mistakes along the way as a result. This flexibility allows us to fit the observed sectoral churning in the data, common across contexts but often overlooked in empirical analyses of sorting.

Our paper relates to the recent work by Hicks et al. (2017), which evaluates relative productivities of workers across sectors using, in part, our same longitudinal data. We complement this evidence by studying household returns to non-farm enterprise rather than the relative earnings of individual workers across sectors, as well as by extending their fixed effects approach to allow for dynamic correlated random coefficients. Our model nests both the fixed effects approach and a model of sorting without learning. This allows us to test and reject the ability of these simpler frameworks to match the patterns in the data, which validates the importance of learning and misallocation in household sectoral choice.

The remainder of the paper is organized as follows. In Section 2, we present the data and motivate our research design with empirical facts. Then, we introduce our model and estimation strategy in Section 3. We present the results of our analysis in Section 4 and conclude in Section 5.

2 Data and Motivation

2.1 IFLS

We use the Indonesian Family Life Survey (IFLS), a longitudinal household survey that began in 1993, with four follow-ups conducted in 1997, 2000, 2007, and 2014 (Strauss et al., 2016) . The sample is representative of the 13 provinces that were selected to be included in the first survey wave (corresponding to over 80% of the Indonesian population). The IFLS collected detailed information about a wide array of household and individual characteristics, including basic demographics, educational attainment, physical health, cognitive ability, risk aversion, and most importantly for this paper, business ownership and income from various sources. Specifically, the main respondent for each household is asked about the household's ownership of and income from household enterprise (both farm and non-farm), and each household member aged 15 or older is asked to report their individual wage income.

We are interested in total annual household income, which we calculate as the sum of profits from non-farm enterprise, profits from farm enterprise (both of which can be negative or positive), and all household members' wage income.[2] After this, we restrict to households with non-missing non-farm enterprise profits, farm enterprise profits, and wage income in all five waves. This leaves us with 3217 households in a balanced panel sample.

This paper focuses on the household-level decision to sort into non-farm enterprise. We therefore generate an indicator equal to one for households who report owning a non-farm enterprise. Over the five survey waves, between 29% to 42% of households owned a non-farm enterprise (reported in Table 1). For brevity, we occasionally refer to these households as "enterprise" households and all other households as "non-enterprise households," but it should be noted that the businesses owned by our "enterprise" households are strictly non-farm.

In Table 1, we also report total annual household income in millions of 2015 Indonesian rupiahs. In 1993, average household income was approximately 7 million rupiahs (around 485 USD), but by 2014, this increased to approximately 19 million.

2.2 Preliminary Evidence

Basic descriptive exercises reveal significant churning in and out of enterprise. In Figure 1, we illustrate the share of households in enterprise and non-enterprise, with five shades of red that represent enterprise households and five shades of blue that represent non-enterprise households. The darkness of a color indicates the number of times a household has switched. In

[2] Given the importance of this income variable for our analysis, we first drop outliers in each wave (specifically, the top 5% and bottom 5% of the income distribution), which we suspect suffer from reporting errors – a common method for trimming self-reported incomes. We show, however, that our results are not sensitive to this choice, demonstrating robustness of main results to winsorizing the top and bottom 1% rather than trimming as an alternative in the appendix.

Table 1: Summary Statistics

			Year		
	1993	**1997**	**2000**	**2007**	**2014**
Share in Enterprise	0.29	0.31	0.42	0.40	0.36
	(0.46)	(0.46)	(0.49)	(0.49)	(0.48)
Total Household Income	6.71	8.78	10.5	13.7	18.9
	(7.89)	(9.61)	(10.3)	(13.9)	(20.5)
Household Size	4.66	4.54	4.54	4.08	3.79
	(2.00)	(1.89)	(1.92)	(1.86)	(1.88)
No. Females Aged 15-59	1.36	1.37	1.38	1.30	1.21
	(0.80)	(0.78)	(0.82)	(0.84)	(0.85)
No. Males Aged 15-59	1.25	1.24	1.29	1.23	1.10
	(0.86)	(0.87)	(0.92)	(0.93)	(0.91)

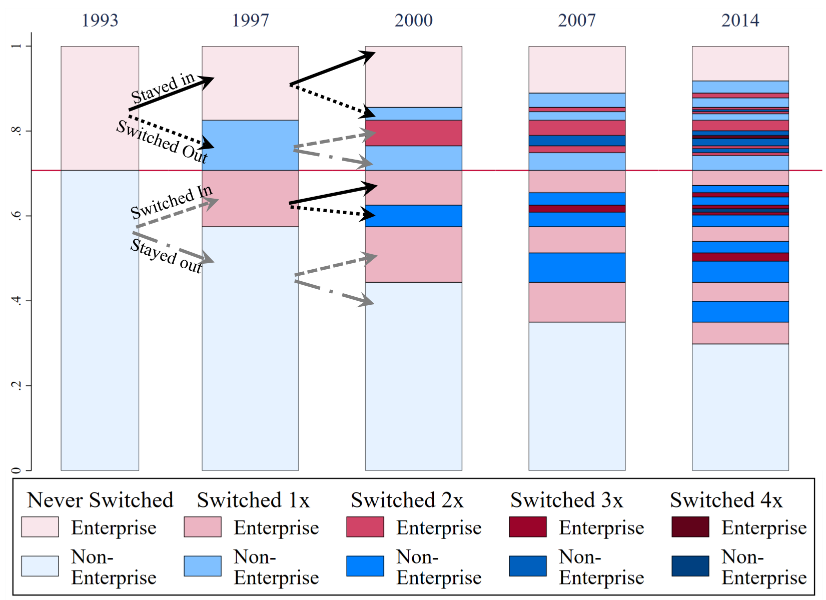
Notes: N=3217. Sample consists of IFLS households with non-missing income information in all five waves of the IFLS.

1993, when we do not have any previous information on enterprise status, all households have never switched according to our data and are therefore represented by the lightest shades of red (for those currently in enterprise) and blue (for those currently in non-enterprise). In 1997, however, about 40% of the households who were in enterprise in 1993 switched out of enterprise in 1997 (represented by a slightly darker shade of blue because they switched once). At the same time, close to 20% of the 1993 non-enterprise households switched into enterprise in 1997 (represented by a slightly darker shade of red).

This switching behavior continues across the remaining 3 waves. By 2014, it is clear that over half of households have switched at least once (any color that is not the lightest red or blue represents a household that has switched). There are many households that have switched more than once, and even some that have switched four times. In short, switching sectors is common.

We next ask whether switching ever slows down. That is, does switching decline with the amount of time a household spends in a sector? Figure 2 shows that it does. Among households that have been in their current sector

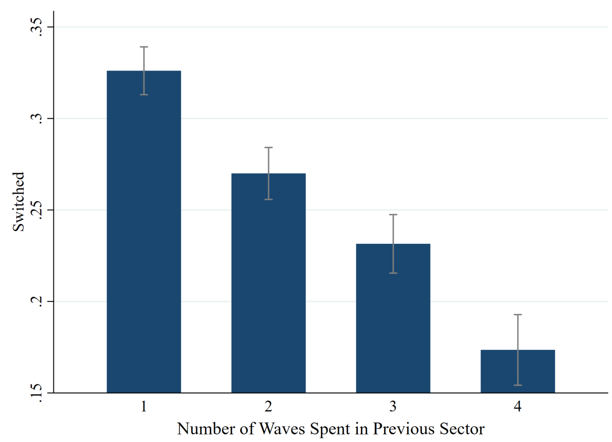
Figure 1: Churning Across Sectors Over Time



Notes: Sample consists of IFLS households with non-missing income information in all five waves of the IFLS. Shades of red represent households that are in non-farm enterprise in the relevant wave, while shades of blue represent households that are not. Color darkness captures the number of times a household has switched prior to that wave.

for only one wave (starting from when we first observe them in 1993), over 30% of households switched sectors. This share drops with the cumulative number of waves spent in the previous sector. This suggests that, though sectoral switching is common, households' switching decisions appear to exhibit convergence, such that longer time spent in a given sector yields a lower probability of switching out. In the appendix, we show this pattern holds in both directions (i.e., for both enterprise and non-enterprise households (see Figure A1)). The patterns depicted in these figures motivate the model we develop in the next section.

Figure 2: Switching by Number of Waves Spent in Previous Sector



Notes: Sample consists of IFLS households with non-missing income information in all five waves of the IFLS. Error bars denote 95% confidence intervals.

3 Model

3.1 Sectoral Choice

In this section, we outline a Roy (1951) model of sectoral choice, where household i in period t chooses whether to go into enterprise (denoted by superscript E) or stay in the default sector (denoted by superscript N). Sector-specific income Y_{it} is determined by the following equations:

$$
\begin{aligned}
Y_{it}^{E} &= \beta_t^{E} + \eta_i^{E} \\
Y_{it}^{N} &= \beta_t^{N} + \eta_i^{N}.
\end{aligned} \tag{1}
$$

β_t^E is average income in the enterprise sector and β_t^N is average income in the non-enterprise sector. η_i^E is the unobserved, heterogeneous component of enterprise-specific productivity, while η_i^N is the corresponding component for the non-enterprise sector.

We can rewrite both η_i^E and η_i^N as a function of relative productivity $(\eta_i^E - \eta_i^N)$, and absolute advantage, τ_i, which we define as the component of the household-specific productivity that has the same effect on the household's productivity in both sectors. (Accordingly, τ_i does not affect the sectoral choice.) Specifically, we rewrite each sector-specific productivity term in the following way:

$$
\begin{aligned}
\eta_i^{E} &= (1+\phi)\eta_i + \tau_i \\
\eta_i^{N} &= \eta_i + \tau_i,
\end{aligned} \tag{2}
$$

where both ϕ and η_i depend on projection coefficients, b_N and b_E.[3] We define $\phi \equiv b_E/b_N - 1$, and $\eta_i \equiv b_N(\eta_i^E - \eta_i^N)$.

The equations in (2) show that a household's sector-specific productivity is a function of both relative productivity and absolute advantage. Importantly, the parameter ϕ depends on the covariance between enterprise and non-enterprise productivity in the population as a whole, $Cov(\eta_i^E, \eta_i^N)$, and therefore, summarizes the nature of sorting in the population.

[3]Since with 2 sectors only the relative magnitude of η_i^N and η_i^E can be identified, we will define, following Lemieux (1998) and Suri (2011), η_i^N and η_i^E in terms of the household's relative productivity in enterprise over non-enterprise activity $(\eta_i^E - \eta_i^N)$ using the following projections: $\eta_i^N = b_N(\eta_i^E - \eta_i^N) + \tau_i$ and $\eta_i^E = b_E(\eta_i^E - \eta_i^N) + \tau_i$, where $b_E = (\sigma_E^2 - \sigma_{EF})/(\sigma_E^2 + \sigma_N^2 - 2\sigma_{EF})$, $b_N = (\sigma_{EF} - \sigma_N^2)/(\sigma_E^2 + \sigma_N^2 - 2\sigma_{EF})$, with $\sigma_{EF} \equiv Cov(\eta_i^E, \eta_i^N)$, $\sigma_E^2 \equiv Var(\eta_i^E)$, and $\sigma_N^2 \equiv Var(\eta_i^N)$.

As we discuss in the appendix (section B), positive ϕ means that there is positive selection (using the terminology used in Borjas (1987)): households who are productive in enterprise tend to also be productive in non-enterprise, and the more productive households select into enterprise because the variance of productivity is higher in this sector. On the other hand, when $-1 < \phi < 0$, we have negative selection: like in the previous case, households who are productive in enterprise tend to also be productive in non-enterprise, but the more productive households select into non-enterprise because of the higher variance in the non-enterprise sector. Finally, when $\phi < -1$, we have selection on comparative advantage: households who are productive in the enterprise sector tend to be less productive in the non-enterprise sector and therefore select into enterprise (and vice versa).

Let D_{it} represent a dummy equal to one for households in the enterprise sector at time t. Combining equations (1) and (2), we arrive at the following generalized income equation:

$$Y_{it} = \alpha_t + \beta D_{it} + \eta_i(1 + \phi D_{it}) + \tau_i, \tag{3}$$

where $\alpha_t \equiv \beta_t^N$ and $\beta \equiv (\beta_t^E - \beta_t^N)$, which we assume to be constant over time.[4] Estimation of the parameters β and ϕ is complicated by the fact that D_{it} is endogenous. Households will choose $D_{it} = 1$ if they expect higher earnings in the enterprise sector (that is, if $\phi\eta_i > -\beta$) . In the next subsection, we discuss what households know about their own η_i, and how this knowledge evolves over time.

[4] As we discuss later, when we estimate the model we will explicitly purge all outcome variables and regressors of variation in means across communities and within communities over time, using community fixed effects that vary across time periods (essentially, community-by-time dummies). These fixed effects will account for changes in relative output prices across sectors, as long as relative prices do not vary within a community in a single year. Under these conditions, extending the analysis to estimate a time-varying β seems of little empirical benefit.

3.2 Learning

We assume that households know the population average earning in both sectors (α_t, β), their own absolute advantage (τ_i), and ϕ, but have imperfect information about their comparative advantage (η_i).[5] In particular, we introduce an additive productivity shock, ε_{it}, to η_i in equation (3) and assume that $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2 = 1/h_\varepsilon)$. That is, the household only observes the sum of η_i and ε_{it}, but not either individually. The generalized income equation then becomes:

$$Y_{it} = \alpha_t + \beta D_{it} + (\eta_i + \varepsilon_{it})(1 + \phi D_{it}) + \tau_i \tag{4}$$

Households hold the initial belief that $\eta_i \sim N(m_{i0}, \sigma^2 = 1/h)$; and this belief is refined each period using output observations, Y_{it}. That is, from Y_{it}, households can compute

$$l_{it} = \frac{Y_{it} - \alpha_t - \beta D_{it} - \tau_i}{(1 + \phi D_{it})} = \eta_i + \varepsilon_{it}, \tag{5}$$

a noisy signal of their relative productivity η_i, which is independent of the their period t sectoral choice. Let $l_i^t = (l_{i1}, ..., l_{it})$ denote the history of household i's normalized relative productivity observations through period t. Then, the posterior distribution of η_i given history l_i^t is distributed $N(m_t(l_i^t), 1/h_t)$, where

$$m_t(l_i^t) = \frac{hm_{i0} + h_\varepsilon(l_{i1} + ... + l_{it})}{h + th_\varepsilon}, \qquad \text{and} \qquad h_t = h + th_\varepsilon \tag{6}$$

Note that the specific learning mechanism proposed here allows households to learn about returns to enterprise each period, irrespective of the sector the household has chosen that period. This learning structure is borrowed from Gibbons et al. (2005) who use it to study learning about comparative advantage in a model of occupational choice.[6] The bidirectional churning and

[5] As we explain below, ϕ can be thought of as the value of skills in each sector, where the skills are captured by the comparative advantage component.

[6] They, in turn, borrow heavily from the classic development in DeGroot (1970). Please see these previous works for more in depth discussion of this framework.

convergence observed in the raw data motivates the use of this approach in our setting (see Figures 1 and A1).

The intuition behind this proposed mechanism is that relative productivity, η_i, is an index of fundamental skills which affect productivity in both sectors, but is valued differentially across the two sectors (e.g., managerial skill). Assuming that the household knows ϕ but not η_i corresponds to assuming the household knows how much each sector values these skills but not their own skill stock. Accordingly, households can learn about their stock through production in either sector.

For example, suppose that η_i represents the household's managerial skill. The default sector rewards managerial skill in its relation to input inventory management and efficient resource allocation. However, the non-farm enterprise sector, corresponding to the household's running a retail shop for example, rewards managerial ability more heavily. Studies on small and medium enterprises in developing countries have emphasized the relative importance of such skill in determining enterprise productivity as well as extensive margin enterprise participation decisions (Bloom et al., 2013; Bruhn et al., 2010; Calderon et al., 2020). Accordingly, enterprise earnings relative to non-enterprise earnings are increasing in managerial skill. The assumptions of the model imply that the household recognizes that enterprise rewards managerial ability more than non-enterprise does; however, the household is unsure of its specific stock of managerial skill.

Of course, an excellent manager might still be able to earn more in the non-enterprise sector than someone with the same access to resources but worse skill in allocating those resources. Therefore, a household that initially believes it is bad at management will operate in the non-enterprise sector to start, where this lack of managerial skill is less penalized; however, should this household find this period that it is better able to manage its agricultural inputs (for example) than it expected, it will decide to open a retail shop next period, knowing that the retail business is very lucrative for a household with strong managerial ability. The mechanism, of course, works in the opposite direction as well. We should note that, to the degree that both sectors reward some skills (e.g., work ethic) *equally*, these skills are represented by τ_i and will affect household income in both sectors, but will not affect the return to switching sectors.

Household i will choose the enterprise sector in period t if $E[Y_{it}^E - Y_{it}^N] > 0$, and choose the non-enterprise sector otherwise. That is, household i will choose the enterprise sector in period t (i.e., $D_{it} = 1$) if and only if $\phi m_i^{t-1} > -\beta$.

3.3 Estimation

Allowing for measurement error in equation (4), our estimating equation is the following:

$$Y_{it} = \alpha_t + \beta D_{it} + (\eta_i + \varepsilon_{it})(1 + \phi D_{it}) + \tau_i + \zeta_{it} \tag{7}$$

where measurement error ζ_{it} is assumed mean independent of sector and input decisions conditional on η_i and τ_i. That is, in particular, we will assume $E(D_{it}|\zeta_{it}, \eta_i, \tau_i) = E(D_{it}|\eta_i, \tau_i)$.

As discussed above, D_{it} will depend on the mean of the household's prior distribution on η_i coming into period t, $m_{i,t-1}$, which we cannot observe. Accordingly, OLS estimates of β will be biased. We now develop a strategy which allows us to consistently estimate β, recover ϕ, and validate the importance of learning dynamics in this empirical context.

In particular, in order to recover consistent estimates of β, we must purge the composite unobserved term, $(\eta_i + \varepsilon_{it})(1 + \phi D_{it}) + \tau_i + \zeta_{it}$, of its correlation with D_{it}. We know from section 3.2 that the portion of $(\eta_i + \varepsilon_{it})$ which correlates with sectoral choices is $m_{i,t-1}$. We will begin by decomposing $m_{i,t-1}$ into two components which have distinct effects on the household's history of sectoral choices. Note that the Bayesian updating of beliefs implies that the mean of the prior distribution is a martingale. That is, the law of motion for

$m_{i,t}$ is

$$m_{i,t} = m_{i,t-1} + \xi_{it} \quad \Rightarrow \quad m_{i,t-1} = m_{i0} + \sum_{k=1}^{t-1} \xi_{ik}, \tag{8}$$

where ξ_{it} is a noise term orthogonal to $m_{i,t-1}$. Then, denoting $\tilde{m}_{i,t-1} \equiv \sum_{k=1}^{t-1} \xi_{ik}$ as the sum of the signals received up to period $t-1$, we have

$$Y_{it} = \alpha_t + \beta D_{it} + (m_{i0} + \tilde{m}_{i,t-1} + \omega_{it})(1 + \phi D_{it}) + v_{it}, \tag{9}$$

where $v_{it} \equiv \tau_i + \zeta_{it}$ is orthogonal to sectoral choice in period t, D_{it}, by construction and $\omega_{it} \equiv \eta_i + \varepsilon_{it} - (m_{i0} + \tilde{m}_{i,t-1})$ is orthogonal to D_{it} by nature of the martingale structure of $m_{i,t-1}$.

Extending the approaches developed by Chamberlain (1982, 1984), Islam (1995), and Suri (2011), we can overcome the endogeneity of D_{it} by projecting m_{i0} and $\tilde{m}_{i,t-1}$ onto the history of sectoral choices. In particular, the law of motion of the prior, as expressed in equation (8), suggests that the initial belief, m_{i0}, will affect sectoral choices in all periods. On the other hand, the cumulative update, $\tilde{m}_{i,t-1}$, will only affect sectoral choices in period t onwards.

We have five waves of data and therefore four cumulative updates. The projection of the initial belief, m_{i0}, which appears in the estimating equation for all periods, will include the entire history of sectoral choices as follows:[7]

$$m_{i0} = \lambda_0 + \prod_{k=1}^{5} (1 + \lambda_k D_{ik}) - 1 + \psi_{i0} \tag{10}$$

where ψ_{it} is projection error in period t. The projection of each cumulative

[7]If we expand m_0, we get: $m_0 = \lambda_0 + \lambda_1 D_1 + \lambda_2 D_2 + \lambda_3 D_3 + \lambda_4 D_4 + \lambda_5 D_5 + \lambda_{12} D_1 D_2 + \lambda_{13} D_1 D_3 + \lambda_{14} D_1 D_4 + \lambda_{15} D_1 D_5 + \lambda_{23} D_2 D_3 + \lambda_{24} D_2 D_4 + \lambda_{25} D_2 D_5 + \lambda_{34} D_3 D_4 + \lambda_{35} D_3 D_5 + \lambda_{45} D_4 D_5 + \lambda_{123} D_1 D_2 D_3 + \lambda_{124} D_1 D_2 D_4 + \lambda_{125} D_1 D_2 D_5 + \lambda_{134} D_1 D_3 D_4 + \lambda_{135} D_1 D_3 D_5 + \lambda_{145} D_1 D_4 D_5 + \lambda_{234} D_2 D_3 D_4 + \lambda_{235} D_2 D_3 D_5 + \lambda_{245} D_2 D_4 D_5 + \lambda_{345} D_3 D_4 D_5 + \lambda_{1234} D_1 D_2 D_3 D_4 + \lambda_{1235} D_1 D_2 D_3 D_5 + \lambda_{1245} D_1 D_2 D_4 D_5 + \lambda_{1345} D_1 D_3 D_4 D_5 + \lambda_{2345} D_2 D_3 D_4 D_5 + \lambda_{12345} D_1 D_2 D_3 D_4 D_5 + \psi_{i0}$, where $\lambda_{ijklm} = \lambda_i \lambda_j \lambda_k \lambda_l \lambda_m$.

update, $\tilde{m}_{it}$, includes only the sectoral choices in $t+1$ and onward:

$$\begin{aligned}
\tilde{m}_{i1} &= \theta_{20} + \theta_{22} D_{i2} + \theta_{23} D_{i3} + \theta_{24} D_{i4} + \theta_{25} D_{i5} + \psi_{i1} \\
\tilde{m}_{i2} &= \theta_{30} + \theta_{33} D_{i3} + \theta_{34} D_{i4} + \theta_{35} D_{i5} + \psi_{i2} \\
\tilde{m}_{i3} &= \theta_{40} + \theta_{44} D_{i4} + \theta_{45} D_{i5} + \psi_{i3} \\
\tilde{m}_{i4} &= \theta_{50} + \theta_{55} D_{i5} + \psi_{i4}.
\end{aligned} \tag{11}$$

Note that the martingale structure of the prior on η_i implies that learning is *efficient*; that is, all information the household will use to make its decision at time t is fully summarized in the initial condition m_{i0} and the sum of the orthogonal updates to period $t-1$, $\tilde{m}_{i,t-1}$. In other words, the path by which the prior reaches $m_{i,t-1}$ will not, conditional on $m_{i,t-1}$ itself, affect sectoral choice in period t, D_{it}. Most importantly, the path by which the sum of the updates reaches $\tilde{m}_{i,t-1}$ will not, conditional on both the initial belief m_{i0} and $\tilde{m}_{i,t-1}$ itself, affect D_{it}. Therefore, we need not include past sectoral choices nor the interactions of future sectoral choices in the update projections in (11).

Note also that the relative sizes of h and h_ϵ will determine the degree to which the initial condition, m_{i0}, or subsequent updates, $\tilde{m}_{i,t-1}$, correlate more strongly with choices across periods. We do not explicitly discuss this relationship further as the estimation will approach this issue agnostically. That is, the estimation will allow the data to show (in the projection coefficients) the degree to which initial conditions and subsequent updates affect choices without restricting *a priori* the relative magnitudes of these correlations. If, for example, a large dispersion in the initial conditions effectively makes their impact on production decisions negligible, the coefficients in equation (10) will be estimated as indistinguishable from 0, while those from the equations in (11) might be estimated with larger magnitudes and more precision.

Plugging projections (10) and (11) into equation (9), and grouping terms, we can now express each Y_t as a function of all sectoral choices (see equation (15) in Appendix section C).[8] This results in the following reduced form

[8]It is important that we properly specify the projections in (10) and (11). That is, we must include all necessary elements of the history of sectoral choices in order to ensure that

regressions, where income in each period depends on all five D_{it} as well as their double, triple, quadruple, and quintuple interactions:

$$Y_{it} \quad = \gamma_0^t + \prod_{k=1}^{5}(1 + \gamma_k^t D_{ik}) - 1 + \nu_{it}. \tag{12}$$

If we define $\gamma_{ijklm}^t \equiv \gamma_i^t \gamma_j^t \gamma_k^t \gamma_l^t \gamma_m^t$, each equation has 32 reduced form coefficients to be estimated.[9] Following Chamberlain (1982, 1984), we will first estimate these reduced form coefficients by seemingly unrelated regressions (SUR) and then estimate from these coefficients the structural parameters of the model using minimum distance. After normalizing each of the intercepts in equations (10), (11), and (12),[10] there are 43 structural parameters of the model (31 λ coefficients, 10 θ coefficients, β, and ϕ), to be identified from the 155 reduced form coefficients using the minimum distance restrictions implied by the model. The minimum distance restrictions are reported in Appendix section C.2. To help with the exposition of the strategy, we expand all equations for the two-period case in Appendix section D.

For simplicity, we have not included any covariates in the exposition above,

the projection errors (ψ) are, indeed, orthogonal to current choices.

[9]Expanding, we obtain: $Y_{it} = \gamma_0^t + \gamma_1^t D_1 + \gamma_2^t D_2 + \gamma_3^t D_3 + \gamma_4^t D_4 + \gamma_5^t D_5 + \gamma_{12}^t D_1 D_2 + \gamma_{13}^t D_1 D_3 + \gamma_{14}^t D_1 D_4 + \gamma_{15}^t D_1 D_5 + \gamma_{23}^t D_2 D_3 + \gamma_{24}^t D_2 D_4 + \gamma_{25}^t D_2 D_5 + \gamma_{34}^t D_3 D_4 + \gamma_{35}^t D_3 D_5 + \gamma_{45}^t D_4 D_5 + \gamma_{123}^t D_1 D_2 D_3 + \gamma_{124}^t D_1 D_2 D_4 + \gamma_{125}^t D_1 D_2 D_5 + \gamma_{134}^t D_1 D_3 D_4 + \gamma_{135}^t D_1 D_3 D_5 + \gamma_{145}^t D_1 D_4 D_5 + \gamma_{234}^t D_2 D_3 D_4 + \gamma_{235}^t D_2 D_3 D_5 + \gamma_{245}^t D_2 D_4 D_5 + \gamma_{345}^t D_3 D_4 D_5 + \gamma_{1234}^t D_1 D_2 D_3 D_4 + \gamma_{1235}^t D_1 D_2 D_3 D_5 + \gamma_{1245}^t D_1 D_2 D_4 D_5 + \gamma_{1345}^t D_1 D_3 D_4 D_5 + \gamma_{2345}^t D_2 D_3 D_4 D_5 + \gamma_{12345}^t D_1 D_2 D_3 D_4 D_5 + \psi_{i0}$, where $\gamma_{ijklm}^t = \gamma_i^t \gamma_j^t \gamma_k^t \gamma_l^t \gamma_m^t$.

[10]We normalize the intercepts such that the estimates of the projection coefficients are mean zero, as follows:

$$\begin{aligned}
\lambda_0 &= 1 - \prod_{t=1}^{5}(1 + \lambda_t \bar{D}_t) \\
\theta_{20} &= -\theta_{22}\bar{D}_2 - \theta_{23}\bar{D}_3 - \theta_{24}\bar{D}_4 - \theta_{25}\bar{D}_5 \\
\theta_{30} &= -\theta_{33}\bar{D}_3 - \theta_{34}\bar{D}_4 - \theta_{35}\bar{D}_5 \\
\theta_{40} &= -\theta_{44}\bar{D}_4 - \theta_{45}\bar{D}_5 \\
\theta_{50} &= -\theta_{55}\bar{D}_5,
\end{aligned}$$

where $\bar{D}_t$ is the sample mean of the enterprise dummy in period t. An analogous exercise is conducted for the reduced form regressions in (12).

although one could argue that there are household-level characteristics which are correlated with household income and also sectoral choice D_{it}. Though the inclusion of covariates will affect reduced form expressions (12), it will not affect the relationships between the reduced form coefficients on the choices and the structural parameters of interest. We control for community fixed effects and household composition variables (number of household members, number of women aged 15-59, and number of men aged 15-59) in each equation of the first stage SUR estimation. Note that by allowing each community effect to vary across waves, we are also able to account for local community-level demand shocks and price fluctuations that may affect switching decisions but do not convey any information about household-level perceptions of relative ability across sectors.

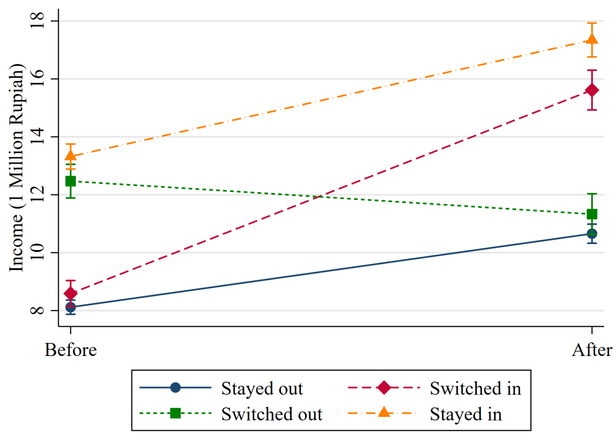
3.4 Identification

Note that identification of the structural parameters, such as β, ϕ, the λ's and θ's, comes from a comparison of the income evolutions across households with different sectoral choice histories. That is, we observe in the data the conditional sample mean of income for each enterprise history in each period (i.e. $E(Y_{it}|D_{i1}, D_{i2}, D_{i3}, D_{i4}, D_{i5})$). The econometric strategy recovers the contribution that each choice in the optimized history of the household makes to the trajectory of household income.

To more clearly convey this intuition, we plot in Figure 3 the evolution of realized incomes for four groups of households: period t enterprise households who stay in enterprise in $t+1$, period t enterprise households who switch out of enterprise in $t+1$, period t non-enterprise households who switch into enterprise in $t+1$, and period t non-enterprise households who stay out of enterprise in $t+1$. To generate this figure, we include all transitions between waves (such that each household appears multiple times, potentially in different groups), and calculate average income across all households in each group, in the "before" period (t) and the "after" period ($t + 1$).

The identification comes from comparing across households with marginally

Figure 3: Income by Switch Status



Notes: This figure treats each household transition as a separate observation, which means that each household has four observations (one for each transition: 1993-1997, 1997-2000, 2000-2007, and 2007-2014). “Stayed out” includes households in non-enterprise in both t and $t + 1$. “Switched In” includes households in non-enterprise in t and enterprise in $t + 1$. “Switched Out” includes households in enterprise in t and non-enterprise in $t + 1$. “Stayed In” includes households in enterprise in both t and $t + 1$. Error bars denote 95% confidence intervals.

different histories (e.g., households that stay in throughout the panel vs. households that switch into the enterprise sector after the first wave and stay in thereafter). The intuition maps roughly to differencing the slopes of the four lines in Figure 3. In a preliminary inspection of Figure 3, we see that households that stay in the enterprise sector have higher income to start than all other types of households. Additionally, households that stay in enterprise have income that seems to grow more steeply than does that of households that stay out in both periods or switch out of the enterprise sector. However, households that switch into the enterprise sector have low income to start but rise most steeply, indicating that these households have large returns to enterprise, while other households may not necessarily have the same returns. Identification in the full model then expands the intuition of this comparison to involve each wave-specific income for each household type denoted by a specific sequence of sectoral choices.

The main identifying assumption is sequential exogeneity: the current period's shock to productivity is assumed to be mean zero, conditional on the prior at the beginning of period. If households can predict future productivity shocks (e.g., good rains next year, infrastructure expansion in the village in the near future, rising demand for a specific good in village) and respond to them in their sector and input decisions, the update projection, as specified, will not fully account for the endogeneity in these choices. Note, however, that these future predictions only matter if they are household-specific as community by time variation is projected off in the first stage. Specifically, there are no λ's and θ's included in the estimation to capture correlations between future idiosyncratic shocks and past household sectoral choices. These correlations are assumed to be zero in order to be able to identify the model with multiple endogenous choices and a small number of periods. Specifically, relaxing this assumption further in a model with heterogeneous returns leads to an incidental parameters problem causing the model to not be fully identified.[11]

[11] Though this paper contributes to the literature on panel data estimators of correlated random coefficients models by relaxing the strict exogeneity assumption to sequential exogeneity to allow for dynamics, we leave it to future work to relax the sequential exogeneity assumption further to allow for correlations of regressors with both past and future shocks.

3.5 Nested Models

The model described above is a dynamic correlated random coefficients (DCRC) model that allows for heterogeneous returns to enterprise and imperfect information. In addition to estimating this preferred model, we also estimate nested models which impose additional restrictions on the relationships between η_i and the endogenous choices, D_{it} . Specifically, we estimate (1) a correlated random coefficients (CRC) model of heterogeneous returns to enterprise with perfect information and (2) a simple fixed effects model with homogeneous returns and perfect information, which is equivalent to a correlated random effects (CRE) model.

In the CRC model, households are assumed to have perfect information about their relative productivity η_i, which means there is no longer an additive productivity shock, ε_{it}, nor any updating of expectations about η_i. As we discuss in Appendix section E.1, this translates into a restricted version of the DCRC model described above, in which all θ coefficients from equation (11) are assumed equal to zero. Models of this sort have been used to study agricultural technology adoption (Suri, 2011) and returns to schooling (Heckman and Vytlacil, 1998).

In the CRE model, in addition to perfect information about η_i, households are assumed to have homogeneous returns. Because a household's return to enterprise no longer depends on their relative productivity η_i, ϕ is assumed to be zero. As we show in Appendix section E.2, these restrictions lead to a version of the model where ϕ, all θ coefficients in equation (11), and all λ coefficients in equation (12) – except for λ_1, λ_2, λ_3, λ_4, λ_5 – are equal to zero. This corresponds to the original model studied by Chamberlain (1984) and has been used recently to study sorting among wage laborers in our same empirical context (Hicks et al., 2017). By estimating these nested models in comparison to the preferred DCRC model, we are able to test whether simpler models are rich enough to capture key features of the data.

4 Results

4.1 Structural Minimum Distance Estimates

In Table 2, we present the minimum distance estimates of β and ϕ. The first column displays estimates from our preferred DCRC model. We estimate an average return to enterprise (β) of approximately 1.5 million rupiah, which is about one-fifth of the average household income in 1993. ϕ is estimated to be -2.5. Significantly less than 1, this estimate implies that households sort based on comparative advantage in this context. That is, households productive in enterprise tend to be less productive in non-enterprise and vice versa.

The DCRC model fits the data well. Table C1 reports individual tests of each of the minimum distance restrictions. We compare the γ coefficients estimated by the SUR to those implied by the structural estimates and minimum distance restrictions (listed in section C.2). Differences are small in magnitude and the vast majority are not statistically different from zero. We also fail to reject a joint test of the minimum distance restrictions at the 5% level.

We next compare our preferred estimates of β and ϕ to those from the two nested models: the CRC model of heterogeneous returns and perfect information, and the CRE model of homogeneous returns and perfect information. Both restricted models substantially over-estimate the average return to enterprise. The CRC model estimates an average return of 2.7 million rupiah (column 2), while the CRE model estimates a return of approximately 3 million rupiah (column 3) – double the magnitude of the DCRC estimate. While the CRE model assumes ϕ to be equal to zero, the CRC model underestimates the magnitude of ϕ relative to our preferred estimate (-1 compared to -2.5).[12]

In short, ignoring heterogeneity in returns and dynamics results in an overestimation of the average return to enterprise and an underestimation of the extent to which households sort based on comparative advantage. Notably, the DCRC model fits the data better than either the CRC or the CRE. When

[12] β^{CRC} and β^{CRE} are statistically larger than β^{DCRC} at the 1% level (the p-values are 0.002 and close to 0 respectively). ϕ^{CRC} is larger than ϕ^{DCRC} at the 1% level (p-value 0.0078).

we conduct a joint test of the minimum distance restrictions imposed by each of the three models, we easily reject the null (with p-values close to zero) for both the CRC and CRE but, as mentioned above, fail to reject the null at the 5% level for the DCRC.

Table 2: Structural Estimates

	Specification		
	(1)	(2)	(3)
	DCRC	CRC	CRE
β	1.498	2.695	3.112
	(0.313)	(0.273)	(0.203)
ϕ	-2.542	-1.076	
	(0.588)	(0.143)	

Notes: Structural parameters estimated using minimum distance. Standard errors in parentheses. Column 1 reports estimates from the full DCRC model (with heterogeneous returns and imperfect information), column 2 reports estimates from the CRC model (with heterogeneous returns and perfect information), and column 3 reports estimates from the CRE model (with homogeneous returns and perfect information).

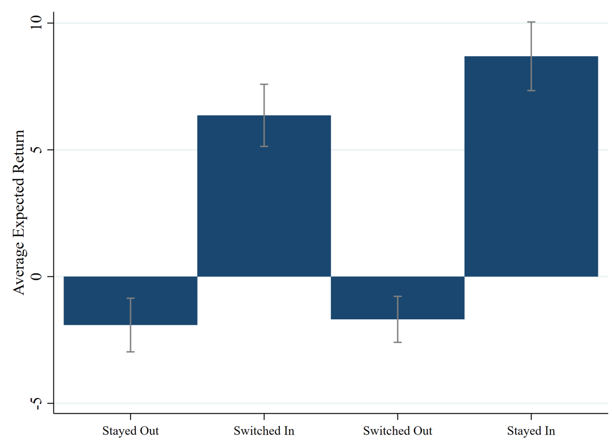
4.2 Expected Returns

We next examine how sorting and switching behavior is governed by a household's expected returns to enterprise. The ability to recover and interpret these patterns is, perhaps, the main strength of our empirical approach. Other approaches to recovering β and even ϕ would not allow for recovery of each household's expected returns at each decision point and analysis of whether these expectations correspond to subsequent choices in ways consistent with the intuition of the model.[13]

[13]Though approaches to estimating DCRC models are quite limited in the literature, instrumental variables approaches, for example, used to estimate CRC models (Heckman and Vytlacil, 1998) would not recover these additional parameters. Even to estimate static heterogeneous returns, it would likely be infeasible to find a rich enough set of instruments across such a large set of household types over such a long panel. That is, one would need instruments that predict switching in both directions across households with different relative abilities across different waves just to recover β and ϕ even in the absence of dynamics. For example, price fluctuations alone would not, in general, be enough.

First, we calculate $\beta + \phi m_{it}$ for each household, for periods $t = 1$ to 4. This represents a household's expected return to enterprise, based on what they have learned up until period t about their relative productivity η_i. In Figure 4, we average these returns across households in four different groups: those who stayed out of enteprise, those who switched into enterprise, those who switched out of enterprise, and those who stayed in enterprise. As expected, returns to enterprise are higher for households in non-enterprise who switch into enterprise compared to those who stay out. Returns are also higher for enterprise households who stay in enterprise compared to those who switch out. Figure A2 in the appendix calculates these returns by wave, and separately for current enterprise households and current non-enterprise households – both groups show similar patterns, consistent with both the patterns in the raw data and the learning structure assumed in the model.

Figure 4: Expected Returns by Switch Status



Notes: The figure reports the average return to enterprise ($\beta + \phi m_{it}$) across $t = 1$ to 4 and all households in each category. "Stayed out" includes households in non-enterprise in both t and $t + 1$. "Switched In" includes households in non-enterprise in t and enterprise in $t + 1$. "Switched Out" includes households in enterprise in t and non-enterprise in $t + 1$. "Stayed In" includes households in enterprise in both t and $t + 1$. Error bars denote 95% confidence intervals. Standard errors are calculated analytically (see Appendix F).

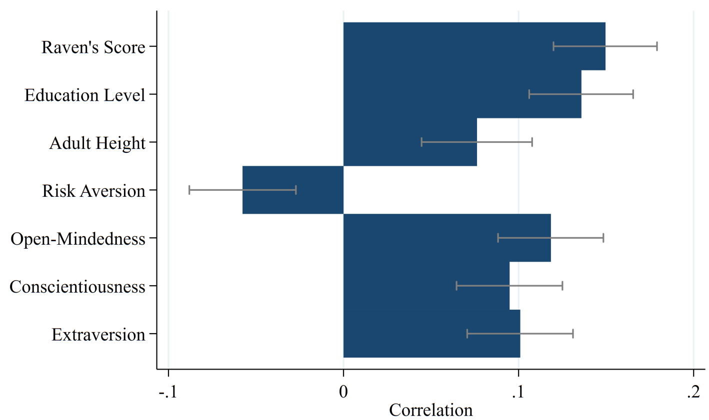
In short, the expected returns estimated by the model are consistent with households' sorting behavior. Note that though the results are fully consistent with the model intuition, the estimated pattern is not mechanical. The estimation strategy does not restrict in any way these recovered correlations between income evolutions and the sequence of choices. It would be entirely possible to recover estimates suggesting that future earnings evolutions for households who switch into enterprise are in fact negative, as might be the case if non-farm enterprise were overwhelmingly a stop-gap smoothing mechanisms for farm households in the sample or if the set of entrepreneurial households were made up of mostly reluctant, subsistence entrepreneurs (Adhvaryu et al., 2020; Schoar, 2010). Similarly, we could have found that only households that stayed in expected large gains while households that switched in expected substantially smaller or negligible gains, suggesting that productivity in the new sector accrues over time as in the case of learning by doing (Foster and Rosenzweig, 1995). As such, we interpret the internally consistent pattern of estimates here as a resounding confirmation of the intuition of the model and structure assumed.

Using these estimated returns, we next explore what types of households tend to have high returns to enterprise. To do this, we take each household's final return ($\beta + m_{i4}$) – which is the household's most informed or precise estimate of its return – and calculate its correlation with various household-level characteristics. We take these household characteristics from the 2014 wave of the IFLS because $\beta + m_{i4}$ is a household's perceived return going into this last wave and because this wave includes variables not found in the others (like personality traits). We first use lasso to select predictors of final returns from a large set of household-level characteristics covering a wide range of areas: cognitive ability, educational attainment, physical health, risk aversion, mental health, and personality traits (see Appendix section G.1 for a description of all variables). Then, for each of the seven variables that were selected, we calculate its correlation with the estimated final return.

These correlations, reported in Figure 5, are statistically significant and have the expected signs.[14] Returns to enterprise are positively correlated with

[14]In a multivariate regression that includes all variables, however, only Raven's scores,

Figure 5: Expected Returns and Household Characteristics



Notes: Each bar illustrates the correlation between the listed household level characteristic, taken from the 2014 wave of the IFLS, and the final return to to enterprise $(\beta + m_{i4})$. Error bars denote 95% confidence intervals. These variables were selected from a larger set of variables (listed in Appendix G.1) using lasso. All variables in this figure represent the maximum value across household members, with the exception of risk aversion, which is the average score across household members (as selected by lasso).

cognitive ability (measured by Raven's test scores), educational attainment, adult height, open-mindedness, conscientiousness, and extraversion. In addition, returns are negatively correlated with risk aversion.

It is important to note that these variables explain only a small percentage of the variation in returns. In a multivariate regression that includes all seven variables, the adjusted R-squared is less than 0.02.[15] In other words, returns to enterprise are driven primarily by unobservables, which could explain why it is difficult for households to calculate their returns to enterprise and therefore why suboptimal sorting decisions are common, as we discuss in the following sub-section.

4.3 Misallocation

Because households switch in and out of enterprise as they learn more information about their η_i, many households spend time in a sector which is suboptimal. To identify households that are misallocated, we use the household's beliefs about its relative productivity going into the final period (m_{i4}), and calculate its expected return to enterprise using this value ($\beta + m_{i4}$). Households with a positive return should be in the enterprise sector, while households with a negative return should be in non-enterprise.[16] Based on this information, we characterize households as misallocated if they are not in their optimal sector. Figure 6 shows that a large share of households are misallocated in each wave. This share declines from 35% in 1993 to 27% in 2007, indicating that households are learning about their true η_i and becoming increasingly likely to select their optimal sector.[17]

We next explore the costs of this misallocation, represented by the absolute value of enterprise returns (calculated using final beliefs about η_i, as described

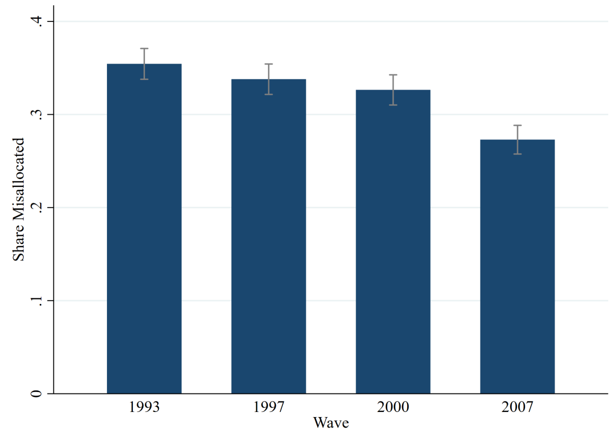
risk aversion, and open-mindedness yield statistically significant coefficients.

[15]In a multivariate regressions with all 27 variables originally included in the lasso, the adjusted R-squared is only 0.016.

[16]Note that the underlying incomes and, as a result, these estimated returns are in terms of *net* earnings. As such, any costs of engaging in either activity are already accounted for.

[17]Remember that the sample is a balanced panel such that these patterns are not driven by the entry of new households.

Figure 6: Share of Households Misallocated

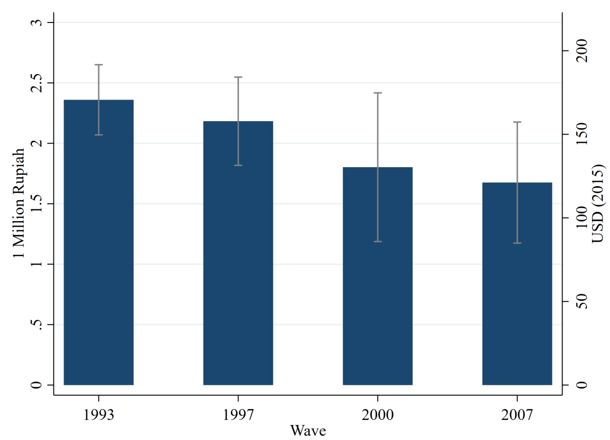


Notes: Misallocated households are defined as those with final returns $(\beta + m_{i4})$ greater than zero but in the non-enterprise sector, or those with final returns less than zero but in the enterprise sector. Error bars denote 95% confidence intervals.

above) among misallocated households. Misallocated households who are currently in non-enterprise but should be in enterprise have a positive return, which represents unrealized income gains due to their misallocation. Similarly, misallocated households who are currently in enterprise but should be in non-enterprise have a negative return, the absolute value of which represents how much more they could have earned if they had chosen enterprise instead. We sum all of these misallocation amounts for each wave and divide by the total number of households in the sample. We plot these values in Figure 7. Misallocation leads to losses of around 2.4 million rupiah (about 170 USD) per household in 1993. This declines over time, driven both by reductions in the share of misallocated households and the extent of their misallocation. That is, as households converge over time and beliefs become more precise, fewer households are misallocated and the remaining misallocated households are more "marginally" misallocated with smaller average forgone earnings per misallocated household.

We can also express these amounts as a fraction of total potential income (which is equal to a misallocated household's realized income plus their return). As we show in Appendix Figure A3, misallocated amounts correspond to almost 50% of misallocated households' potential income overall in 1993. Put differently, misallocated households earn nearly 50% less than they could have had they been in their optimal sector. This figure decreases to around 30% in 2007. Though misallocation appears in both sectors, these lost earnings percentages are higher for those who should be (but are not) in the enterprise sector than those who should be in the non-enterprise sector (panel B of Figure A3).

Figure 7: Average Misallocated Income



Notes: A household's misallocated income is equal to zero if they are not misallocated, and equal to the absolute value of their estimated final return $(\beta + m_{i4})$ if they are misallocated. Standard errors are calculated analytically (see Appendix F).

Finally, to check whether trimming of outliers in the preparation of the final balanced household panel sample impacted the estimates, we also estimate the model on the full available sample (approximately 4000 households), winsorizing income at the 1st and 99th percentiles in each wave as an alternative way to deal with outliers in reported incomes.[18] Estimating the DCRC model on this dataset yields very similar estimates of β and ϕ (1.1 and -2,

[18]The main results are estimated from a dataset in which we trim the top and bottom 5%

both statistically significant at the 5% level). We also obtain almost identical estimates of misallocated shares in each year, as shown in Appendix Figure A4.

5 Conclusion

We use a dynamic sectoral sorting framework to study the developing country household's decision to engage in non-farm enterprise. Using an extension of projection-based panel methods to estimate the resulting generalized earnings equation with dynamic correlated random coefficients on the national panel sample from the IFLS, we find that households select into non-farm enterprise on the basis of comparative advantage. We also document churning along the sectoral margin and show it is, at least in part, due to substantial learning about relative abilities across sectors and slow convergence to optimal sectors. Many households spend substantial amounts of time in a sector which is sub-optimal for them, earning up to 50% less on average than they could have if they were properly sorted across sectors.

The misallocation of households across sectors along with the high degree of churning emphasized by our dynamic approach also helps to explain some persistent patterns from the large set of interventions aimed at microenterprises over more than a decade. Experimental interventions offering financial, managerial, and human capital to existing enterprises and potential entrepreneurs yield evidence that interest in adopting these interventions is often low, at least among a large subset of the study sample; and impacts are at best heterogeneous and at worst null or negative.[19] Moreover, attrition from study

of the household income distribution in each year due to suspected reporting errors, which are common in these types of data.

[19]Mixed results have been found with respect to the effects of access to microcredit on entrepreneurial performance and entry (Banerjee et al., 2015; Crépon et al., 2015; De Mel et al., 2008; Dupas and Robinson, 2013; Kaboski and Townsend, 2012; Meager, 2019). Some research on the contribution of skills and entrepreneurial ability has found small positive effects on performance of existing enterprises (Bloom et al., 2013; Bruhn et al., 2010; Calderon et al., 2020; Karlan and Valdivia, 2011) , while others have found insignificant effects or even negative effects (De Mel et al., 2012; Karlan et al., 2015), and yet others substantial heterogeneity (Carter et al., 2019).

samples is high (McKenzie and Woodruff, 2014). The heterogeneity in returns to enterprise we find implies that some households might benefit greatly from consulting, training, and even financing interventions, but many others would benefit little and be less likely to adopt or comply. In addition, the uncertainty about returns and consequently future enterprise activity conveyed by the learning patterns we estimate would further reduce the incentives for households to invest newly available (or newly cheap) credit into enterprises and/or additional effort into improving business practices.

The results of this study have several important implications for policy and future research. Interventions aimed at encouraging entrepreneurship among persistently non-entrepreneurial households might be misguided and inefficient. Some households might simply be better at farming and wage employment than entrepreneurial activities. On the other hand, financial, managerial, and human capital interventions aimed at improving performance among existing enterprises in developing contexts should be targeted to persistently entrepreneurial households that have demonstrated a commitment to the sector over time. This would not only serve to reduce issues with high attrition and low adoption and compliance, but would likely lead to larger improvements among these more committed and better-suited households. That said, our results indicate that predicting suitability from even a rich set of observables is difficult, but rather that relative abilities are revealed slowly over time. How to hasten this learning and convergence process would be a promising avenue for future research.

References

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2017). Misallocation, selection and productivity: A quantitative analysis with panel data from china. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. (2019a). No line left behind: Assortative matching inside the firm. *Available at SSRN 3462873*.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2020). Booms, busts, and household enterprise: Evidence from coffee farmers in tanzania. *The World Bank Economic Review*.

Adhvaryu, A. and Nyshadham, A. (2017). Health, enterprise, and labor complementarity in the household. *Journal of development economics*, 126:91–111.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2019b). Managerial quality and productivity dynamics. Technical report, Mimeo, University of Michigan, Boston College and Harvard Business School.

Alvarez-Cuadrado, F., Amodio, F., and Porschke, M. (2019). Selection and absolute advantage in farming and entrepreneurship: Microeconomic evidence and macroeconomic implications.

Atkin, D. and Donaldson, D. (2015). Who's getting globalized? the size and implications of intra-national trade costs. Technical report, National Bureau of Economic Research.

Atkin, D., Khandelwal, A. K., and Osman, A. (2017). Exporting and firm performance: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 132(2):551–615.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015). The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7(1):22–53.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*, 1(51):51.

Bloom, N., Mahajan, A., McKenzie, D., and Roberts, J. (2010). Why do firms in developing countries have low productivity? *American Economic Review*, 100(2):619–23.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, pages 1351–1408.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. Technical report, National Bureau of Economic Research.

Bruhn, M., Karlan, D., and Schoar, A. (2010). What capital is missing in developing countries? *American Economic Review*, 100(2):629–33.

Calderon, G., Cunha, J. M., and Giorgi, G. D. (2020). Business literacy and development: Evidence from a randomized controlled trial in rural mexico. *Economic Development and Cultural Change*, 68(2):507–540.

Carter, M. R., Tjernström, E., and Toledo, P. (2019). Heterogeneous impact dynamics of a rural business development program in nicaragua. *Journal of Development Economics*, 138:77–98.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of econometrics*, 18(1):5–46.

Chamberlain, G. (1984). Panel data. *Handbook of econometrics*, 2:1247–1318.

Crépon, B. and Mairesse, J. (2008). The chamberlain approach to panel data: an overview and some simulations. In *The Econometrics of Panel Data*, pages 113–183. Springer.

Crépon, B., Devoto, F., Duflo, E., and Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco. *American Economic Journal: Applied Economics*, 7(1):123–50.

De Mel, S., McKenzie, D., and Woodruff, C. (2008). *Who are the microenterprise owners? Evidence from Sri Lanka on Tokman v. de Soto*. The World Bank.

De Mel, S., McKenzie, D., and Woodruff, C. (2012). *Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka*. The World Bank.

DeGroot, M. (1970). *Optimal Statistical Decisions*. McGraw Hill, New York.

Dupas, P. and Robinson, J. (2013). Savings constraints and microenterprise development: Evidence from a field experiment in kenya. *American Economic Journal: Applied Economics*, 5(1):163–92.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, 103(6):1176–1209.

Frankenberg, E. and Karoly, L. A. (1995). The 1993 indonesian family life survey: Overview and field report.

Frankenberg, E. and Thomas, D. (2000). The indonesia family life survey (ifls): Study design and results from waves 1 and 2.

Gibbons, R., Katz, L. F., Lemieux, T., and Parent, D. (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of labor economics*, 23(4):681–724.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1):83–116.

Heckman, J. and Vytlacil, E. (1998). Instrumental variables methods for the correlated random coefficient model: Estimating the average rate of return to schooling when the return is correlated with schooling. *Journal of Human Resources*, pages 974–987.

Hicks, J. H., Kleemans, M., Li, N. Y., and Miguel, E. (2017). Reevaluating agricultural productivity gaps with longitudinal microdata. Technical report, National Bureau of Economic Research.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Islam, N. (1995). Growth empirics: a panel data approach. *The quarterly journal of economics*, 110(4):1127–1170.

Kaboski, J. P. and Townsend, R. M. (2012). The impact of credit on village economies. *American Economic Journal: Applied Economics*, 4(2):98–133.

Karlan, D., Knight, R., and Udry, C. (2015). Consulting and capital experiments with microenterprise tailors in ghana. *Journal of Economic Behavior and Organization*, 118:281–302.

Karlan, D. and Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and statistics*, 93(2):510–527.

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

Lemieux, T. (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.

McKenzie, D. and Woodruff, C. (2014). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1):48–82.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, 11(1):57–91.

Molina, T. (2016). Reporting heterogeneity and health disparities across gender and education levels: Evidence from four countries. *Demography*, 53(2):295–323.

Restuccia, D. and Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, 1(16):1–10.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of monetary economics*, 55(2):234–250.

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

Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy*, 10(1):57–81.

Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y., and Witoelar, F. (2004). The third wave of the indonesia family life survey (ifls3): Overview and field report. *NIA/NICHD.*

Strauss, J., Witoelar, F., and Sikoki, B. (2016). The fifth wave of the indonesia family life survey: Overview and field report: Volume 1.

Strauss, J., Witoelar, F., Sikoki, B., and Wattie, A. M. (2009). The fourth wave of the indonesian family life survey (ifls4): overview and field report. *RAND Corporation*.

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

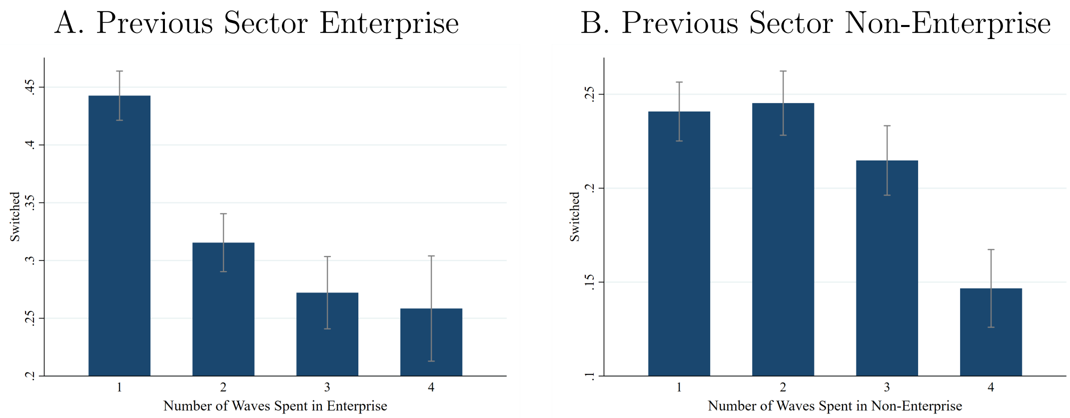
Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Zhang, W., O'Brien, N., Forrest, J. I., Salters, K. A., Patterson, T. L., Montaner, J. S., Hogg, R. S., and Lima, V. D. (2012). Validating a shortened depression scale (10 item ces-d) among hiv-positive people in british columbia, canada. *PloS one*, 7(7):e40793.

Appendix

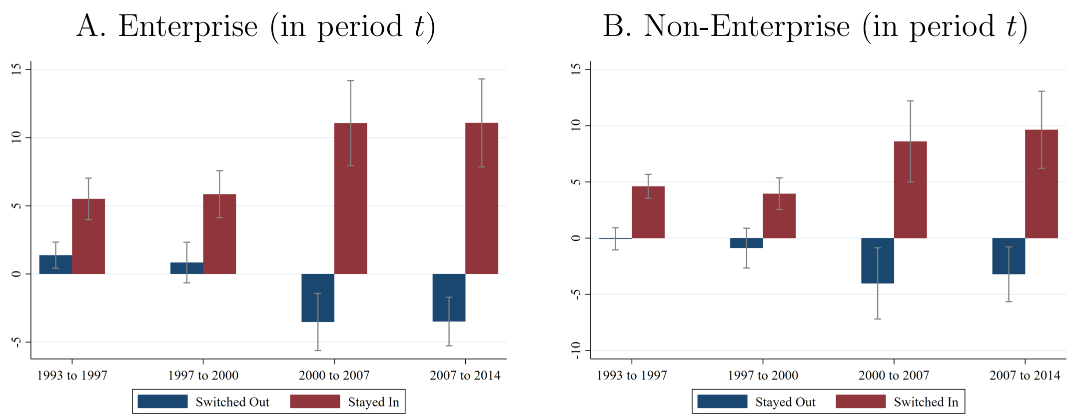
A Appendix Figures

Figure A1: Switching by Number of Waves Spent in Previous Sector: Enterprise and Non-Enterprise



Notes: Sample consists of IFLS households with non-missing income information in all five waves of the IFLS. Error bars denote 95% confidence intervals.

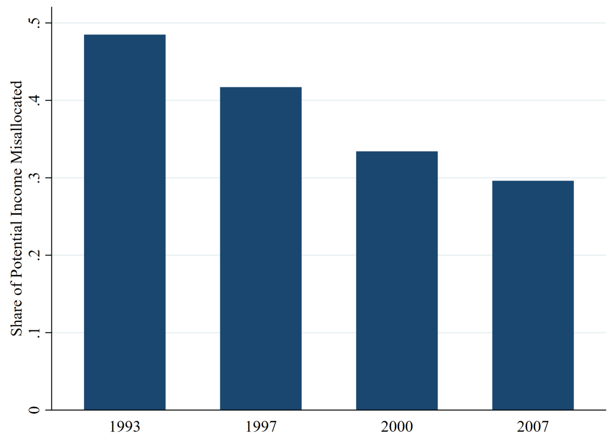
Figure A2: Expected Returns by Switch Status, Wave, and Current Sector



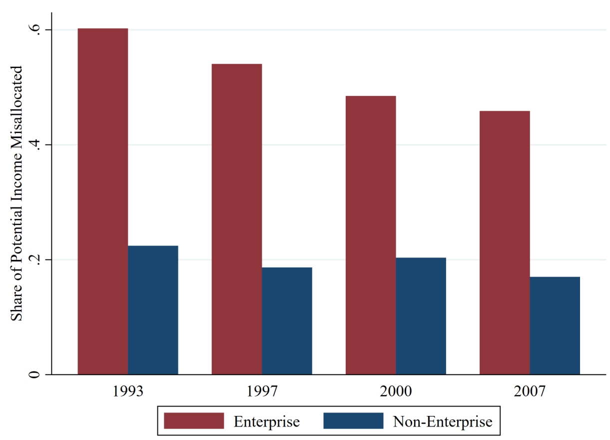
Notes: Figure reports the average return to enterprise ($\beta + \phi m_{it}$), separately for each transition, across all households in each category. "Stayed out" includes households in non-enterprise in both t and $t+1$. "Switched In" includes households in non-enterprise in t and enterprise in $t+1$. "Switched Out" includes households in enterprise in t and non-enterprise in $t+1$. "Stayed In" includes households in enterprise in both t and $t+1$. Error bars denote 95% confidence intervals. Standard errors are calculated analytically (see Appendix F).

Figure A3: Share of Potential Income Misallocated

A. Overall

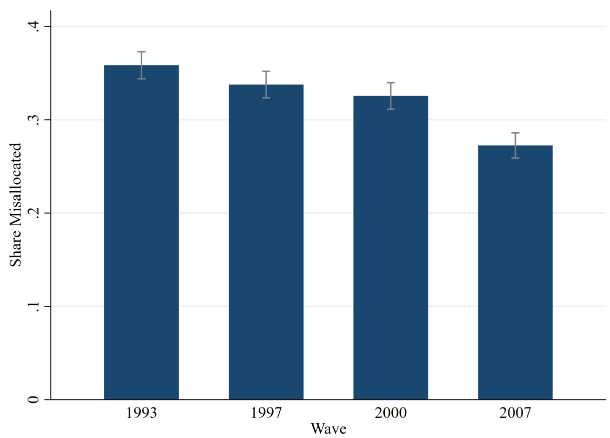


B. By Final Sector



Notes: Misallocated income is defined as the absolute value of final returns ($\beta + m_{i4}$) among misallocated households). The share of potential income misallocated is equal to the sum of all misallocated income divided by the sum of realized income among misallocated households.

Figure A4: Misallocated Shares, Full Sample



Notes: This figure uses an untrimmed sample of households, winsorizing income at the 1st and 99th percentiles in each year. Estimates (and standard errors) from the DCRC model using this sample are $\beta = 1.095$ (0.504), $\phi = -1.966$ (0.373). Misallocated households are defined as those with final returns ($\beta + m_{i4}$) greater than zero but in the non-enterprise sector, or those with final returns less than zero but in the enterprise sector. Error bars denote 95% confidence intervals.

B Interpretation of ϕ

In this section, we discuss how ϕ governs the nature of selection in the Roy model described in section 3.1. We combine equations (1) and (2) and suppress t subscripts to express income in the enterprise and non-enterprise sectors as follows:

$$
\begin{aligned}
Y_i^E &= \beta^E + (1+\phi)\eta_i + \tau_i \\
Y_i^N &= \beta^N + \eta_i + \tau_i.
\end{aligned}
$$

Unconditional expected income in the enterprise and non-enterprise sector is

$$
\begin{aligned}
E[Y_i^E] &= \beta^E + (1+\phi)E[\eta_i] + E[\tau_i] \\
E[Y_i^N] &= \beta^N + E[\eta_i] + E[\tau_i].
\end{aligned}
$$

As we discuss in section 3.1, households select into enterprise based on their η_i; specifically, households with $\phi\eta_i > -\beta$ (where $\beta \equiv \beta^E - \beta^N$) will choose to go into enterprise ($D_i = 1$). Therefore, conditional average enterprise and non-enterprise income, among those who select into enterprise, is the following:

$$
\begin{aligned}
E[Y_i^E|D_i = 1] &= E[Y_i^E|\phi\eta_i > -\beta] \\
&= \beta_t^E + (1+\phi)E[\eta_i|\phi\eta_i > -\beta] + E[\tau_i|\phi\eta_i > -\beta] \\
&= \beta_t^E + (1+\phi)E[\eta_i|\phi\eta_i > -\beta] + E[\tau_i] \\
E[Y_i^N|D_i = 1] &= E[Y_i^N|\phi\eta_i > -\beta] \\
&= \beta_t^N + E[\eta_i|\phi\eta_i > -\beta] + E[\tau_i|\phi\eta_i > -\beta] \\
&= \beta_t^N + E[\eta_i|\phi\eta_i > -\beta] + E[\tau_i],
\end{aligned}
$$

where the last step is due to the independence of τ and η.

In order to characterize sorting in the same way that Borjas (1987) does (distinguishing between positive selection, negative selection, and "refugee sorting" or sorting on comparative advantage), we focus on the same income differentials that he uses. The first differential of interest is the difference between average enterprise income among households that select into enterprise and unconditional average enterprise income (labeled Q_1 in Borjas (1987) and defined by equation (13) below). The second differential of interest is the difference between average non-enterprise income among households that select into enterprise and unconditional average non-enterprise income (labeled Q_0 in Borjas (1987) and defined by equation (14) below). Positive selection is defined as the case when $Q_1 > 0$ and $Q_0 > 0$, negative selection when $Q_1 < 0$ and $Q_0 < 0$, and sorting on comparative advantage when $Q_1 > 0$ and $Q_0 < 0$.

$$E[Y_i^E|D_i = 1] - E[Y_i^E] = (1+\phi)\left(E[\eta_i|\phi\eta_i > -\beta] - E[\eta_i]\right) \tag{13}$$
$$E[Y_i^N|D_i = 1] - E[Y_i^N] = \left(E[\eta_i|\phi\eta_i > -\beta] - E[\eta_i]\right). \tag{14}$$

B.1 Case 1: $\phi > 0$

When $\phi > 0$, average enterprise income among those who select into enterprise is higher than the population average of enterprise income, as shown below. Average non-enterprise income is also higher among those who select into enterprise. This means that enterprise households are positively selected.

$$E[Y_i^E|D_i = 1] - E[Y_i^E] = \overbrace{(1+\phi)}^{>0}\overbrace{\left(E[\eta_i|\eta_i > -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{>0} > 0$$
$$E[Y_i^N|D_i = 1] - E[Y_i^N] = \overbrace{\left(E[\eta_i|\eta_i > -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{>0} > 0.$$

B.2 Case 2: $-1 < \phi < 0$

When $-1 < \phi < 0$, we have negative selection. Both average enterprise income and average non-enterprise income among those who select into enterprise are lower than population averages. Those who select into enterprise tend to be less productive in both enterprise and non-enterprise.

$$E[Y_i^E|D_i = 1] - E[Y_i^E] = \overbrace{(1+\phi)}^{>0}\overbrace{\left(E[\eta_i|\eta_i < -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{<0} < 0$$

$$E[Y_i^N|D_i = 1] - E[Y_i^N] = \overbrace{\left(E[\eta_i|\eta_i < -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{<0} < 0.$$

B.3 Case 3: $\phi < -1$

Finally, when $\phi < -1$, average enterprise income among those who select into enterprise is higher than the population average of enterprise income. However, average non-enterprise income is lower among those who select into enterprise. This implies sorting based on comparative advantage: productive enterprise households would have low productivity in non-enterprise, while productive non-enterprise households would have low productivity in enterprise.

$$E[Y_i^E|D_i = 1] - E[Y_i^E] = \overbrace{(1+\phi)}^{<0}\overbrace{\left(E[\eta_i|\eta_i < -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{<0} > 0$$

$$E[Y_i^N|D_i = 1] - E[Y_i^N] = \overbrace{\left(E[\eta_i|\eta_i < -\frac{\beta}{\phi}] - E[\eta_i]\right)}^{<0} < 0.$$

C Additional Equations

C.1 Income Equations

The following equations express income in each period (Y_{it}) as a function of each period's sectoral choice (D_{it}) and structural parameters.

$$
\begin{aligned}
Y_{i1} =& \alpha_1 + \beta D_{i1} + (\lambda_0 + \prod_{t=1}^{5}(1 + \lambda_t D_{it}) - 1)(1 + \phi D_{i1}) + \\
& (\omega_{i1} + \psi_{i0})(1 + \phi D_{i1}) + \nu_{i1} \\
Y_{i2} =& \alpha_2 + \beta D_{i2} + (\lambda_0 + \prod_{t=1}^{5}(1 + \lambda_t D_{it}) - 1 + \theta_{20} + \sum_{t=2}^{5} \theta_{2t} D_{it})(1 + \phi D_{i2}) + \\
& (\omega_{i2} + \psi_{i0} + \psi_{i1})(1 + \phi D_{i2}) + \nu_{i2} \\
Y_{i3} =& \alpha_3 + \beta D_{i3} + (\lambda_0 + \prod_{t=1}^{5}(1 + \lambda_t D_{it}) - 1 + \theta_{30} + \sum_{t=3}^{5} \theta_{3t} D_{it})(1 + \phi D_{i3}) + \\
& (\omega_{i3} + \psi_{i0} + \psi_{i1} + \psi_{i2})(1 + \phi D_{i3}) + \nu_{i3} \\
Y_{i4} =& \alpha_4 + \beta D_{i4} + (\lambda_0 + \prod_{t=1}^{5}(1 + \lambda_t D_{it}) - 1 + \theta_{40} + \sum_{t=4}^{5} \theta_{4t} D_{it})(1 + \phi D_{i4}) + \\
& (\omega_{i4} + \psi_{i0} + \psi_{i1} + \psi_{i2} + \psi_{i3})(1 + \phi D_{i4}) + \nu_{i4} \\
Y_{i5} =& \alpha_5 + \beta D_{i5} + (\lambda_0 + \prod_{t=1}^{5}(1 + \lambda_t D_{it}) - 1 + \theta_{50} + \theta_{55} D_{i5})(1 + \phi D_{i5}) + \\
& (\omega_{i5} + \psi_{i0} + \psi_{i1} + \psi_{i2} + \psi_{i3} + \psi_{i4})(1 + \phi D_{i5}) + \nu_{i5}
\end{aligned} \tag{15}
$$

C.2 Minimum Distance Restrictions

The minimum distance restrictions are as follows.

$\gamma_1^1 = \beta + \phi\lambda_0 + \lambda_1 + \phi\lambda_1$

$\gamma_2^1 = \lambda_2$

$\gamma_3^1 = \lambda_3$

$$\gamma_4^1 = \lambda_4$$

$$\gamma_5^1 = \lambda_5$$

$$\gamma_{12}^1 = \phi\lambda_2 + \lambda_{12} + \phi\lambda_{12}$$

$$\gamma_{13}^1 = \phi\lambda_3 + \lambda_{13} + \phi\lambda_{13}$$

$$\gamma_{14}^1 = \phi\lambda_4 + \lambda_{14} + \phi\lambda_{14}$$

$$\gamma_{15}^1 = \phi\lambda_5 + \lambda_{15} + \phi\lambda_{15}$$

$$\gamma_{23}^1 = \lambda_{23}$$

$$\gamma_{24}^1 = \lambda_{24}$$

$$\gamma_{25}^1 = \lambda_{25}$$

$$\gamma_{34}^1 = \lambda_{34}$$

$$\gamma_{35}^1 = \lambda_{35}$$

$$\gamma_{45}^1 = \lambda_{45}$$

$$\gamma_{123}^1 = \phi\lambda_{23} + \lambda_{123} + \phi\lambda_{123}$$

$$\gamma_{124}^1 = \phi\lambda_{24} + \lambda_{124} + \phi\lambda_{124}$$

$$\gamma_{125}^1 = \phi\lambda_{25} + \lambda_{125} + \phi\lambda_{125}$$

$$\gamma_{134}^1 = \phi\lambda_{34} + \lambda_{134} + \phi\lambda_{134}$$

$$\gamma_{135}^1 = \phi\lambda_{35} + \lambda_{135} + \phi\lambda_{135}$$

$$\gamma_{145}^1 = \phi\lambda_{45} + \lambda_{145} + \phi\lambda_{145}$$

$$\gamma_{234}^1 = \lambda_{234}$$

$\gamma^1_{235} = \lambda_{235}$

$\gamma^1_{245} = \lambda_{245}$

$\gamma^1_{345} = \lambda_{345}$

$\gamma^1_{1234} = \phi\lambda_{234} + \lambda_{1234} + \phi\lambda_{1234}$

$\gamma^1_{1235} = \phi\lambda_{235} + \lambda_{1235} + \phi\lambda_{1235}$

$\gamma^1_{1245} = \phi\lambda_{245} + \lambda_{1245} + \phi\lambda_{1245}$

$\gamma^1_{1345} = \phi\lambda_{345} + \lambda_{1345} + \phi\lambda_{1345}$

$\gamma^1_{2345} = \lambda_{2345}$

$\gamma^1_{12345} = \phi\lambda_{2345} + \lambda_{12345} + \phi\lambda_{12345}$

$\gamma^2_{1} = \lambda_{1}$

$\gamma^2_{2} = \beta + \phi\theta_{20} + \theta_{22} + \phi\theta_{22} + \phi\lambda_0 + \lambda_2 + \phi\lambda_2$

$\gamma^2_{3} = \theta_{23} + \lambda_3$

$\gamma^2_{4} = \theta_{24} + \lambda_4$

$\gamma^2_{5} = \theta_{25} + \lambda_5$

$\gamma^2_{12} = \phi\lambda_1 + \lambda_{12} + \phi\lambda_{12}$

$\gamma^2_{13} = \lambda_{13}$

$\gamma^2_{14} = \lambda_{14}$

$\gamma^2_{15} = \lambda_{15}$

$\gamma^2_{23} = \phi\theta_{23} + \phi\lambda_3 + \lambda_{23} + \phi\lambda_{23}$

$$\gamma_{24}^2 = \phi\theta_{24} + \phi\lambda_4 + \lambda_{24} + \phi\lambda_{24}$$

$$\gamma_{25}^2 = \phi\theta_{25} + \phi\lambda_5 + \lambda_{25} + \phi\lambda_{25}$$

$$\gamma_{34}^2 = \lambda_{34}$$

$$\gamma_{35}^2 = \lambda_{35}$$

$$\gamma_{45}^2 = \lambda_{45}$$

$$\gamma_{123}^2 = \phi\lambda_{13} + \lambda_{123} + \phi\lambda_{123}$$

$$\gamma_{124}^2 = \phi\lambda_{14} + \lambda_{124} + \phi\lambda_{124}$$

$$\gamma_{125}^2 = \phi\lambda_{15} + \lambda_{125} + \phi\lambda_{125}$$

$$\gamma_{134}^2 = \lambda_{134}$$

$$\gamma_{135}^2 = \lambda_{135}$$

$$\gamma_{145}^2 = \lambda_{145}$$

$$\gamma_{234}^2 = \phi\lambda_{34} + \lambda_{234} + \phi\lambda_{234}$$

$$\gamma_{235}^2 = \phi\lambda_{35} + \lambda_{235} + \phi\lambda_{235}$$

$$\gamma_{245}^2 = \phi\lambda_{45} + \lambda_{245} + \phi\lambda_{245}$$

$$\gamma_{345}^2 = \lambda_{345}$$

$$\gamma_{1234}^2 = \phi\lambda_{134} + \lambda_{1234} + \phi\lambda_{1234}$$

$$\gamma_{1235}^2 = \phi\lambda_{135} + \lambda_{1235} + \phi\lambda_{1235}$$

$$\gamma_{1245}^2 = \phi\lambda_{145} + \lambda_{1245} + \phi\lambda_{1245}$$

$$\gamma_{1345}^2 = \lambda_{1345}$$

$$\gamma^2_{2345} = \phi\lambda_{345} + \lambda_{2345} + \phi\lambda_{2345}$$

$$\gamma^2_{12345} = \phi\lambda_{1345} + \lambda_{12345} + \phi\lambda_{12345}$$

$$\gamma^3_1 = \lambda_1$$

$$\gamma^3_2 = \lambda_2$$

$$\gamma^3_3 = \beta + \phi\theta_{30} + \theta_{33} + \phi\theta_{33} + \phi\lambda_0 + \lambda_3 + \phi\lambda_3$$

$$\gamma^3_4 = \theta_{34} + \lambda_4$$

$$\gamma^3_5 = \theta_{35} + \lambda_5$$

$$\gamma^3_{12} = \lambda_{12}$$

$$\gamma^3_{13} = \phi\lambda_1 + \lambda_{13} + \phi\lambda_{13}$$

$$\gamma^3_{14} = \lambda_{14}$$

$$\gamma^3_{15} = \lambda_{15}$$

$$\gamma^3_{23} = \phi\lambda_2 + \lambda_{23} + \phi\lambda_{23}$$

$$\gamma^3_{24} = \lambda_{24}$$

$$\gamma^3_{25} = \lambda_{25}$$

$$\gamma^3_{34} = \phi\theta_{34} + \phi\lambda_4 + \lambda_{34} + \phi\lambda_{34}$$

$$\gamma^3_{35} = \phi\theta_{35} + \phi\lambda_5 + \lambda_{35} + \phi\lambda_{35}$$

$$\gamma^3_{45} = \lambda_{45}$$

$$\gamma^3_{123} = \phi\lambda_{12} + \lambda_{123} + \phi\lambda_{123}$$

$$\gamma^3_{124} = \lambda_{124}$$

$$\gamma^3_{125} = \lambda_{125}$$

$$\gamma^3_{134} = \phi\lambda_{14} + \lambda_{134} + \phi\lambda_{134}$$

$$\gamma^3_{135} = \phi\lambda_{15} + \lambda_{135} + \phi\lambda_{135}$$

$$\gamma^3_{145} = \lambda_{145}$$

$$\gamma^3_{234} = \phi\lambda_{24} + \lambda_{234} + \phi\lambda_{234}$$

$$\gamma^3_{235} = \phi\lambda_{25} + \lambda_{235} + \phi\lambda_{235}$$

$$\gamma^3_{245} = \lambda_{245}$$

$$\gamma^3_{345} = \phi\lambda_{45} + \lambda_{345} + \phi\lambda_{345}$$

$$\gamma^3_{1234} = \phi\lambda_{124} + \lambda_{1234} + \phi\lambda_{1234}$$

$$\gamma^3_{1235} = \phi\lambda_{125} + \lambda_{1235} + \phi\lambda_{1235}$$

$$\gamma^3_{1245} = \lambda_{1245}$$

$$\gamma^3_{1345} = \phi\lambda_{145} + \lambda_{1345} + \phi\lambda_{1345}$$

$$\gamma^3_{2345} = \phi\lambda_{245} + \lambda_{2345} + \phi\lambda_{2345}$$

$$\gamma^3_{12345} = \phi\lambda_{1245} + \lambda_{12345} + \phi\lambda_{12345}$$

$$\gamma^4_{1} = \lambda_{1}$$

$$\gamma^4_{2} = \lambda_{2}$$

$$\gamma^4_{3} = \lambda_{3}$$

$$\gamma^4_{4} = \beta + \phi\theta_{40} + \theta_{44} + \phi\theta_{44} + \phi\lambda_0 + \lambda_4 + \phi\lambda_4$$

$$\gamma^4_{5} = \theta_{45} + \lambda_5$$

$$\gamma^4_{12} = \lambda_{12}$$

$$\gamma^4_{13} = \lambda_{13}$$

$$\gamma^4_{14} = \phi\lambda_1 + \lambda_{14} + \phi\lambda_{14}$$

$$\gamma^4_{15} = \lambda_{15}$$

$$\gamma^4_{23} = \lambda_{23}$$

$$\gamma^4_{24} = \phi\lambda_2 + \lambda_{24} + \phi\lambda_{24}$$

$$\gamma^4_{25} = \lambda_{25}$$

$$\gamma^4_{34} = \phi\lambda_3 + \lambda_{34} + \phi\lambda_{34}$$

$$\gamma^4_{35} = \lambda_{35}$$

$$\gamma^4_{45} = \phi\theta_{45} + \phi\lambda_5 + \lambda_{45} + \phi\lambda_{45}$$

$$\gamma^4_{123} = \lambda_{123}$$

$$\gamma^4_{124} = \phi\lambda_{12} + \lambda_{124} + \phi\lambda_{124}$$

$$\gamma^4_{125} = \lambda_{125}$$

$$\gamma^4_{134} = \phi\lambda_{13} + \lambda_{134} + \phi\lambda_{134}$$

$$\gamma^4_{135} = \lambda_{135}$$

$$\gamma^4_{145} = \phi\lambda_{15} + \lambda_{145} + \phi\lambda_{145}$$

$$\gamma^4_{234} = \phi\lambda_{23} + \lambda_{234} + \phi\lambda_{234}$$

$$\gamma^4_{235} = \lambda_{235}$$

$$\gamma^4_{245} = \phi\lambda_{25} + \lambda_{245} + \phi\lambda_{245}$$

$$\gamma^4_{345} = \phi\lambda_{35} + \lambda_{345} + \phi\lambda_{345}$$

$$\gamma^4_{1234} = \phi\lambda_{123} + \lambda_{1234} + \phi\lambda_{1234}$$

$$\gamma^4_{1235} = \lambda_{1235}$$

$$\gamma^4_{1245} = \phi\lambda_{125} + \lambda_{1245} + \phi\lambda_{1245}$$

$$\gamma^4_{1345} = \phi\lambda_{135} + \lambda_{1345} + \phi\lambda_{1345}$$

$$\gamma^4_{2345} = \phi\lambda_{235} + \lambda_{2345} + \phi\lambda_{2345}$$

$$\gamma^4_{12345} = \phi\lambda_{2345} + \lambda_{12345} + \phi\lambda_{12345}$$

$$\gamma^5_{1} = \lambda_{1}$$

$$\gamma^5_{2} = \lambda_{2}$$

$$\gamma^5_{3} = \lambda_{3}$$

$$\gamma^5_{4} = \lambda_{4}$$

$$\gamma^5_{5} = \beta + \phi\theta_{50} + \theta_{55} + \phi\theta_{55} + \phi\lambda_0 + \lambda_5 + \phi\lambda_5$$

$$\gamma^5_{12} = \lambda_{12}$$

$$\gamma^5_{13} = \lambda_{13}$$

$$\gamma^5_{14} = \lambda_{14}$$

$$\gamma^5_{15} = \phi\lambda_1 + \lambda_{15} + \phi\lambda_{15}$$

$$\gamma^5_{23} = \lambda_{23}$$

$$\gamma^5_{24} = \lambda_{24}$$

$$\gamma^5_{25} = \phi\lambda_2 + \lambda_{25} + \phi\lambda_{25}$$

$\gamma^5_{34} = \lambda_{34}$

$\gamma^5_{35} = \phi\lambda_3 + \lambda_{35} + \phi\lambda_{35}$

$\gamma^5_{45} = \phi\lambda_4 + \lambda_{45} + \phi\lambda_{45}$

$\gamma^5_{123} = \lambda_{123}$

$\gamma^5_{124} = \lambda_{124}$

$\gamma^5_{125} = \phi\lambda_{12} + \lambda_{125} + \phi\lambda_{125}$

$\gamma^5_{134} = \lambda_{134}$

$\gamma^5_{135} = \phi\lambda_{13} + \lambda_{135} + \phi\lambda_{135}$

$\gamma^5_{145} = \phi\lambda_{14} + \lambda_{145} + \phi\lambda_{145}$

$\gamma^5_{234} = \lambda_{234}$

$\gamma^5_{235} = \phi\lambda_{23} + \lambda_{235} + \phi\lambda_{235}$

$\gamma^5_{245} = \phi\lambda_{24} + \lambda_{245} + \phi\lambda_{245}$

$\gamma^5_{345} = \phi\lambda_{34} + \lambda_{345} + \phi\lambda_{345}$

$\gamma^5_{1234} = \lambda_{1234}$

$\gamma^5_{1235} = \phi\lambda_{123} + \lambda_{1235} + \phi\lambda_{1235}$

$\gamma^5_{1245} = \phi\lambda_{124} + \lambda_{1245} + \phi\lambda_{1245}$

$\gamma^5_{1345} = \phi\lambda_{134} + \lambda_{1345} + \phi\lambda_{1345}$

$\gamma^5_{2345} = \phi\lambda_{234} + \lambda_{2345} + \phi\lambda_{2345}$

$\gamma^5_{12345} = \phi\lambda_{1234} + \lambda_{12345} + \phi\lambda_{12345}$

Table C1 reports the γ coefficients estimated by the SUR, the value of

γ that is implied by the structural parameter estimates and the minimum distance restrictions above, and the t-statistics for the differences between the two.

Table C1: Tests of Minimum Distance Restrictions

	SUR		Structural		Diff
	γ	SE	γ	SE	t
γ^1_1	1.56	0.7	1.88	0.52	-0.38
γ^1_2	0.2	0.75	-0.29	0.39	0.57
γ^1_3	1.45	1.59	1.37	0.86	0.05
γ^1_4	-0.43	0.57	0.02	0.37	-0.66
γ^1_5	2.52	1.59	-0.19	0.93	1.47
γ^1_{12}	-0.49	1.14	0.11	0.51	-0.48
γ^1_{13}	-2.39	2.43	-1.16	1.21	-0.46
γ^1_{23}	0.87	0.56	0.61	0.42	0.38
γ^1_{14}	-1.14	1.38	-0.55	1.02	-0.35
γ^1_{24}	0.41	1.88	-0.74	0.79	0.57
γ^1_{34}	0.29	3.27	0.31	1.83	-0.01
γ^1_{15}	1.23	1.21	-0.72	0.6	1.44
γ^1_{25}	-2.26	2.42	-0.33	1.47	-0.68
γ^1_{35}	-1.38	2.43	0.73	0.99	-0.8
γ^1_{45}	1.34	4.09	1.01	2.24	0.07
γ^1_{123}	0.56	0.65	0.36	0.45	0.26
γ^1_{124}	1.15	1.95	-1.25	1.11	1.07
γ^1_{134}	0.79	1.71	-0.34	0.75	0.6
γ^1_{234}	-3.91	3.2	-0.69	1.63	-0.9
γ^1_{125}	1.53	1.24	0	0.58	1.12
γ^1_{135}	-4.19	2.89	0.98	1.49	-1.59
γ^1_{235}	-2.56	2.4	-0.87	1.03	-0.65
γ^1_{145}	6.79	4.41	2.57	2.18	0.86
γ^1_{245}	-1.2	0.99	-0.38	0.59	-0.71
γ^1_{345}	2.09	3.08	1.32	1.72	0.22
γ^1_{1234}	-2.09	2.67	0.15	1.12	-0.77
γ^1_{1235}	3.42	5.31	-0.12	2.84	0.59
γ^1_{1245}	-2.83	1.81	0.15	0.87	-1.49
γ^1_{1345}	3.75	4.12	1.15	2.22	0.56
γ^1_{2345}	4.4	3.53	0.44	1.47	1.03
γ^1_{12345}	-6.92	6.49	-2.59	3.41	-0.59
γ^2_1	-1.6	0.75	-0.42	0.38	-1.4
γ^2_2	3.57	1.07	2.56	0.62	0.82
γ^2_3	0.24	1.83	1.7	0.87	-0.72
γ^2_4	0.15	0.71	0.5	0.43	-0.43
γ^2_5	0.95	1.55	0.08	0.61	0.52
γ^2_{12}	-1.28	1.71	-1.44	0.85	0.09
γ^2_{13}	0.25	2.71	-1.1	1.16	0.46
γ^2_{23}	1.02	0.73	-0.01	0.43	1.2
γ^2_{14}	1.08	1.94	-0.65	0.71	0.84
γ^2_{24}	2.03	2.13	1.16	1.24	0.35
γ^2_{34}	0.36	4.08	0.08	1.88	0.06
γ^2_{15}	-0.58	1.33	-0.72	0.6	0.1
γ^2_{25}	-1.09	2.94	1.4	1.09	-0.8
γ^2_{35}	-2.76	2.88	0.7	1.37	-1.08
γ^2_{45}	0.76	5.04	-0.69	2.25	0.26
γ^2_{123}	1.56	0.75	0.39	0.47	1.32
γ^2_{124}	1.08	2.35	0.22	0.75	0.35
γ^2_{134}	1.99	2.36	-0.47	1.14	0.94
γ^2_{234}	-4.97	3.92	-2.11	1.62	-0.67
γ^2_{125}	-2.42	1.47	0	0.58	-1.53
γ^2_{135}	-1.42	3.31	-0.64	1.06	-0.22

	SUR		Structural		Diff
	γ	SE	γ	SE	t
γ^2_{235}	-1.28	3.19	1.34	1.43	-0.75
γ^2_{145}	5.63	5.17	1.97	2.18	0.65
γ^2_{245}	-2.99	1.2	-0.38	0.59	-1.95
γ^2_{345}	-0.73	3.47	-0.23	1.2	-0.14
γ^2_{1234}	-0.45	3.62	0.74	1.6	-0.3
γ^2_{1235}	2.72	6.39	0.83	2.79	0.27
γ^2_{1245}	4	2.09	0.15	0.87	1.7
γ^2_{1345}	-0.88	4.63	-0.99	1.62	0.02
γ^2_{2345}	1.23	4.56	-1.06	2.04	0.46
γ^2_{12345}	-1.88	7.68	1.05	3.37	-0.35
γ^3_1	-1.58	0.85	-0.42	0.38	-1.25
γ^3_2	-0.47	1.06	-0.29	0.39	-0.16
γ^3_3	-0.28	1.68	-0.42	0.62	0.08
γ^3_4	1.51	0.79	1.23	0.54	0.29
γ^3_5	1.32	1.69	0.93	0.86	0.21
γ^3_{12}	0.23	1.69	0.56	0.71	-0.18
γ^3_{13}	0.85	2.68	0.18	1.16	0.23
γ^3_{23}	1.64	0.78	0.08	0.43	1.75
γ^3_{14}	-1.53	1.69	-0.65	0.71	-0.48
γ^3_{24}	-1.98	2.29	-0.74	0.79	-0.51
γ^3_{34}	5.93	3.91	1.02	1.32	1.19
γ^3_{15}	2.33	1.57	0.9	0.93	0.78
γ^3_{25}	-1.12	2.9	-0.51	1.37	-0.19
γ^3_{35}	1.06	3.12	0.76	1.23	0.09
γ^3_{45}	-3.38	5.01	0.28	2	-0.68
γ^3_{123}	-0.39	0.79	0.19	0.47	-0.63
γ^3_{124}	5.79	2.45	0.22	0.75	2.17
γ^3_{134}	0.41	2	-0.34	0.75	0.35
γ^3_{234}	-3.84	3.73	1	1.21	-1.23
γ^3_{125}	1.09	1.56	-0.49	0.86	0.89
γ^3_{135}	-5.38	3.4	0.41	1.35	-1.58
γ^3_{235}	3.93	3.24	2.2	1.36	0.49
γ^3_{145}	1.73	5.23	-2.19	2.08	0.7
γ^3_{245}	0.25	1.31	-0.38	0.59	0.44
γ^3_{345}	-3.77	3.56	-0.23	1.2	-0.94
γ^3_{1234}	1.61	3.42	0.15	1.12	0.41
γ^3_{1235}	-1.59	6.4	-0.17	2.04	-0.21
γ^3_{1245}	-3.25	2.38	0.74	1.15	-1.51
γ^3_{1345}	7.64	4.84	2.1	1.98	1.06
γ^3_{2345}	-4.48	4.77	-1.05	1.86	-0.67
γ^3_{12345}	1	7.92	-1.05	3	0.24
γ^4_1	-1.32	1.07	-0.42	0.38	-0.79
γ^4_2	2.68	1.4	-0.29	0.39	2.04
γ^4_3	-0.98	2.31	-0.42	0.62	-0.24
γ^4_4	1.16	1.15	0.02	0.37	0.94
γ^4_5	0.45	2.32	0.08	0.61	0.15
γ^4_{12}	-1.61	2.24	0.11	0.51	-0.75
γ^4_{13}	-1.28	3.66	0.57	0.9	-0.49
γ^4_{23}	6.07	1.08	2.67	0.71	2.63
γ^4_{14}	0.61	2.82	2.06	1.08	-0.48
γ^4_{24}	-3.84	2.75	1.87	1.17	-1.91
γ^4_{34}	5.6	5.3	-0.51	1.93	1.08

	SUR		Structural		Diff
	γ	SE	γ	SE	t
γ^4_{15}	-0.4	2.12	1.05	0.9	-0.63
γ^4_{25}	-4.94	4.28	-2.38	1.5	-0.56
γ^4_{35}	2.49	3.98	-1.4	1.42	0.92
γ^4_{45}	2.94	6.79	1.42	2.37	0.21
γ^4_{123}	0.17	0.99	-0.05	0.52	0.2
γ^4_{124}	3.86	2.86	0.22	0.75	1.23
γ^4_{134}	1.53	2.99	-0.34	0.75	0.61
γ^4_{234}	-6.07	5	1	1.21	-1.37
γ^4_{125}	-0.05	2.11	0	0.58	-0.02
γ^4_{135}	-4.95	4.2	-0.64	1.06	-0.99
γ^4_{235}	-4.08	4.22	-0.87	1.03	-0.74
γ^4_{145}	11.77	6.74	-0.23	1.61	1.73
γ^4_{245}	-0.17	1.84	0.73	1.04	-0.43
γ^4_{345}	-3.76	4.43	-0.22	1.65	-0.75
γ^4_{1234}	0.07	4.84	0.63	1.52	-0.11
γ^4_{1235}	-0.27	8.67	-2.29	2.81	0.22
γ^4_{1245}	0.43	3.23	-0.23	1.19	0.19
γ^4_{1345}	9.02	6.21	3.15	2.15	0.89
γ^4_{2345}	2.29	6.39	1.54	1.96	0.11
γ^4_{12345}	-8.73	10.58	-0.88	3.38	-0.71
γ^5_1	-1.12	1.87	-0.42	0.38	-0.37
γ^5_2	2.27	2.13	-0.29	0.39	1.18
γ^5_3	-3.29	3.69	-0.42	0.62	-0.77
γ^5_4	1.01	1.48	0.02	0.37	0.65
γ^5_5	-0.86	3.17	0.08	0.61	-0.29
γ^5_{12}	0.01	3.31	0.11	0.51	-0.03
γ^5_{13}	0.49	5.3	0.57	0.9	-0.02
γ^5_{23}	3.46	1.68	0.61	0.42	1.65
γ^5_{14}	-4.81	3.7	-0.65	0.71	-1.1
γ^5_{24}	-7.46	3.9	-0.74	0.79	-1.69
γ^5_{34}	9.38	6.73	1.02	1.32	1.22
γ^5_{15}	-1.87	2.91	-0.72	0.6	-0.39
γ^5_{25}	3.85	5.9	1.4	1.09	0.41
γ^5_{35}	3.28	5.53	0.73	0.99	0.45
γ^5_{45}	-5.92	8.98	-1.86	1.71	-0.44
γ^5_{123}	6.06	1.67	3.97	0.84	1.12
γ^5_{124}	3.64	4.73	0.71	1.14	0.6
γ^5_{134}	3.65	4.74	1.24	1.16	0.49
γ^5_{234}	-5.17	8	-0.48	1.76	-0.57
γ^5_{125}	-1.11	3.28	-0.06	0.91	-0.31
γ^5_{135}	-6.48	6.87	0.77	1.46	-1.03
γ^5_{235}	-4.94	7.13	1.07	1.47	-0.83
γ^5_{145}	19.12	11.36	-1.1	2.28	1.75
γ^5_{245}	-2.68	2.78	-0.96	0.9	-0.59
γ^5_{345}	6.04	9	2	1.65	0.44
γ^5_{1234}	-1.6	7.06	1.66	1.53	-0.45
γ^5_{1235}	-1.99	13.2	-2.33	2.74	0.03
γ^5_{1245}	5.23	4.84	1.6	1.2	0.73
γ^5_{1345}	-0.99	11.28	-2.03	2.13	0.09
γ^5_{2345}	4.22	9.64	-2.54	1.97	0.69
γ^5_{12345}	-10.72	16.49	3.26	3.37	-0.83

Notes: The "SUR" column reports γ coefficients and standard errors estimated by the SUR. The "Structural" column reports the value of γ implied by the structural parameter estimates and minimum distance restrictions, as well as standard errors calculated using the delta method. The "Diff t" column reports the t-statistic for the difference between the two.

D Two-Period Model Equations

Consider the generalized income equations with learning in the case of 2 periods:

$$
\begin{aligned}
Y_{i1} &= \alpha_1 + \beta D_{i1} + (m_{i0} + \omega_{i1})(1 + \phi D_{i1}) + v_{i1} \\
Y_{i2} &= \alpha_2 + \beta D_{i2} + (m_{i0} + \tilde{m}_{i,1} + \omega_{i2})(1 + \phi D_{i2}) + v_{i2}
\end{aligned}
$$

The projections are then:

$$
\begin{aligned}
m_{i0} &= \lambda_0 + \prod_{k=1}^{2}(1 - \lambda_k D_{ik}) - 1 + \psi_{i0} \\
m_{i0} &= \lambda_0 + \lambda_1 D_{i1} + \lambda_2 D_{i2} + \lambda_1 \lambda_2 D_{i1} D_{i2} + \psi_{i0} \\
m_{i0} &= \lambda_0 + \lambda_1 D_{i1} + \lambda_2 D_{i2} + \lambda_{12} D_{i1} D_{i2} + \psi_{i0} \\
\tilde{m}_{i1} &= \theta_{20} + \theta_{22} D_{i2} + \psi_{i1}
\end{aligned}
$$

Replacing the projections in the income equations and grouping terms allows us to obtain the following reduced form equations:

$$
\begin{aligned}
Y_{i1} =& \alpha_1 + \beta D_{i1} + (\lambda_0 + \lambda_1 D_{i1} + \lambda_2 D_{i2} + \lambda_{12} D_{i1} D_{i2} + \psi_{i0} + \omega_{i1})(1 + \phi D_{i1}) + v_{i1} \\
Y_{i1} =& \overbrace{\alpha_1 + \lambda_0}^{\gamma_0^1} + \overbrace{[\beta + (1+\phi)\lambda_1 + \lambda_0 \phi]}^{\gamma_1^1} D_{i1} + \overbrace{[\lambda_2]}^{\gamma_2^1} D_{i2} + \overbrace{[(1+\phi)\lambda_{12} + \lambda_2 \phi]}^{\gamma_{12}^1} D_{i1} D_{i2} \\
&+ \underbrace{(\psi_{i0} + \omega_{i1})}_{\perp D_{i1}} \phi D_{i1} + \underbrace{\psi_{i0} + \omega_{i1} + v_{i1}}_{u_{i1}} \\
Y_{i1} =& \gamma_0^1 + \gamma_1^1 D_{i1} + \gamma_2^1 D_{i2} + \gamma_{12}^1 D_{i1} D_{i2} + u_{i1} \qquad (16)
\end{aligned}
$$

$$
\begin{aligned}
Y_{i2} =& \alpha_2 + \beta D_{i2} + \\
& (\lambda_0 + \lambda_1 D_{i1} + \lambda_2 D_{i2} + \lambda_{12} D_{i1} D_{i2} + \psi_{i0} + \theta_{20} + \theta_{22} D_{i2} + \psi_{i1} + \omega_{i2})(1 + \phi D_{i2}) + v_{i2} \\
Y_{i2} =& \overbrace{\alpha_2 + \lambda_0 + \theta_{20}}^{\gamma_0^2} + \overbrace{[\lambda_1]}^{\gamma_1^2} D_{i1} + \overbrace{[\beta + (1+\phi)(\lambda_2 + \theta_{22}) + \phi(\lambda_0 + \theta_{20})]}^{\gamma_2^2} D_{i2} + \\
& \underbrace{[(1+\phi)\lambda_{12} + \lambda_1 \phi]}_{\gamma_{12}^2} D_{i1} D_{i2} + \underbrace{(\psi_{i0} + \psi_{i1} + \omega_{i2})}_{\perp D_{i2}} \phi D_{i2} + \underbrace{\psi_{i0} + \psi_{i1} + \omega_{i2} + v_{i2}}_{u_{i1}} \\
Y_{i2} =& \gamma_0^2 + \gamma_1^2 D_{i1} + \gamma_2^2 D_{i2} + \gamma_{12}^2 D_{i1} D_{i2} + u_{i2} \qquad (17)
\end{aligned}
$$

Equations (16) and (17) are estimated by a seemingly unrelated regression which allows us to recover estimates for the γ coefficients. We then estimate the structural parameters through minimum distance where the minimum distance restrictions are:

$$
\begin{aligned}
\gamma_1^1 &= \beta + (1+\phi)\lambda_1 + \lambda_0 \phi \\
\gamma_2^1 &= \lambda_2 \\
\gamma_{12}^1 &= (1+\phi)\lambda_{12} + \lambda_2 \phi \\
\gamma_1^2 &= \lambda_1 \\
\gamma_2^2 &= \beta + (1+\phi)(\lambda_2 + \theta_{22}) + \phi(\lambda_0 + \theta_{20}) \\
\gamma_{12}^2 &= (1+\phi)\lambda_{12} + \lambda_1 \phi
\end{aligned}
$$

The exact same steps are taken to obtain estimates for the γ parameters and the minimum distance restrictions for the 5-period model.

E Nested Models

In this section, we show how our preferred model, a DCRC model of heterogeneous returns to enterprise with imperfect information, nests the following restricted models: heterogeneous returns with perfect information and a simple fixed effects model with homogeneous returns and perfect information. We describe the coefficient restrictions that need to be made to the DCRC model to arrive at each of the nested models.

E.1 Heterogeneous Returns with Perfect Information: CRC

With perfect information, the model becomes a static CRC model. The estimating equation is nearly the same as in the DCRC model:

$$Y_{it} = \alpha_t + \beta D_{it} + \eta_i(1 + \phi D_{it}) + v_{it}.$$

However, now the household is assumed to have perfect information about its relative productivity, η_i; hence, there is no longer an additive productivity shock, ε_{it}. Therefore, the relationship between η_i and the history of sectoral choices is static. Note, however, that v_{it} could still include exogenous, transitory shocks that shift households from period to period above and below the cutoff for enterprise entry. That is, households will sort into a particular enterprise history on the basis of η_i and their expectations of Y_{it}^N and Y_{it}^E; however, these expectations will not evolve over time as they do in the imperfect information case.

Accordingly, we need only a single projection in which we project η_i onto the sectoral choice dummies and all of their interactions, as in equation (10):

$$\eta_i = \lambda_0 + \prod_{k=1}^{5}(1 + \lambda_k D_{ik}) - 1 + \psi_{i0}.$$

Because households no longer update their expectations over time, the cumulative updates $\tilde{m}_{it}$ are irrelevant, which means that the θ coefficients in equation (11) are all equal to zero. The CRC model is therefore a restricted version of the DCRC model where all θ coefficients are assumed to be zero. This model has 33 (instead of 43) structural parameters that we estimate from 155 reduced form coefficients (γ) using minimum distance.

E.2 Homogeneous Returns with Perfect Information: CRE

The most restricted model imposes both that returns to enterprise are homogeneous and that households have perfect information about their earnings in both sectors. That is, the only source of heterogeneity is additive and fixed over time. This amounts to assuming that the data generating process is a simple household fixed effects or CRE model. Under these assumptions, the estimating equation becomes

$$Y_{it} = \alpha_t + \beta D_{it} + \eta_i + v_{it}.$$

We now need only a single projection of η_i on the five sectoral choice dummies:

$$\eta_i = \lambda_0 + \lambda_1 D_{i1} + \lambda_2 D_{i2} + \lambda_3 D_{i3} + \lambda_4 D_{i4} + \lambda_5 D_{i5} + \psi_{i0}.$$

Note that we have not included the interactions of sectoral choice dummies across periods. This is because, once we assume that η_i has no effect on the return to enterprise, the changes in choices over time will no longer depend on the initial belief, though the choice in each period still will. As in the CRC model above, all θ coefficients are assumed to be equal to zero. Therefore, the CRE model is a restricted version of the DCRC model where ϕ, all θ coefficients in equation (11), and all λ coefficients in equation (12) – except for λ_1, λ_2, λ_3, λ_4, λ_5 – are assumed to be zero. This model has 5 structural parameters which we estimate from 25 reduced form parameters using minimum distance.

F Standard Errors

In Figures 4 and A2, we report error bars for average returns ($\beta + \phi m_{it}$) across various combinations of household types and waves. In this section, we describe how we obtain the required standard errors.

We denote estimated average returns for a particular group of households in a particular wave as $\hat{f}$. To estimate $\hat{f}$, we use estimates of the parameters β, ϕ, and some combination of the λ and θ parameters that are required to estimate m_{it}. In short, $\hat{f}$ is a non-linear function of estimated parameters and household decisions D_{it}. We define

$$\hat{f} = \frac{1}{N}\sum_{i=1}^{N} h(X_i, \hat{\rho}),$$

where $\hat{\rho}$ represents a vector of the estimated structural parameters, X_i is vector of household i's sectoral decisions, and $h(.)$ is a continuous and differentiable function. We can define $\tilde{f}$ as the sample average return calculated using the true parameter vector (ρ_0):

$$\tilde{f} = \frac{1}{N}\sum_{i=1}^{N} h(X_i, \rho_0),$$

and the population average return as

$$f = E[h(X, \rho_0)],$$

where the expectation is over the joint distribution of X.

If we decompose the difference between the estimated $\hat{f}$ and the population parameter f into two parts:

$$(\hat{f} - f) = (\hat{f} - \tilde{f}) + (\tilde{f} - f),$$

then it can be shown that the variance of $(\hat{f} - f)$ is the sum of two terms: the variance of $(\tilde{f} - f)$ and $(\hat{f} - \tilde{f})$ (See Molina (2016) for details). Specifically,

$$\text{Var}(\hat{f} - f) = \frac{\sigma^2}{N} + \frac{s^2}{N},$$

where (using the delta method)

$$\frac{\sigma^2}{N} = \frac{1}{N} E\left[\nabla h(\rho_0)\right]' V E\left[\nabla h(\rho_0)\right]$$

and

$$\frac{s^2}{N} = \frac{1}{N} \text{Var}(h(X, \rho_0)).$$

G Data Appendix

G.1 Selecting Household Characteristics

As described in section 4.2, we use lasso to select a set of household-level predictors of returns to enterprise from a wide range of variables. Below, we describe all 27 variables included in the lasso.

- Years of educational attainment (average and maximum): We calculate both average and maximum educational attainment across all household members.

- Raven's test z-score (average and maximum): The IFLS administered a test of cognitive ability (which included questions from the Raven's test of fluid intelligence as well as a few math questions). Different versions of the test were given to respondents aged 7-14 and 15-59. We calculate the version-specific z-score for each respondent and average across all household members. We also calculate the maximum.

- Risk aversion score (average and maximum): This is a five-point score generated from a set of five questions asked of those aged 15 and older, where a score of 5 represents the highest level of risk aversion. Each question offers two hypothetical options: receiving 4 million rupiah for certain, or a lottery with a higher expected value. We calculate the average and maximum score across all household members.

- Height (average and maximum): The IFLS measures height for all household members. Restricting to adults aged 20-65, we standardize height separately for men and women. We calculate the average and maximum z-score across all adults.

- Self-reported health (average and maximum): All respondents aged 15 and older are asked whether they consider themselves very healthy, somewhat healthy, somewhat unhealthy, or unhealthy. We assign a 4 to very healthy and 1 to unhealthy, and calculate both the average and maximum.

- Share of very healthy adults: We calculate the share of household members aged 15 and older who consider themselves very healthy.

- Share of somewhat healthy adults: We calculate the share of household members aged 15 and older who consider themselves very healthy or somewhat healthy.

- Physical functioning (average and maximum): The IFLS asks all respondents aged 15 and older whether they can easily, can with difficulty, or cannot at all do 23 physical activity tasks (including activities of daily living, instrumental activities of daily living, and other physical tasks). We calculate the share of activities a respondent "can easily" do. We then calculate the average share and maximum share for each household.

- Mental health score (average and maximum): To measure mental health (for respondents aged 15 and older), the IFLS includes a 10-question version of the CES-D questionnaire designed to help identify clinical depression. We sum the responses to all 10 questions, which generates a score ranging from 0 to 30 points, where higher numbers are associated with a higher severity of depressive symptoms. We calculate the average and maximum score for each household.

- Share of members with depressive symptoms: Using the 10-question CES-D questionnaire described above, we calculate the share of (adult) household members with a score of 10 or greater, a cutoff that is used as an indicator of significant depressive symptoms (Zhang et al., 2012).

- Big 5 personality traits (open-mindedness, conscientiousness, extraversion, agreeableness, negative emotionality – average and maximum): The IFLS includes the Big Five Index 15 (BFI 15), a set of 15 questions about the respondents' personality, three for each of the five personality traits. We use these to create a five-point score for each of the five personality traits. We calculate the average and maximum score for each household, for each of the five personality traits.