Spatial Information Frictions in Rural-Urban Migration*

Tasso Adamopoulos
York University

Zahra Shahidi
York University

November 2026

Abstract

We study the role of information frictions for the low and slow out-migration of workers from rural to urban areas in developing countries. Using panel data from Indonesia we document that migration flows are highly persistent across space and that current district-level migration is correlated with incomes of previous migrants from the same districts. Using a difference-in-differences design and exploiting natural disaster events at origin as push-shocks, we show that rural districts with improved internet access exhibit higher migration to urban areas. We develop a quantitative spatial model of imperfect information with noisy and dispersed information across rural districts about urban earnings opportunities. Our theory features localized social learning, where rural residents learn about urban opportunities from previous migrants from their own districts. We estimate the model to the panel data from Indonesia and identify information frictions using indirect inference, with easily computable second moments in micro-level datasets. We allow for other migration frictions to disentangle the role of information per se. We find that our theory can account for low aggregate migration and the spatial persistence of migration over time. Information frictions account for about 1/5 of total migration costs and generate a welfare loss of 2.47 percent for rural workers. We validate the learning mechanism and the model's quantitative predictions using elicited beliefs from experimental evidence in Kenya.

JEL classification: O11, O13, O40, Q10, R23.

Keywords: information frictions, rural-urban migration, welfare, development.

*Contact: Tasso Adamopoulos, aadamo@yorku.ca; Zahra Shahidi, zhash@yorku.ca. We have also benefited from comments at the SED in Copenhagen, the EEA in Dublin, the CEA in Montréal and Vancouver, and York University. Adamopoulos gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

Consider a young worker in rural Java who is thinking about moving to Jakarta. She may know little about what she would earn there, relying instead on stories from a cousin who left a few years ago or a neighbor's son. Her decision depends on what these experiences lead her to believe about urban opportunities. When information travels through such local connections, workers in different rural locations may have very different perceptions of the gains from moving.

These individual decisions shape a central feature of economic development: the transition of workers out of agriculture and rural locations and into manufacturing and services in urban locations (Harris and Todaro, 1970; Herrendorf et al., 2014). Developing countries, however, have been slow in making this transition, with large shares of the population still living and working in rural areas and in agriculture, where productivity is low (Gollin et al., 2002; Caselli, 2005; Restuccia et al., 2008; Adamopoulos and Restuccia, 2014, 2022). This occurs despite substantial rural-urban income and consumption gaps (Lagakos, 2020), a major source of inequality in developing countries.[1] There are many potential explicit and implicit barriers that inhibit rural to urban migration. Understanding what these barriers are has important implications for welfare, structural change, development and policy design.

We study the role that information frictions play in this slow transition. Rural workers may stay where they are because they know little about what they could earn in the city, and experimental evidence shows that rural households underestimate urban earnings and migrate more when they are given information about them (Baseler, 2023; Barnett-Howell et al., 2026). We ask how much such frictions matter for aggregate migration, for its pattern across space, and relative to other barriers to labor mobility. We find that with perfect information migration would be about 40 percent higher and much less persistent across districts, and that information frictions account for about one-fifth of the migration response to removing all barriers.

The literature has studied a large set of other explicit or implicit frictions, such as poor road networks (Asher and Novosad, 2020; Morten and Oliveira, 2024; Adamopoulos, 2025), land use rights (De Janvry et al., 2015; Adamopoulos et al., 2024), lack of insurance markets (Munshi and Rosenzweig, 2016), travel cost and migration risk (Bryan et al., 2014), migration subsidies as insur-

[1] These gaps go together with differences in labor productivity between agriculture and non-agriculture, known as agricultural productivity gaps (Young, 2013; Gollin et al., 2014; Herrendorf and Schoellman, 2018). While there are measurement issues, Gollin et al. (2014) show that large agricultural productivity gaps remain even after controlling for observable factors such as hours of work, human capital and cost of living. Mis-measurement is part of the story (Alvarez-Cuadrado et al., 2019), but migrants still see an increase in their income (Pulido and Swiecki, 2021) and consumption (Beegle et al., 2011; Bryan et al., 2014), and the gaps are persistent, with past sectoral exports predicting current gaps (Gollin et al., 2025). See Lagakos (2020) for an excellent summary of this literature.

ance (Lagakos et al., 2023), risk sharing and community network spillovers (Morten, 2019; Meghir et al., 2022), market access risk (Brooks and Donovan, 2020), and amenities (Bryan and Morten, 2019). Much of this important work uses a micro-to-macro approach of combining structural models with observational or experimental micro-level evidence. We also combine micro evidence with a structural model but study a distinct friction, the lack of information about true urban earnings opportunities.

Using micro-level data from Indonesia we construct a district-level panel and document two observations about permanent migration. First, as in the cross-country evidence of Adamopoulos and Shahidi (2025), migration flows are geographically concentrated and highly persistent over time: the top fifth of districts accounts for 57 percent of migration flows, and the probability that an initially high migration district will remain high migration later on is 0.76. Second, subsequent district-level migration is positively correlated with the urban incomes of previous migrants of the same district, with a correlation coefficient of about 0.24. We then employ a difference-in-differences approach and exploit natural disaster events at origin as push-shocks,[2] to show that rural districts with improved internet access exhibit higher migration to urban areas and an amplified migration response to natural disasters. The premise is that internet access allows rural individuals not only to directly access job ads but also to connect on social media with family and friends in urban locations and learn from them.

Motivated by this evidence we build a spatial equilibrium model of rural-urban migration in which information about urban earnings is imperfect and spatially dispersed. Rural workers do not observe urban productivity, so they make their migration decision based on their *perceived* return rather than the actual one. The theory features localized social learning: workers learn from the urban earnings of previous migrants from their own district. These earnings are a noisy signal, because they confound urban productivity with the migrants' own ability, and the signal is more precise the more previous migrants the district has.[3] To account for the heterogeneity in migration decisions across individuals within rural districts we allow for idiosyncratic uncertainty, as captured by a noisy private signal about urban earnings opportunities. Because workers are risk averse, the uncertainty that remains after learning lowers the perceived return to migration. We also allow for other district-specific barriers to labor mobility, which capture all determinants of migration other than information. This allows us to disentangle the role of information frictions relative to all other labor mobility frictions across space.

[2] Gröger and Zylberberg (2016) show how migration can serve as a coping mechanism after catastrophic shocks.

[3] A further possible source of imprecision, on which we do not explicitly put structure, is that previous migrants may deliberately mis-represent their earnings to avoid the "community tax," the social pressure to share income with relatives and other community members (Squires, 2024; Boltz et al., 2020; Jakiela and Ozier, 2016; Boltz et al., 2019). Baseler (2023) finds such income hiding among Kenyan migrants.

The theory has two implications for the spatial pattern of migration. First, what workers believe about urban earnings depends on their district's own migration history, so districts that face the same urban opportunities can differ in their migration. Second, our framework endogenizes imperfect sorting of workers across space through heterogeneous beliefs. Under full information, unidimensional ability implies perfect sorting. With information frictions, high-ability but pessimistic workers may remain in rural areas, while low-ability but optimistic workers may migrate, generating spatial misallocation of talent. Whereas a large literature attributes part of urban-rural wage gaps to the efficient sorting of individuals across sectors on the basis of observable or unobservable characteristics (Lagakos and Waugh, 2013; Herrendorf and Schoellman, 2018; Young, 2013; Hamory et al., 2021; Alvarez, 2020; Adamopoulos et al., 2022), in our framework imperfect sorting arises endogenously from heterogeneity in beliefs and learning dynamics.

We estimate the model to match the micro-level data and cross-district moments from Indonesia. Because beliefs are not observed, we discipline the extent of information frictions using indirect inference, with easily computable second moments from the micro-level data: the correlation across districts between the earnings of previous migrants and subsequent migration, and the dispersion of rural-urban wage gaps among migrants from the same district. District-specific barriers are chosen so that the model matches observed district-level migration, and we measure the role of information by removing information frictions while keeping these barriers in place.

Moving to perfect information raises aggregate migration from 10.5 to 14.7 percent and lowers persistence, the correlation of district migration flows over time, from 0.80 to 0.26. The second-period consumption-equivalent gain is 2.47 percent for the rural-origin population, and real GDP per capita rises by 1.4 percent. Most of these effects come from correcting workers' beliefs, while removing the penalty that risk-averse workers attach to uncertainty accounts for a small part. The increase in migration is about one-fifth of the increase obtained when all migration barriers are removed. Our results thus relate to the literature on migration costs across space (Tombe and Zhu, 2019; Caliendo et al., 2019; Bryan and Morten, 2019; Eckert and Peters, 2022). While this literature is concerned with the overall migration costs, we put structure on one particular factor, information frictions. For policy, our results imply that giving rural households credible information about urban earnings can raise migration and welfare.

We validate the learning mechanism and quantitative results using reported beliefs and a randomized information experiment in Kenya. Rural workers there underestimate urban earnings by 40 percent on average, and those with greater exposure to migrants underestimate them less, as local learning implies. Estimating information frictions in Kenyan villages directly from reported beliefs yields migration, persistence and welfare effects close to those obtained by indirect inference.

Both calibrations closely match the untargeted experimental migration response, with an elasticity of 1.88 compared with 2.00.

We thus offer a framework for measuring information frictions in migration when beliefs are not observed. Localized learning leaves a trace in standard migration and earnings moments, and we use that trace to quantify the friction. The agreement between the two Kenyan calibrations supports the use of this approach in other settings where such moments are available but beliefs are not. In this respect, our paper is related to a large literature in economics that studies the role of information frictions for a variety of macroeconomic phenomena and aggregate outcomes, such as resource misallocation across firms and aggregate productivity (David et al., 2016), international trade (Allen, 2014), inflation forecasts (Coibion and Gorodnichenko, 2015), female labor force participation (Fogli and Veldkamp, 2011), asset pricing (Veldkamp, 2006), business cycles (Angeletos and La'O, 2013), among many others. Closest to our approach are David et al. (2016), who also infer information frictions from observed outcomes rather than from measured beliefs, and Fogli and Veldkamp (2011), where agents learn from those near them, so that information differs across locations.

Our paper also complements the experimental evidence on information and migration. Baseler (2023) shows that rural families in Kenya systematically underestimate the urban earnings of their own family members, and that providing rural workers with earnings information in nearby cities leads to a 40 percent increase in migration to the capital and an improvement in welfare. Barnett-Howell et al. (2026) find, in a cluster-randomized experiment across five Kenyan counties, that providing information on urban earnings privately, through a mentor in the city, or at a public village meeting raises beliefs about the returns to migration, with each delivery mode increasing migration to Nairobi by about 2 percentage points on a base of 17%. Bryan et al. (2014), in an experiment in rural Bangladesh, find that providing households with the value of a return bus ticket increases temporary migration and consumption, but simply providing households with information about urban opportunities does not. Interestingly, households that did migrate were more likely to migrate later even after the transport subsidy was no longer offered. One possible reading of this evidence is that first-hand experience may provide more direct and credible information about urban opportunities than information from non-migrants, consistent with the idea of learning. These experiments establish that information changes beliefs and migration where it is provided; we measure how much information frictions matter in the aggregate, across space, and relative to other barriers.

The rest of the paper is organized as follows. Section 2 describes the data and the empirical evidence. Section 3 presents the model, and Section 4 its estimation and the counterfactual

experiments. Section 5 validates the mechanism with the Kenyan data, and Section 6 concludes.

2 Data

We use individual-level survey data from the Indonesian Family Life Survey (IFLS), and community-level survey data from the accompanying Community and Facility (COMFAS).[4] The IFLS is a longitudinal survey, particularly suitable to studying migration because it makes a deliberate effort to track respondents that have moved. We use multiple waves of the data, corresponding to years 1997, 2000, 2007, 2014. The IFLS data contain information on a rich set of individual characteristics and include a migration questionnaire, where individuals are surveyed on their migration histories. This allows us to extract detailed information on individual labor supply, residence location, migration, rural and urban incomes.

Migration data In each year of the survey, we focus on *permanent* migrations (greater than 6 months) of individuals that are currently living in urban locations, have moved at least once since age 12, and that were above 15 years old when they moved. In each year of the survey we capture only the migrations that have occurred since the last time surveyed, which precludes double counting over time. We determine their previous location as the location they were at the age of 12 if they moved only once, and as the destination of the previous migration, before moving to the current urban location. This approach accurately captures the locations migrants have come from by survey year.

Natural disaster data We use two sections from the individual-level survey of the IFLS to construct our natural disaster variable: (i) "MIGRATION" questionnaires, where individuals are asked about the reasons for their migration, where they can answer natural disaster; (ii) "NATURAL DISASTER" questionnaires, where individuals are asked explicitly about the hardships they faced from natural disasters. We focus on natural disaster events of the following type: flood, landslide/mudslide, mudflow, volcanic eruption, earthquake, tsunami, windstorm, forest fire, fire, drought. A district is classified as having faced a natural disaster if at least one household in that district reported such an event, and our variable $Disasters_{it}$ is the total number of such disasters in district i in period t.

[4] https://www.rand.org/well-being/social-and-behavioral-policy/data/FLS/IFLS/study.html

Internet access data We use answers of village officials to the COMFAS community questionnaires, who are asked about the year in which the village office first received internet access. We use this as the year of the infrastructure investment for internet connectivity as the investment is a prerequisite for internet access by the village office, which is among the first to receive access.

We conduct our analysis at the province level (first administrative level in Indonesia) and the district (or *kabupaten*) level (second administrative level in Indonesia).

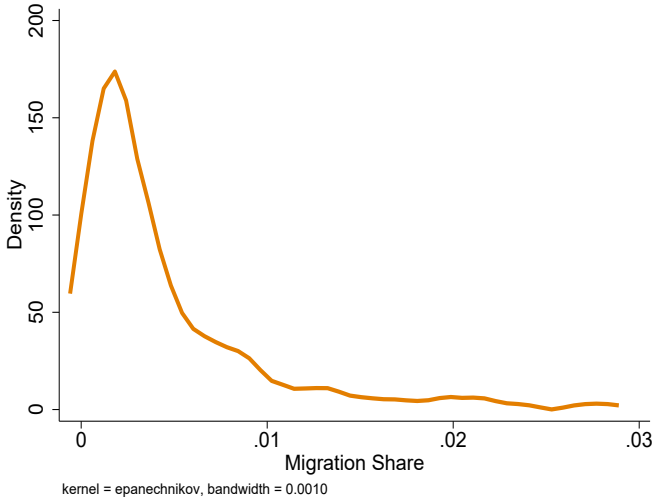
2.1 Empirical Observations

We aggregate individual rural-urban migration transitions at the district level (*kabupaten*) to trace out-migration from rural districts to urban centers. This allows us to construct a district-level panel of rural-to-urban migration. Let M_i be the number of migrants from district i, and $M = \sum_i M_i$ as the total number of migrants. Then we define the migration flow as M_i/M, which captures which districts are major sources of migrants at the national level, i.e., the national migration contribution share.

1. Geographic Clustering of Migration Flows To examine the degree of concentration of migration flows across districts within Indonesia we focus on two measures. First, we consider a kernel density estimate of the spatial distribution of migration flows across districts. Figure 1 shows the distribution of migration flows across districts in Indonesia (2nd administrative level). The density illustrates that there is a handful of districts from which most of the migration flows emanate, while from other districts there is little or no migration occurring. Second, we rank districts by their migration flows and sort them in quintile bins. The share of migration flows accounted for by the top 20% districts is 56.8%. These observations indicate considerable geographic clustering of migration flows across districts within Indonesia.

2. Spatial Persistence of Migration Flows Over Time To gauge the persistence in migration flows across space within Indonesia we compute a transition matrix from initial to subsequent migration flows. We first order districts according to the initial migration flows and sort them into quintiles of the distribution of initial migration flows (five bins). Then we sort districts according to their subsequent migration flows, and organize them into quintiles of the subsequent distribution of migration flows. An entry in the transition matrix reflects the probability that a district moves across bins of the migration flow distribution from "today" to "tomorrow." High values for the diagonal elements in the transition matrix indicate that districts tend to stay in their initial quintiles of migration flows over time. Table 1 displays the transition matrix for districts in Indonesia. The

Figure 1: Distribution of Migration Flows Across Districts



Notes: The figure depicts a kernel density estimate of migration shares across districts within Indonesia. "Migration Share" refers to the share of a district's migration flow in the total migration flows within the country. The kernel is an Epanechnikov kernel. Based on individual micro-level migration data from IFLS.

first diagonal element from the top shows the probability that a high migration district in the initial period is a high migration district in the subsequent period. This probability is 0.76, indicating high persistence of migration flows over time for high migration districts. High persistence is observed also at the bottom end of the transition matrix. There tends to be a bit more mobility though in the middle of the distribution.

Table 1: Transition Matrix of Migration Flows

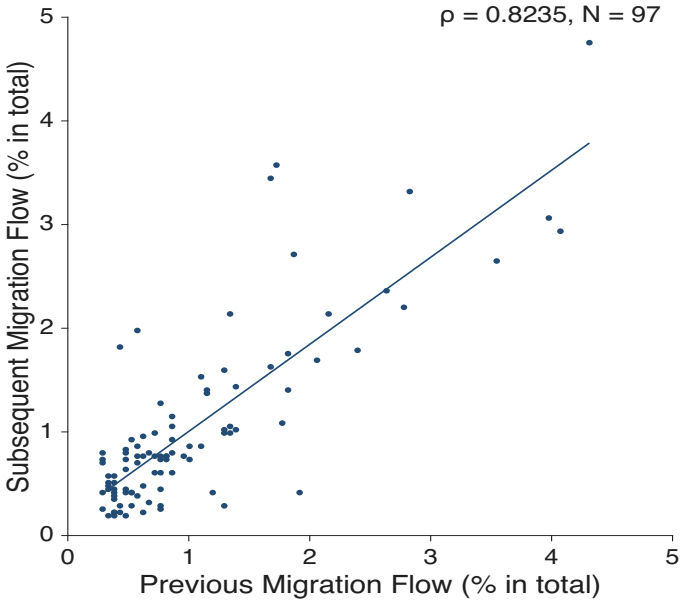
	Q1	Q2	Q3	Q4	Q5
Q1	0.76	0.22	0.02	0.00	0.00
Q2	0.16	0.34	0.32	0.16	0.02
Q3	0.07	0.27	0.30	0.25	0.11
Q4	0.02	0.14	0.20	0.32	0.32
Q5	0.00	0.02	0.16	0.27	0.55

Note: Each element shows the probability that a district within Indonesia moves from one quintile in the initial period to another in the subsequent period. Diagonal elements show the probability that districts remain in the same quintiles over time. Estimated from IFLS data.

Figure 2 shows migration flows by district, as a share of total migration, between 2007 (vertical

axis) and 2000 (horizontal axis). There is substantial persistence over time, with a correlation coefficient of 0.82: new migration occurs primarily from the same districts as old migration. This correlation may reflect systematic differences across districts in fundamentals, such as size, lack of opportunities, skills, or composition of the labor force. In Figure 3 we control both previous and subsequent migration flows for observable characteristics, population, income, education, age, gender, sector and time fixed effects. The persistence pattern cannot all be explained away by fundamentals, as migration flows continue to be highly correlated across periods.

Figure 2: Spatial Persistence of Migration Flows

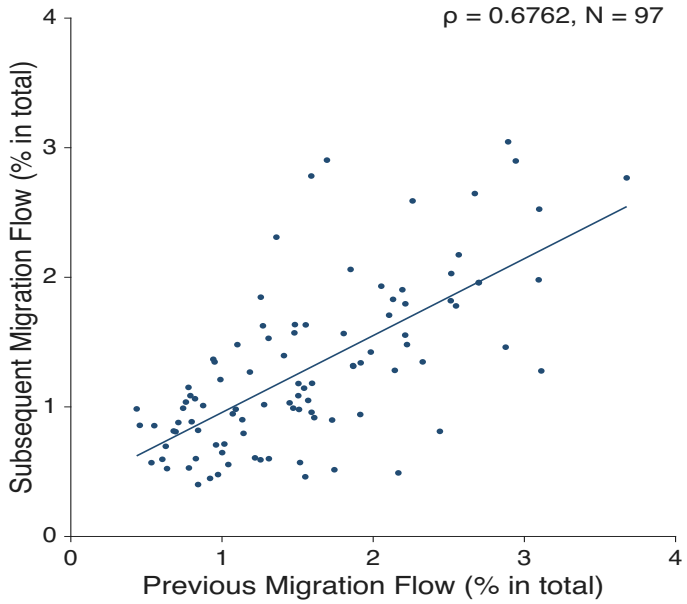


Notes: The y-axis represents the share of migrants from a district as a share in total migration over 2007 (period t). The x-axis represents the share of migrants from a district as a share in total migration over 2000 (period $t-1$).

One possible explanation behind the persistence of migration across districts is that migration histories generate spatially uneven information networks about urban opportunities. We next motivate this channel empirically using internet connectivity across districts as a proxy for information access.

3. Correlation of Migration Flows and Previous Migrant Income As a measure of "responsiveness" of migration to urban income, we estimate the correlation of previous migrant income by administrative unit with the subsequent migration flows from that administrative unit. We assign individuals living and working in urban areas to rural administrative units if they migrated from those administrative units in the previous migration period (wave of the survey). Our mea-

Figure 3: Spatial Persistence of Migration Flows Controlling for Fundamentals



Notes: The y-axis represents the share of migrants from a district as a share in total migration over 2007 (period t). The x-axis represents the share of migrants from a district as a share in total migration over 2000 (period $t-1$). Migration flows in each period are controlled for population, income, education, age, gender, sector, and time fixed effects.

sure of income is monthly earnings, and we compute for each administrative unit the mean over all permanent migrants from that administrative unit. Subsequent migration flows from administrative units are the migration flows in the next wave of the survey. Table 2 shows the estimated correlation across provinces and across districts. They are both positive, with the province based one being higher at 0.30. However, as a caveat we note that there are fewer observations per district to estimate mean income in the district-level correlation.

Table 2: Correlation of Previous Migrant Earnings and Migration Flows

	Provinces	Districts
Correlation	0.30	0.24

Note: Previous migrant earnings for each administrative unit are the average earnings in urban areas of those that have migrated in the previous wave. Estimated from IFLS data.

While the above observations do not speak to information per se, in the next section we employ a difference-in-differences framework that connects migration flows to internet connectivity, as a

proxy for information access.

2.2 Difference-in-Differences Estimates

We estimate the effect of internet availability, as a proxy to information access, on out-migration flows across districts at the second administrative level (*kabupaten*). We employ a difference-in-differences specification with continuous treatment and staggered timing, as well as time and district fixed effects. The continuous treatment captures the number of villages within a district that receive internet access in a given period. The staggered timing allows for variation in the timing districts received internet access, as the infrastructure was not rolled out simultaneously for all districts. Time fixed effects control for aggregate time-varying factors affecting all districts in Indonesia. District fixed effects, control for observed and unobserved district characteristics with time invariant effects, allowing to exploit within district variation over time.

We consider two alternative difference-in-differences specifications. First, we examine directly the effect of the internet infrastructure investments on district migration. Second, we use district-level natural disaster events as push shocks for migration, and examine the role that internet access plays in mediating the extent of that migration.

We use the community-survey data from COMFAS on the first year the village administrative office received internet as the year the village internet infrastructure investment was completed (as without the infrastructure the village office could not get internet). We construct our internet access variable for district i at time t as,

$$InternetAccess_{it} = Access_{it} \times n_{it}$$

where $Access_{it}$ is a dummy variable that takes the value of 1 in the period that the district first receives internet access, and stays at that value thereafter. In other words, if a district gets treated with internet availability, it remains treated thereafter. n_{it} is a continuous variable that measures the number of villages within district i that received internet access for the first time in period t. This variable measures the intensity of treatment in each period.

We obtain estimates of the effect of internet access on migration (dependent variable) using the following difference-in-differences specification,

$$Migration_{it} = \alpha + \beta_0 Disasters_{it} + \beta_1 InternetAccess_{it} + \delta\left(Disasters_{it} \cdot InternetAccess_{it}\right) + \varphi_i + \iota_t + \varepsilon_{it} \tag{1}$$

where $Migration_{it}$ is the natural logarithm of number of migrants from district i in period t;

Table 3: Estimated Effects of Internet Access on Migration

	Dependent Variable (in logs): Migration	
	w/o Disaster shocks	w/ Disaster shocks
	(1)	(2)
$InternetAccess_{it}$	0.539^{***}	0.403^{***}
	(9.50)	(5.24)
$(Disasters_{it} \cdot InternetAccess_{it})$		0.0028^{*}
		(1.83)
$Disasters_{it}$		0.085^{***}
		(3.41)
Intercept	-0.79	-0.61
	(-0.15)	(-0.12)
District FE	yes	yes
Year FE	yes	yes
Observations	740	740
Adjusted R^2	0.81	0.81

Note: All columns contain estimates from OLS regressions of district log migration on treatment, including time fixed effects, and district fixed effects in each case. The estimates in column (1) abstract from natural disasters, while those in column (2) include them. The sample is a panel of districts with three different time periods, 1997-2000, 2000-2007, 2007-2014. t-statistics are in the parentheses, ***, **, and * represent significance at the 1% ($p < 0.01$), 5% ($p < 0.05$), and 10% ($p < 0.10$) level respectively.

$Disasters_{it}$ is a count variable that measures the number of natural disasters that occurred in district i in period t; ι_t are time fixed effects, and φ_i are district fixed effects. We consider three periods in our estimation, $1997-2000$, $2000-2007$, and $2007-2014$.[5] The key coefficient we are after without natural disaster shocks is β_1, and with natural disaster shocks the coefficient of the interaction term with internet access δ. β_1, captures the average response across districts to being exposed to new internet infrastructure investments. It measures the differential change in migration after obtaining internet access attributable to the extra number of villages that first receive internet. δ captures how large the effect of a natural disaster is on migration for districts that have relatively more internet access.

Table 3 reports the results from the difference-in-differences analysis. Columns (1) and (2) show the results of improved internet access for districts on the number of permanent migrants (in

[5]We do not consider the data before 1997 because only one survey village had internet access at that time.

logarithmic form). In column (1) we report the results in the absence of natural disaster shocks, while in column (2) the results with natural disasters as push shocks. The results show that improved internet access has a highly statistically significant positive effect. In other words, when the number of villages in a district that receive internet access for the first time increases, migration increases as well. Specifically, in column (1) a one-unit increase in the treatment variable (i.e., one more village receiving internet access) is associated with a 71% increase in migration outflows in that district, holding other factors constant. The effect of internet access with natural disaster shocks is similar, but the specification in column (2) allows us to see the effect of natural disasters mediated through internet access. We find that the effect of natural disasters on migration is larger in districts with more internet access. Specifically, for each additional village that receives internet access, the migration response to a natural disaster increases by 0.28% per disaster, amplifying the migration response to natural disasters.

Empirically, one challenge in identifying the causal effects of new internet connectivity on migration is the potential endogenous placement of internet infrastructure investments. The distribution of new internet connections across districts, their timing as well as their intensity, may not be random, but may depend on district proximity to the capital, their economic prospects or the political connections of their administrative authorities to decision making bodies.

Indonesia is an archipelago with a multitude of islands. While internet expansion was more pronounced on the islands of Java and Sumatra (ICT Watch Indonesia, 2014),[6] there was deliberate government effort for a fibre optic development program for the whole country (Nusantara 21 and Palapa Ring Program), particularly reaching remote rural and under-served communities (Setiawan et al, 2017).[7] Another initiative sought to expand internet provision using satellite connections (Subdistrict Internet Service Center (PLIK) and Mobile PLIK (MPLIK)). On net, it is not clear that there was bias in a particular direction. In addition, the district fixed effects in specification (1) account to some extent for permanent differences across districts.

3 A Spatial Model of Information Frictions and Migration

Consider a spatial two-period economy that consists of a finite number of multiple rural locations R, indexed by $r \in \{1, 2, ...R\}$, and a single urban location, indexed u. Time is discrete and is indexed by $t = \{0, 1\}$. There is a single good in each period that can be produced in any of the R rural

[6] "A Brief Overview of Indonesia Internet Landscape, ICT Watch Indonesia, 2014, https://www.gp-digital.org/.

[7] "Development of the Information and Communication Technology Service Industry in Indonesia," Ahmad Setiawan, Ms Kautsarina, Onny Rafizan, Ashwin Sasongko Sastrosubroto, AJTDE - Vol 5, No 3 - August 2017, https://telsoc.org/journal/ajtde-v5-n3/a96.

locations or the urban location. The good produced in each location is consumed locally and there is no inter-regional trade. The economy is populated by a finite number of individuals N, indexed by $i \in \{1, 2, ..., N\}$, that inelastically supply one unit of labor and consume in the location where they reside each period. Individuals can migrate from a rural location to the urban location over time.

Individual heterogeneity Individuals are identical in terms of their rural ability, which we normalize to 1 for all individuals in any rural location $r \in \{1, 2, ...R\}$. However, individuals are heterogeneous with respect to their permanent ability in the urban location, denoted s^i for individual i. In each rural location r, and the urban location, urban ability is drawn from a known distribution with cdf $F(s)$, that is common across locations. In other words, locations do not differ a priori with respect to the potential urban skills of their population. We assume that $F(\cdot)$ is a log-normal distribution with mean μ_s and variance σ_s^2, that is, $s^i \sim L\mathcal{N}(\mu_s, \sigma_s^2)$ with $e^i \equiv \log s^i - \mu_s \sim \mathcal{N}(0, \sigma_s^2)$. The distribution of skills is fixed over time.

Initial distribution of population At time $t = 0$ there is an exogenous initial distribution of the population across locations. Let $N_{u,0}$ be the number of individuals in the urban location and $N_{\bar{r},0} = N - N_{u,0}$ be the total population across all rural locations combined, where $\bar{r}$ denotes the total rural sector. The initial number of individuals in each rural location r is $n_{r,0}$. Rural locations however differ exogenously in the initial number of individuals in the urban location originating from each rural location. In particular, denote by $M_{r,0}$ the number of urban workers at time $t = 0$ that have previously (outside the model) migrated from rural location $r \in \{1, 2, ...R\}$. The entire initial urban population at $t = 0$ has previously migrated from one of the rural locations. In other words,

$$\sum_{r=1}^{R} M_{r,0} = N_{u,0}$$

We take the initial distribution of previous migrants in rural location of origin $\{M_{r,0}\}_{r=1}^{R}$ as given in what follows. Note however that the distribution of migrants in the second period, at time $t = 1$, is endogenous based on the dynamic migration decisions at time $t = 0$. Given the timing, in our framework the migration decision is a one-time irreversible decision, i.e., if an individual migrates to the urban location they stay there. Also there is no migration between rural locations, i.e., all migration is uni-directional to the urban location. To isolate the effect of different numbers of previous migrants from different quality we assume that migrants from every rural location do not differ systematically with respect to their skills, i.e., the mean ability and its dispersion for previous migrants is the same for each location r, and equal to the population mean μ_s and

dispersion σ_s. We make this assumption to keep the exposition of the learning problem simple. In the quantitative analysis, we account for the fact that previous migrants are a selected group, as explained in Section 4.1.

Production In each location there is a large number of homogeneous, perfectly competitive firms producing the economy's single good. In each period t, the representative firm in rural location $r \in \{1, 2, ..., R\}$ produces the single good according to a constant returns to scale technology that is linear in labor,

$$Y_{r,t} = A_r N_{r,t}$$

where $N_{r,t}$ is labor hired and A_r is the constant productivity in rural location r. Rural locations differ in their TFP A_r, which implies differences in *actual* income gaps and agricultural productivity gaps across space. The representative firm in the urban location produces the single good according to a technology that is linear in effective units of labor,

$$Y_{u,t} = A_u S_{u,t}$$

where A_u is urban productivity, and,

$$S_{u,t} = S_{u,t-1} + \sum_{r=1}^{R} \sum_{i \in H_{r,t}} s^i$$

is the total of the previous effective labor and the sum of effective units of labor from all R rural locations over the set of individuals that have migrated from each rural location r. $H_{r,t}$ denotes the set of individuals i that migrate to the urban location from rural location r at time t. Given this notation we can re-write the number of individuals working in the urban location in each period as,

$$N_{u,t} = N_{u,t-1} + \sum_{r=1}^{R} \sum_{i \in H_{r,t}} 1$$

Note that $S_{u,-1} = N_{u,-1} = 0$.

The representative firms in each rural location and the urban location face static profit maximization problems. The first order conditions to these problems imply that the wage rate in each location is equal to productivity in each period,

$$w_{r,t} = A_r$$

$$w_{u,t} = A_u$$

where $w_{r,t}$ is the wage rate in rural location r, and $w_{u,t}$ is the real wage per effective unit of labor in the urban location in period t. Note, that the representative firm in the urban location observes the (constant) urban productivity A_u when making production decisions. However, as explained next urban productivity is unobserved by workers.

Individual Preferences Individuals have preferences over the consumption of the single good in each period. The lifetime utility of individual i is given by,

$$u\left(c_0^i\right) + \beta u\left(c_1^i\right)$$

where $\beta < 1$ is the subjective discount factor and c_t^i is individual i's consumption of the single good in period t. The instantaneous utility function takes the constant relative risk aversion (CRRA) form given by,

$$u\left(c\right) = \frac{1}{a}\left(1 - e^{-a\ln(c)}\right)$$

Here $a > 0$ measures absolute risk aversion with respect to log consumption; relative risk aversion with respect to consumption equals $1 + a$. Individuals in rural locations face uncertainty at time 0 with respect to urban sector productivity in the next period.

Earnings The earnings of individual i endowed with abilities $\{1, s^i\}$ in their rural and urban locations respectively would be,

$$y_{r,t}^i = w_{r,t} = A_r$$

$$y_{u,t}^i = w_{u,t}s^i = A_u s^i$$

The budget constraint of individual i in rural location r is $c_t^i = y_{r,t}^i$, and $c_t^i = y_{u,t}^i$ in the urban location. Each individual i only faces a dynamic migration choice problem (no savings), whereby the individual chooses whether to continue working in their rural location r or switch to the urban location. This choice is captured by an indicator variable $o_{r,t}^i$ with $o_{r,t}^i = 1$ if the individual chooses in period t to work in the urban location next period. Then we can re-write the number of individuals migrating from rural location r to the urban location u from period $t = 0$ to $t = 1$,

$$M_{r,1} = \sum_{i \in H_{r,1}} 1 = \sum_{i} o^i_{r,1}$$

and their effective units of labor supplied in the urban location by $\sum_{i \in H_{r,1}} s^i = \sum_i o^i_{r,1} \cdot s^i$.

Information Structure Individuals in rural locations are uncertain about their earnings in the urban location, that is, they do not know the true value of their urban earnings potential $y^i_{u,t}$. While individuals know their ability in the urban location s^i, they do not observe the common component of urban earnings $w_{u,t}$ that is determined by the urban technology parameter A_u from the firm's problem. Since they are uncertain about their urban earnings potential, in order to make their rural-urban migration decision individuals must form beliefs about the unobserved component of their urban earnings A_u.

Individuals in rural location r at time 0 have three sources of information about A_u: (a) the initial beliefs they hold (prior beliefs); (b) the urban earnings outcomes of individuals from the same rural location r as them that have already moved to the urban location; and, (c) a noisy private signal about true urban productivity. There is no learning from aggregate outcomes.

Prior Beliefs Initial beliefs are identical for all individuals i and for all rural locations r. Individuals in every rural location $r \in \{1, 2, ...R\}$ start at $t = 0$ with prior beliefs about A_u given by $\log(A_u) \sim \mathcal{N}(\mu_0, \sigma_0^2)$.

Spatial Social Learning Individuals learn about A_u by observing in the current period $t = 0$ the current urban earnings of individuals from their information "neighborhood" (same rural location r) that have migrated to the urban location before them. Individuals do not observe the earnings of those that have previously migrated from other rural locations. Individuals cannot observe the urban ability s^i of anyone else from their own or other rural location. Specifically, an individual i in rural location r learns at time 0 from a number of previous migrants $M_{r,0}$ from their location, by observing their earnings $\left\{y^j_{u,0}\right\}_{o^j_{r,0}, j \neq i}$. All individuals in the same information neighborhood (rural location) r receive the same amount of information from previous migrants. The information revealed from the urban earnings of previous migrants is noisy because it confounds urban sector productivity and the idiosyncratic abilities of the previous migrants. In particular, while individuals in r can observe $y^j_{u,0}$ for an individual urban worker from their location $j \in H_{r,0}$ they cannot distinguish the common component A_u from the idiosyncratic term e^j (or s^j). Since e^j is not observed you cannot infer A_u from $y^j_{u,0}$. So each signal from an individual $j \in H_{r,0}$ is a noisy signal about A_u,

$$\log\left(y_{u,0}^{j}\right) = \log A_u + \mu_s + e^j$$

Because of normality, with $M_{r,0}$ signals it is equivalent to an individual in rural location r receiving a noisy location-specific aggregate signal,

$$\log \overline{y}_{r,0} \equiv \frac{1}{M_{r,0}} \sum_{j \in H_{r,0}} \log\left(y_{u,0}^{j}\right) = \log A_u + \mu_s + e_{r,0}$$

where the overall noise in r is captured by

$$e_{r,0} \equiv \frac{1}{M_{r,0}} \sum_{j \in H_{r,0}} e^j \sim \mathcal{N}\left(0, \frac{1}{M_{r,0}}\sigma_s^2\right)$$

Let $\sigma_{e_r}^2 \equiv \frac{\sigma_s^2}{M_{r,0}}$. Note that because the distribution of skills from all rural locations is identical, the dispersion of skills σ_s is the same, and $\sigma_{e_r}^2$ only differs across locations inversely to $M_{r,0}$. The signal of previous migrants' earnings differs across rural locations, but is common to all individuals within each rural location.

Private Signal To capture the fact that the amount of information accessed by individuals within a rural location r may differ across individuals, and that their migration decisions may respond to other factors besides the public (within r) signal of previous migrant earnings, we allow for each individual i to observe a private noisy signal about true urban productivity, $\eta_r^i \sim \mathcal{N}\left(\log A_u, \sigma_{\eta_r}^2\right)$ which is *i.i.d* across individuals in r. Note, that the variance of the noise can differ across locations, hence the r subscript in σ_{η_r}.

Bayesian Updating In each information neighborhood, the content of signals in period $t = 0$ depends on the overall number of previous migrants from that rural location. In addition, individuals receive a private signal each. Individuals use the information on observed urban earnings of their "neighbors" and the private signal to update their priors according to Bayes rule.

With Bayesian updating, individual i in rural location r forms an expectation about urban productivity, conditional on the information set available to them Ω_r^i, which has both idiosyncratic and common components. With Bayesian updating, individual i's posterior mean is a linear combination of the noisy public and private signals, $\log \overline{y}_{r,0}$ and η_r^i, and their prior beliefs μ_0,

$$E\left(\log A_u | \Omega_r^i\right) \equiv m_r^i = \left(1 - \theta_r - \psi_r\right)\mu_0 + \theta_r\left(\log \overline{y}_{r,0} - \mu_s\right) + \psi_r \eta_r^i$$

where each component is weighted by its relative precision,

$$\theta_r \equiv \frac{v_r}{\sigma^2_{e_r}}$$

$$\psi_r \equiv \frac{v_r}{\sigma^2_{\eta_r}}$$

$$1 - \theta_r - \psi_r \equiv \frac{v_r}{\sigma^2_0}$$

and v_r is the posterior variance in rural location r. The posterior precision (inverse of the variance) is the sum of the prior precision and the signal precisions. The posterior variance is given by,

$$v_r = \frac{1}{\frac{1}{\sigma^2_0} + \frac{1}{\sigma^2_{e_r}} + \frac{1}{\sigma^2_{\eta_r}}} = \frac{\sigma^2_0 \sigma^2_{e_r} \sigma^2_{\eta_r}}{\sigma^2_0 \sigma^2_{e_r} + \sigma^2_0 \sigma^2_{\eta_r} + \sigma^2_{\eta_r} \sigma^2_{e_r}}$$

Note that $\{\theta_r, \psi_r, v_r\}$ are common across all individuals in rural location r. However, they differ across rural locations given that the amount and precision of information differs across them.

Distribution of Posterior Beliefs in r Posterior beliefs of individuals in rural location r about the unobservable common component of urban earnings (urban TFP) $\log A_u$ conditional on their information set Ω^i_r are normally distributed,

$$\log A_u | \Omega^i_r \sim \mathcal{N}\left(m^i_r, v_r\right)$$

Cross-sectional distribution of posterior mean in r The cross-sectional mean of the posterior means across all individuals i within rural location r, is given by,

$$\overline{m}_r = \mathbb{E}_i\left[m^i_r \mid \log A_u, e_{r,0}\right] = \mu_0 + (\theta_r + \psi_r)(\log A_u - \mu_0) + \theta_r e_{r,0}.$$

If the prior mean equals realized true urban productivity, $\mu_0 = \log A_u$, this mean simplifies to

$$\overline{m}_r = \mu_0 + \theta_r e_{r,0}, \qquad \mathbb{E}[\overline{m}_r \mid \log A_u = \mu_0] = \mu_0.$$

The expectation averages over public-signal realizations; within a realized location, the common public-signal error remains in the mean.

The cross-sectional variance of the posterior means across all individuals i within r is,

$$\begin{aligned}\mathrm{Var}_i\big(m_r^i \mid \log A_u, e_{r,0}\big) &= \mathbb{E}_i\big[(m_r^i - \overline{m}_r)^2 \mid \log A_u, e_{r,0}\big] \\ &= \psi_r^2 \sigma_{\eta_r}^2.\end{aligned}$$

Therefore, the location-specific cross-sectional distribution of the posterior mean m_r^i across all individuals i in rural location r is normal and given by,

$$m_r^i \mid \log A_u, e_{r,0} \sim \mathcal{N}\left(\overline{m}_r, \psi_r^2 \sigma_{\eta_r}^2\right).$$

When $\mu_0 = \log A_u$, this distribution becomes

$$m_r^i \mid \log A_u = \mu_0, e_{r,0} \sim \mathcal{N}\left(\mu_0 + \theta_r e_{r,0}, \psi_r^2 \sigma_{\eta_r}^2\right).$$

Ex ante distribution of posterior means in r Before urban productivity and signals are realized, consider the prior-predictive distribution of individual posterior means. Here, $\log A_u \sim \mathcal{N}(\mu_0, \sigma_0^2)$, and the public- and private-signal errors are independent of urban productivity and of each other. The ex ante mean is,

$$E\left(m_r^i\right) = \mu_0$$

The ex ante variance is,

$$VAR\left(m_r^i\right) = E\left(m_r^i - \mu_0\right)^2 = \left(\theta_r + \psi_r\right)^2 \sigma_0^2 + \theta_r^2 \sigma_{e_r}^2 + \psi_r^2 \sigma_{\eta_r}^2 = \sigma_0^2 - v_r$$

Therefore, the ex ante distribution of the posterior mean is,

$$m_r^i \sim \mathcal{N}\left(\mu_0, \sigma_0^2 - v_r\right)$$

These moments average over possible realizations of urban productivity and both signal errors, rather than over individuals within one realized location. If we instead condition on realized true productivity equaling the prior mean, $\log A_u = \mu_0$, and average over both signal errors, the distribution becomes

$$m_r^i \mid \log A_u = \mu_0 \sim \mathcal{N}\left(\mu_0, \theta_r^2 \sigma_{e_r}^2 + \psi_r^2 \sigma_{\eta_r}^2\right).$$

Migration Costs Given that information frictions may not be the only barrier to labor mobility of workers from rural locations to the urban center, we allow for residual migration costs. We summarize all other costs to migration by a location-specific barrier, that takes the form of a tax or wedge in urban earnings, denoted ξ_r. This is a catchall for all factors that can inhibit migration to the urban location, even in the absence of information frictions. The budget constraint is $c_t^i = y_{r,t}^i$ for rural residents and $c_t^i = y_{u,t}^i$ for existing urban residents. For an individual who migrates from rural location r to the urban location between $t = 0$ and $t = 1$, the migration wedge enters period-1 consumption as $c_1^i = e^{-\xi_r} y_{u,1}^i = e^{-\xi_r} A_u s^i$.

Migration Problem of Rural Worker A worker in rural location r today ($t = 0$) that decides to stay in the same rural location tomorrow ($t = 1$) will have consumptions $c_0^i = A_r$ at time $t = 0$ and $c_1^i = A_r$ at time $t = 1$. The lifetime welfare of a "stayer" in location r is,

$$V_r^i = u\left(c_0^i\right) + \beta u\left(c_1^i\right) = \frac{1}{a}\left(1 - e^{-a\ln(A_r)}\right) + \frac{\beta}{a}\left(1 - e^{-a\log(A_r)}\right) = \frac{1+\beta}{a}\left(1 - e^{-a\log(A_r)}\right)$$

A worker in rural location r today ($t = 0$) that migrates tomorrow ($t = 1$) to the urban location will have consumption $c_0^i = A_r$ at time $t = 0$ and $c_1^i = e^{-\xi_r} A_u s^i$ at time $t = 1$, but A_u is uncertain. Thus the expected lifetime welfare of a switcher with information set Ω_r^i is given by,

$$E\left(V_u^i\left(s^i\right)|\Omega_r^i\right) = u\left(c_0^i\right) + \beta E\left(u\left(c_1^i\right)|\Omega_r^i\right) =$$

$$= \frac{1}{a}\left(1 - e^{-a\log(A_r)}\right) + \frac{\beta}{a}\left(1 - E\left[e^{a\xi_r} e^{-a\ln(A_u)} e^{-a\log(s^i)}|\Omega_r^i\right]\right) =$$

$$= \frac{1}{a}\left(1 - e^{-a\log(A_r)}\right) + \frac{\beta}{a}\left(1 - e^{a\xi_r} e^{-a\ln(s^i)} E\left[e^{-a\log(A_u)}|\Omega_r^i\right]\right) =$$

$$= \frac{1}{a}\left(1 - e^{-a\log(A_r)}\right) + \frac{\beta}{a}\left(1 - e^{-a\log(s^i)} e^{-am_r^i + \frac{a^2}{2}v_r + a\xi_r}\right)$$

where we have exploited the CRRA property of preferences (exponential) and the normality property of the revised beliefs.

The occupational choice problem of a rural worker is then to pick the occupation that maximizes their expected lifetime utility,

$$V^i = \max_{o^i_{r,1} \in \{0,1\}} \left\{ V^i_r, E\left(V^i_u\left(s^i\right)|\Omega^i_r\right)\right\}$$

The expected benefit of individual i in rural location r from migrating is,

$$EB^i_r = E\left(V^i_u\left(s^i\right)|\Omega^i_r\right) - V^i_r = \frac{\beta}{a}\left(1 - e^{-a\log(s^i)}e^{-am^i_r + \frac{a^2}{2}v_r + a\xi_r}\right) - \frac{\beta}{a}\left(1 - e^{-a\log(A_r)}\right)$$

Individual i will migrate if $EB^i_r > 0$, that is, if the perceived net benefit of migration is positive,

$$b^i_r \equiv \log s^i + m^i_r - \xi_r - \frac{a}{2}v_r - \log A_r > 0$$

to which the expected benefit boils down to.

Number of migrants from location r Given the connection of migration decision to the migration benefit for each individual i in rural location r, we can compute the total migration by location as,

$$\frac{M_{r,1}}{n_{r,1}} = Prob\left\{b^i_r > 0\right\} = 1 - Prob\left\{\log s^i < -m^i_r + \xi_r + \frac{a}{2}v_r + \log A_r\right\}$$

which is the share of workers for which the perceived return is positive.

3.1 Role of information frictions for migration

We characterize the effects of uncertainty and belief heterogeneity on individual migration decisions and the sorting patterns of workers across space. Defining the mean posterior belief in district r as $\overline{m}_r$ and the log-actual urban TFP as $\log(A_u)$ we can re-write the net perceived benefit to migration as,

$$b^i_r \equiv \underbrace{\log s^i + \log A_u - \xi_r - \log A_r}_{\text{actual net return}} - \underbrace{\frac{a}{2}v_r}_{\text{risk penalty}} + \underbrace{\left[\overline{m}_r - \log A_u\right]}_{\text{average belief bias}} + \underbrace{\left[m^i_r - \overline{m}_r\right]}_{\text{belief heterogeneity}} \tag{2}$$

Equation 2 can be used to unpack the effect of information frictions and the heterogeneity in beliefs on the cut-off level of ability above which individuals in r move to the city. The perceived net return can be decomposed into four terms. The first term is the *actual* net return, which would be the return to migration in the perfect information case. The second term with the posterior variance,

that enters negatively, can be thought of as the certainty equivalent risk penalty. It captures that risk averse workers dislike the risk associated with urban opportunities. The third term, captures the average belief bias about urban productivity, which is negative if workers systematically under-estimate urban wages. The last term captures belief heterogeneity across individuals within the same district. It is the difference of individual beliefs from the average, and if individuals are more "pessimistic" then it is also negative.

Perfect Information Under perfect information there is no bias in average beliefs ($\overline{m}_r = \log(A_u)$), no risk discounting of moving to the urban location ($v_r = 0$), and no heterogeneity in beliefs ($m_r^i = \overline{m}_r$). In this case equation 2 becomes,

$$\log s^i > -\log(A_u) + \xi_r + \log A_r \equiv \log \widehat{s}_r^{PI}$$

where the right-hand side is the cut-off level of ability under perfect information $\widehat{s}_r^{PI}$, above which individuals in r migrate to the urban location and below which they stay in the rural location.

Information Frictions without belief heterogeneity In the case of uncertainty with common beliefs across workers within rural locations ($m_r^i = \overline{m}_r$) equation 2 becomes,

$$\log s^i > -\log(A_u) + [\log(A_u) - \overline{m}_r] + \xi_r + \frac{a}{2} v_r + \log A_r \equiv \log \widehat{s}_r^{IF}$$

Uncertainty itself has two effects on the cut-off level of ability $\widehat{s}_r^{IF}$. First, the risk discounting of the urban utility raises the cut-off, and reduces migration. Intuitively, this is because of the interaction of belief uncertainty and risk aversion. Migration depends not just on the posterior mean but also the posterior variance. The term $\exp(-\frac{a}{2} v_r)$ can be thought of as the risk premium, which reduces the marginal perceived gain from migrating for each ability level. Second, if there is a downward bias in beliefs, i.e., individuals systematically under-estimate urban earnings (TFP), $\overline{m}_r < \log(A_u)$, there is an additional increase in the cut-off.

Information Frictions with belief heterogeneity In this case individuals receive an additional private signal η_r^i that affects their posterior mean m_r^i and thus their migration decision. If $m_r^i > \overline{m}_r$, i.e., individual is optimistic ($\eta_r^i > \log A_u$) they are more likely to migrate for given s^i. If $m_r^i < \overline{m}_r$, i.e., individual is pessimistic ($\eta_r^i < \log A_u$) $\rightarrow$ they are less likely to migrate for given s^i. Heterogeneous beliefs make the migration cutoff "fuzzier." In other words our framework generates imperfect sorting endogenously, through the learning environment: some low-ability but optimistic

individuals can migrate and some high-ability but pessimistic individuals can stay behind in the rural location. This happens because migration is based on *perceived* returns rather than actual.

Figures 4, 5 and 6 illustrate how uncertainty and heterogeneous beliefs shape rural-urban migration decisions. In Figure 4 there is perfect information, which means that the risk penalty, the average belief bias and the belief heterogeneity terms are all zero. In the absence of uncertainty migration decisions are characterized by a cutoff level of ability $\widehat{s}^{PI}$, with workers that draw an ability s_i above this migrating and those below staying in their rural location. The dotted vertical line marks the skill at which the urban payoff (red) equals the rural wage (dashed horizontal line).

Figure 4: Migration Decision: Perfect Information

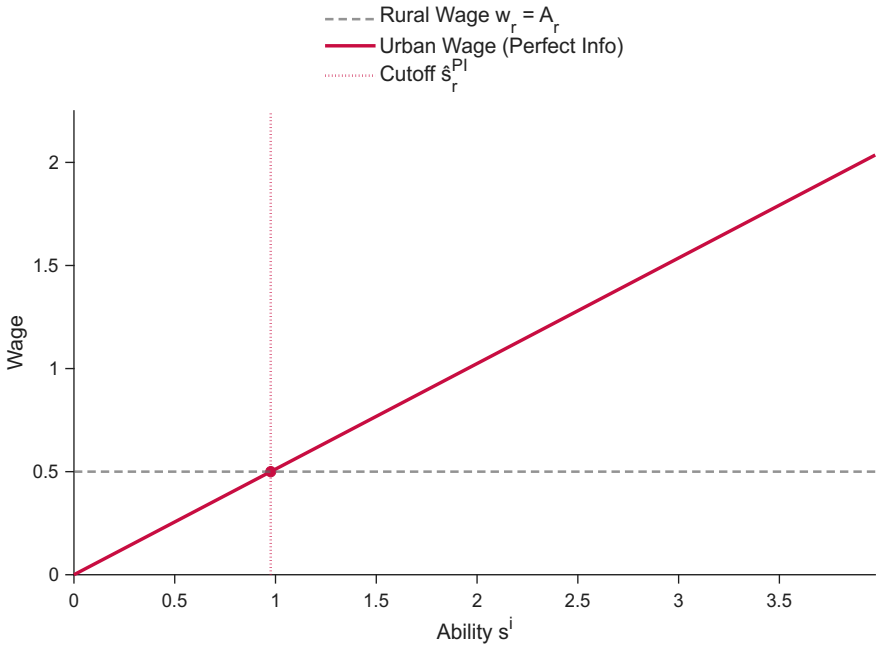
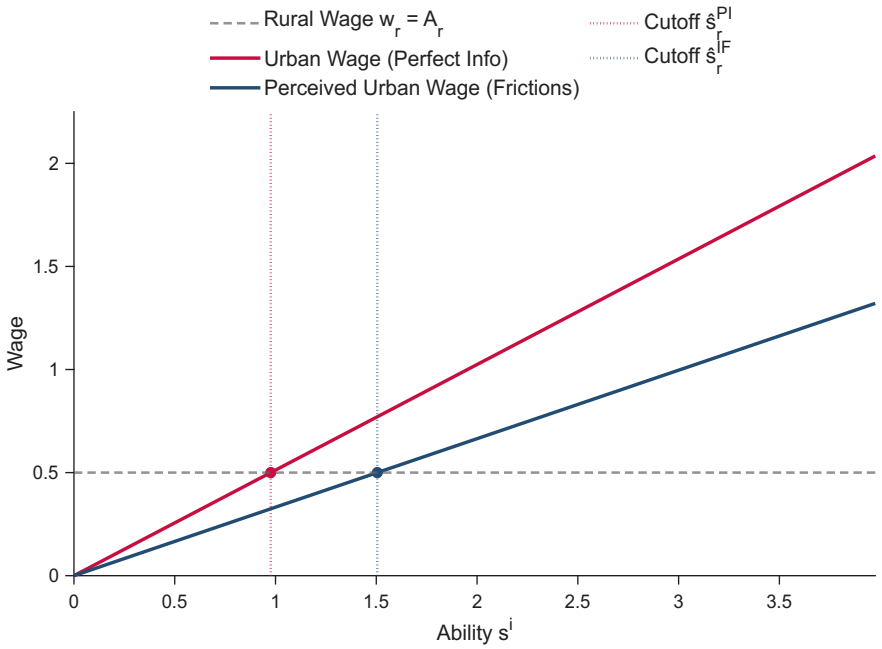


Figure 5 shows how the introduction of uncertainty, under common beliefs across all individuals in a district, can affect the incentive to migrate. The cutoff level of ability above which individuals migrate is now higher $\widehat{s}^{IF} > \widehat{s}^{PI}$, because the risk penalty reduces the perceived return to migration for all s_i, making the urban return line flatter. If average beliefs are also downward biased then the perceived urban return line is even flatter. The two dotted lines bracket the workers who would migrate under perfect information but stay under information frictions. These imply that information frictions inhibit migration, with only the highest ability workers in a rural location moving when all share the same beliefs.

Figure 6 shows that when beliefs are heterogeneous all workers with the same ability s_i do

Figure 5: Migration Decision: Uncertainty, Common Beliefs



not make the same migration choices. Each cross shows a worker's perceived urban payoff, and the blue line gives the payoff at the district's average belief. At the highlighted skill level, only the more optimistic worker migrates because only that worker's perceived payoff exceeds the rural wage. With belief heterogeneity the private signal affects individual posterior means: if a worker receives a signal above its within-district mean, $\eta_r^i > \log A_u$, then $m_r^i > \overline{m}_r$, which means they are more optimistic than average and thus more likely to migrate. This implies that mapping ability to migration is not linear but noisy: some low ability workers that are optimistic may migrate; some high ability workers that are pessimistic may stay. In other words, belief heterogeneity gives rise to imperfect sorting even when actual ability is uni-dimensional.

In our model, these differences in perceived payoffs arise from the signals workers observe. Previous migrants' earnings provide a noisy public signal shared within the district, while private signals differ across individuals. The content and precision of these signals determine how workers update their beliefs and how much uncertainty remains. Common information can shift the district's perceived urban payoff schedule, while private information generates dispersion around it. Signals therefore affect both how many workers migrate and which workers leave.

Figure 6: Migration Decision: Uncertainty, Heterogeneous Beliefs



4 Quantitative Analysis

Our strategy is to estimate the model in the presence of information frictions. We use indirect inference by mapping moments from the panel micro-level data for Indonesia to model moments, in order to identify the extent of information frictions, residual migration barriers, and the other model parameters. We then shut down information frictions and examine the model implications for aggregate migration and the spatial patterns of migration.

We calibrate the model using district-level migration and migrant earnings data for 2000 and 2007, which correspond to $t = 0$ and $t = 1$. The 2000 migrants are the previous migrants $M_{r,0}$, whose urban earnings form the public signal, and the 2007 migrants are $M_{r,1}$. For the 2000 wave, we assume that previous migrants made the same migration decision, facing a barrier $\xi_{r,0}$, but without a public signal, so that their posterior variance is $v_{r,0} = (\tau_0 + \tau_{\eta,r})^{-1}$, where $\tau_0 = \sigma_0^{-2}$ and $\tau_{\eta,r} = \sigma_{\eta_r}^{-2}$ denote prior and private-signal precision. We denote by $\sigma^2_{b,r,t}$ the variance of the migration index $b^i_{r,t}$ across individuals in district r.

The procedure has two distinct stages. In the *identification stage*, we fit district-specific barriers to 2000 migration rates and hold those barriers fixed when predicting migration in 2007. This restriction allows the model to attribute the predicted relationship between previous migrant earn-

ings and subsequent migration to its information channel, conditional on other district observables. In the *baseline-fitting stage*, after identifying the information parameters, we allow district-specific residual changes in barriers between 2000 and 2007 so that the model matches observed 2007 migration rates. The fitted 2007 barriers are then held fixed in information counterfactuals.

4.1 Calibration

We normalize true urban TFP A_u to 1, and back out rural TFP A_r in each rural district r to match average income gap between that district and the city.

We set the preference parameters based on a priori information. We set $a = 1$, corresponding to relative risk aversion of two with respect to consumption. The subjective discount factor β is set to 0.98.

We take as given the initial spatial distribution of permanent migration. The number of previous migrants for each district r as a share of the total urban population, $M_{r,0}/N_{u,0}$ is taken directly from the data. The micro data allow us to identify the rural location from which urban residents have permanently migrated from. We use the observed number of previous migrants from each district, $M_{r,0}$, to construct the public signal. For each matched district, we also observe the relevant origin population $n_{r,t}$, and hence the migration rate $p_{r,t} = M_{r,t}/n_{r,t}$, the mean log earnings of 2000 migrants $\log \overline{y}_{r,0}$, and measures of the within-district variance of migrant wage gaps. We use the matched sample of 91 districts after data-quality filters.

Urban skills follow a log-normal distribution, $s^i \sim L\mathcal{N}(\mu_s, \sigma_s^2)$. We choose the mean μ_s and variance σ_s^2 of the skill distribution to match the mean and variance of log urban earnings among 2000 migrants, $\mu_{s,\text{selected}}$ and $\sigma^2_{s,\text{selected}}$, corrected for selection as described below. The urban skills of the previous migrants as well as those of the current rural residents are drawn from the same time invariant distribution.

Correction for migrant selection Because migrants are positively selected on the migration index, the distribution of their earnings does not generally equal the underlying skill distribution. We therefore correct the observed mean and variance of selected migrant skills before estimating information parameters. Let $q_{r,0} = \Phi^{-1}(1 - p_{r,0})$ denote the standard-normal threshold implied by the observed 2000 migration rate. Define the inverse Mills ratio and the associated truncation adjustment:

$$\lambda_r = \frac{\phi(q_{r,0})}{p_{r,0}}, \qquad \kappa_r = \lambda_r(\lambda_r - q_{r,0}).$$

Let G_r be the observed variance of the migrant wage gap among workers from district r who migrated in period 0. We compute

$$S_r = \frac{G_r}{1-\kappa_r}, \qquad \sigma^2_{b,r,0} = \max\left\{S_r, (1+f)\,\sigma_s^2\right\}, \quad f = 0.01,$$

where S_r inverts the truncated variance of the migration index (the second moment in Section 4.2). The small floor guarantees positive dispersion attributable to the private signal. In the reported baseline no district is affected by this floor. Conditional on a candidate unselected skill variance σ_s^2, the model-implied selection premium and selected variance are

$$E_r^m = \frac{\sigma_s^2 \lambda_r}{\sigma_{b,r,0}}, \qquad V_r^m = \sigma_s^2 - \frac{\sigma_s^4 \kappa_r}{\sigma^2_{b,r,0}}.$$

With weights $\omega_r = M_{r,0}/\sum_j M_{j,0}$, the unselected variance is recovered by matching the pooled variance of selected migrant earnings:

$$\sigma^2_{s,\text{selected}} = \sum_r \omega_r V_r^m + \sum_r \omega_r \left(E_r^m\right)^2 - \left(\sum_r \omega_r E_r^m\right)^2.$$

The unselected mean is then $\mu_s = \mu_{s,\text{selected}} - \sum_r \omega_r E_r^m$.

The observed district migrant earnings signal is also adjusted for the expected skill premium among selected migrants:

$$z_r = \log \overline{y}_{r,0} - \mu_s - E_r^m.$$

The signal therefore attempts to remove the component of observed migrant earnings due to migrant skill and selection, leaving information about urban productivity, $z_r = \log A_u + e_{r,0}$. The precision of the public signal is disciplined by the number of previous migrants:

$$\sigma^2_{e_r} = \frac{V_r^m}{M_{r,0}}.$$

This replaces σ_s^2 in the definition of $\sigma^2_{e_r}$ in Section 3 with the variance of selected migrant earnings, because previous migrants are selected on the migration index. Districts with more previous migrants have more precise public signals, other things equal.

4.2 Identifying Information Frictions

The extent of information frictions in the model is captured by the economy-wide uncertainty

about urban TFP, σ_0^2, and the district-specific uncertainties $\sigma_{\eta_r}^2$ for each district r. We center the common prior at the realized value, $\mu_0 = \log A_u$. These along with residual district-specific migration frictions are jointly estimated.

The main challenge in quantifying the role of information frictions is that we do not directly observe the information set that agents face. Our approach here is to use easily computed empirical moments and indirect inference, through the lens of the model, to back out the extent of uncertainty rural individuals face about urban earnings opportunities.

To identify the intensity of information frictions we use two key sets of moments from the data:

1. **Correlation of previous migrant income and migration flows.** The first moment is the correlation across space, $Corr_r\left(\log \overline{y}_{r,0}, M_{r,1}\right)$, of subsequent migration flows and average urban earnings of previous migrants from the same districts. This empirical moment is informative about the model uncertainty parameter σ_0. Intuitively, if rural workers are facing a lot of uncertainty over urban sector productivity A_u (high σ_0^2), then we would expect migration flows to be very responsive to information about the urban earnings of their peers. For a given level of noise the extent to which rural workers migration decisions comove with urban earnings is a function of the overall level of information. Lower quality information would imply more comovement of migration flows and peer urban earnings. In the data we find that this correlation is positive and equal to 0.242.

2. **Dispersion of rural-urban wage gaps within districts, for each r.** The previous migrant earnings signal in 1. is the same for all rural individuals of the same district. However, we observe considerable heterogeneity in migration decisions even within rural districts. To discipline the extent of within district uncertainty, as captured by the variance of the private noisy signal (σ_{η_r} in district r), the second moment we use for each district is the dispersion of rural-urban wage gaps across individuals that have migrated from that district. Intuitively, the larger the variance of the rural-urban income gaps in a district the larger the uncertainty within the district about the urban return to employment. In particular, we infer σ_{η_r} in each district such that the conditional variance in the model of rural-urban wage gaps – conditional on migrating, matches the corresponding observed empirical moment on wage gaps of previous district-specific migrants in the data. To do this we compute the conditional variance in the model of the left-truncated distribution, exploiting log-normality, in each wave t,

$$VAR\left(b_{r,t}^i | b_{r,t}^i > 0\right) = \sigma_{b,r,t}^2 \left\{1 - \lambda\left(q_{r,t}\right)\left[\lambda\left(q_{r,t}\right) - q_{r,t}\right]\right\}$$

where $b^i_{r,t}$ is the rural-urban income gap or benefit to migrating of individual i in district r, and $\sigma^2_{b,r,t}$ is the unconditional variance of the income gap across individuals,

$$\sigma^2_{b,r,t} = \sigma^2_s + \tau_{\eta,r} v^2_{r,t}$$

with $v_{r,t}$ also being a non-linear function of σ_{η_r}. Within a district, the prior mean and the public signal are common to all individuals, so dispersion comes only from skills and the private signal, which enters $m^i_{r,t}$ with weight $\tau_{\eta,r} v_{r,t}$ (ψ_r in Section 3). The expression therefore holds in both waves; only $v_{r,t}$ differs. Note that $\lambda(q_{r,t}) \equiv E[\zeta | \zeta > q_{r,t}]$ represents the upper tail truncation of a standard normal random variable ζ, where $q_{r,t} = \Phi^{-1}(1 - p_{r,t})$ is the truncation point implied by the migration rate $p_{r,t}$.

We implement the second moment with the wage gaps of the 2000 migrants. For $t = 0$ the truncation term equals $1 - \kappa_r$, so setting the conditional variance equal to the observed G_r gives $\sigma^2_{b,r,0}$ in the selection correction above. Define $X_r = \sigma^2_{b,r,0} - \sigma^2_s$. For any candidate prior precision τ_0, the private-signal precision solves $X_r = \tau_{\eta,r} / (\tau_0 + \tau_{\eta,r})^2$. We take the upper root,

$$\tau_{\eta,r} = \frac{1 - 2\tau_0 X_r + \sqrt{1 - 4\tau_0 X_r}}{2X_r},$$

which requires $1 - 4\tau_0 X_r \geq 0$ for every district. Consequently, feasible values of the prior standard deviation satisfy $\sigma_0 \geq 2\sqrt{\max_r X_r}$. This lower bound restricts the range of σ_0 over which we search, and the choice of the upper root is a maintained assumption. Conditional on σ_0, these equations recover the district-specific private-signal variances $\sigma^2_{\eta_r} = 1/\tau_{\eta,r}$.

Fitting 2000 migration and identifying σ_0 For a candidate σ_0, the posterior mean averaged over private-signal noise at the realized urban productivity is $\overline{m}_{r,0} = v_{r,0}(\tau_0 \mu_0 + \tau_{\eta,r} \log A_u)$. Integrating the normal migration index yields the 2000 migration probability

$$p^{\text{mod}}_{r,0} = \Phi\left(\frac{\mu_s + \overline{m}_{r,0} - \frac{a}{2} v_{r,0} - \log A_r - \xi_{r,0}}{\sigma_{b,r,0}} \right).$$

We choose one barrier for each district to fit its observed migration rate exactly:

$$\xi_{r,0}(\sigma_0) = \mu_s + \overline{m}_{r,0} - \frac{a}{2} v_{r,0} - \log A_r - \sigma_{b,r,0} \Phi^{-1}(p_{r,0}).$$

These barriers are residual wedges: they summarize migration determinants not otherwise accounted for by the model, and they need not be positive. They are not separately identified measures of literal monetary moving costs.

Predicting 2007 migration without changing the 2000 barriers For the same candidate σ_0, we introduce the public signal z_r. Averaging $m^i_{r,1}$ over private-signal noise gives

$$\overline{m}_{r,1} = (1 - \theta_r - \psi_r)\,\mu_0 + \theta_r z_r + \psi_r \log A_u, \qquad \sigma^2_{b,r,1} = \sigma^2_s + \tau_{\eta,r} v^2_{r,1}.$$

Thus, changes in public information affect migration both through the posterior mean and through the reduction in posterior uncertainty. The magnitude of these responses depends on the relative precision of the prior and signals, not only on the observed earnings level. The identifying-stage prediction uses the *unchanged* 2000 barriers:

$$\widetilde{p}_{r,1}(\sigma_0) = \Phi\left(\frac{\mu_s + \overline{m}_{r,1} - \frac{a}{2} v_{r,1} - \log A_r - \xi_{r,0}(\sigma_0)}{\sigma_{b,r,1}}\right), \qquad \widetilde{M}_{r,1}(\sigma_0) = n_{r,1}\widetilde{p}_{r,1}(\sigma_0). \tag{3}$$

We use the 2007 population to convert predicted migration rates into migrant *counts*. Thus, the identifying statistic is based on count levels, not rates or count changes. The empirical target is the unweighted Pearson correlation across districts between 2000 migrant earnings and observed 2007 migrant counts, $\widehat{\rho}_{\text{data}} = Corr_r\left(\log \overline{y}_{r,0}, M_{r,1}\right)$, and its model counterpart is $\rho_{\text{mod}}(\sigma_0) = Corr_r\left(\log \overline{y}_{r,0}, \widetilde{M}_{r,1}(\sigma_0)\right)$. We identify σ_0 by solving the scalar equation $\rho_{\text{mod}}(\widehat{\sigma}_0) = \widehat{\rho}_{\text{data}}$. We solve this equation numerically over the feasible range of σ_0, verifying that the model correlation crosses the empirical target exactly once.[8]

Suppose two districts have similar initial migration propensities after accounting for the fitted 2000 barriers, but previous migrants in one district report higher urban earnings. Its public signal will generally raise the expected payoff to migration relative to the other district. The prior variance σ^2_0 determines the prior's precision relative to private and public information. The model therefore maps a candidate σ_0 into a predicted cross-district relationship between previous earnings and subsequent migrant counts.

Two assumptions matter for this identification. First, we hold the district-specific barriers fitted in 2000 fixed when predicting 2007 migration in equation (3). The correlation therefore identifies σ_0 under the assumption that changes in other migration barriers between 2000 and 2007 are unrelated to previous migrant earnings. Second, the target is based on counts, so variation

[8] We evaluate the moment on a dense grid and refine the solution with a root-finding algorithm.

in district population $n_{r,1}$ also contributes to the model and data correlations, and the moment does not net it out. Specifications that target migration rates or their changes were not feasible, while a target based on the change in migrant counts gives similar welfare results. The estimate also depends on the 2000 selection correction, on the precision assigned to the public signal, on the choice of root for the private-signal precision, and on centering the prior on true urban productivity.

Residual district-level migration frictions We note that while the above estimated information-related frictions have implications for migration they do not account for the entirety of migration flows across districts. We allow the district-specific migration friction ξ_r to account for all remaining migration costs, such that the model exactly matches the migration shares in the data by district. Note that this approach implies that the model also matches the aggregate migration share by construction.

After identifying $\widehat{\sigma}_0$ and the associated private-signal precisions, we allow for additional district-specific barriers in 2007. With public information already incorporated into $\overline{m}_{r,1}$ and $v_{r,1}$, the total barrier that matches the observed 2007 migration rate is

$$\xi_{r,1} = \mu_s + \overline{m}_{r,1} - \frac{a}{2} v_{r,1} - \log A_r - \sigma_{b,r,1} \Phi^{-1}\left(p_{r,1}\right).$$

The residual change is $\Delta\xi_{r,1} = \xi_{r,1} - \xi_{r,0}$. These residuals absorb migration determinants that changed between 2000 and 2007 but are not modeled explicitly. Crucially, these residual barriers are recovered *after* σ_0 has been identified; they do not enter the first-stage correlation used to identify σ_0.

Our estimated model exactly matches the data along several dimensions: the aggregate migration rate, average rural-urban wage gaps and their dispersion by district, the correlation of previous migrant incomes and subsequent migration, as well as the initial and subsequent migration rates. The estimated prior standard deviation is $\widehat{\sigma}_0 = 6.28$ (variance 39.43), with the empirical and model identifying correlations matching to numerical precision. Table 4 reports the estimated parameters. Figure 7 shows the persistence of district-level migration flows in the estimated model. Because the barriers are fitted to observed migration in both years, the model reproduces this persistence by construction; the perfect-information experiment below measures how much of it is due to information frictions. The sample in Figure 7 consists of the 91 districts with all moments required for calibration available, compared with the 97 districts in Figure 2.

Table 4: Estimation of information and learning parameters - Indirect Inference

Parameter	Description	Target	Value
σ_0^2	Variance of prior beliefs	Correlation of 2000 mean log migrant earnings and 2007 migrant counts	39.43
μ_0	Mean of prior beliefs	Prior centered on true log urban productivity	0
σ_s^2	Variance of urban skills	Variance of log migrant earnings, adjusted for selection	0.94
μ_s	Mean of urban skills (log rupiah)	Mean log migrant earnings, adjusted for selection	14.23
$\sigma_{e_r}^2 = \frac{V_r^m}{M_{r,0}}$	Noise variance of the public signal	Selected earnings variance divided by the number of previous migrants	0.011–0.260
$\sigma_{\eta_r}^2$	Private-signal noise variance	Within-district dispersion of previous migrants' rural-urban wage gaps	0.19–22.19

Source: Spatial parameters estimated from district-level data in Indonesia (IFLS). For parameters that vary across districts, the range of lowest and highest values are provided.

External validation To assess the role of the public signal in 2007, we remove that signal while retaining the fitted 2007 barrier $\xi_{r,1}$. The posterior reverts to the prior-plus-private-information posterior, yielding

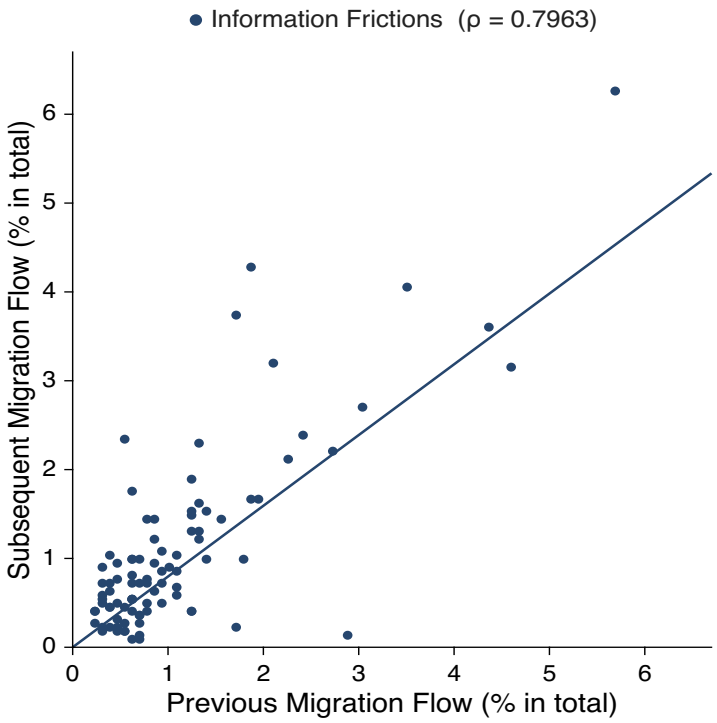
$$p_{r,1}^{\text{no public}} = \Phi\left(\frac{\mu_s + \overline{m}_{r,0} - \frac{a}{2}v_{r,0} - \log A_r - \xi_{r,1}}{\sigma_{b,r,0}}\right).$$

The average within-district log migration response to public information is

$$\beta_{\text{public}}^{\text{mod}} = \frac{1}{R}\sum_{r=1}^{R}\left[\log p_{r,1} - \log p_{r,1}^{\text{no public}}\right].$$

Because the same district population is used in both scenarios, the within-district log difference in counts is identical to the log difference in rates. This statistic is compared with the empirical internet-connectivity difference-in-differences coefficient of Section 2.2, but is *not targeted* in the estimation. The average log migration response to introducing the public signal is 0.247, compared with the internet DiD estimate of 0.403 in column (2) of Table 3. The comparison is imperfect:

Figure 7: Persistence in Migration Flows (estimated model)



Notes: The x-axis represents the initial period migration flow from a district, while the y-axis the subsequent period migration flow. Migration flows by district are measured relative to the total number of migrants in each period.

removing the entire public signal is not equivalent to adding one internet-connected village. The exercise is therefore a qualitative or magnitude-based external validation, not a replication of the treatment effect.

4.3 Quantitative Experiments

We have estimated the model in the presence of information frictions and residual migration costs. To assess the role of information frictions per se and their contribution relative to other mobility barriers we conduct three counterfactual experiments. First, we eliminate information frictions alone, keeping all other migration costs to their estimated values. Second, we eliminate all migration costs, information-induced and residual. Third, we separate the effect of perfect information into correcting belief errors and removing the risk penalty. The results of the first two counterfactuals are presented in Table 5, where the first column displays the statistics from the

Table 5: Effects of Information Frictions

	Baseline	Perfect Information	+ No Migration Costs
Aggregate Migration Rate (%)	10.5	14.7	32.3
Migration Persistence	0.80	0.26	−0.01
Income-Migration Correlation	0.24	−0.44	−0.19
Rural Period-2 Welfare Gain (%)	–	+2.47	+17.21
Δ Rural Output per Worker (%)	–	+4.1	+17.2
Δ Real GDP per Capita (%)	–	+1.4	+4.5

Note: Statistics for the baseline estimated model versus our two counterfactual experiments, "Perfect Information" ($v_r = 0$ for all r) and "+ No Migration Costs" ($\xi_r = 0$ for all r). The migration rate, migration persistence, and income–migration correlation are reported in levels. All remaining statistics are percentage changes relative to the baseline economy.

benchmark estimated economy, and those of the third in Table 6.

4.4 Perfect Information

In our theory, information frictions inhibit rural-to-urban migration and induce persistence in district-level migration flows over time. To assess the overall importance of information frictions we conduct a counterfactual of "perfect information," whereby we eliminate overall uncertainty (by setting $\sigma_0^2 = 0$) and district-specific uncertainty (by setting $\sigma_{\eta_r} = 0$ for each r). This is equivalent to $v_r = 0$ for all r. Under perfect information, rural workers are no longer uncertain about the urban earnings opportunities. They know that urban TFP is equal to its true value A_u, and given their urban skills they can compute exactly their gross return to urban employment, $\log(A_u) + \log(s^i)$. However, all other parameters are the same (including actual income gaps) and residual migration costs are still in place. Holding the 2007 barriers fixed, migration becomes

$$p_{r,1}^{\text{PI}} = \Phi\left(\frac{\mu_s + \log A_u - \log A_r - \xi_{r,1}}{\sigma_s}\right).$$

The aggregate baseline and perfect-information migration rates are computed as population-weighted district rates, $\sum_r n_{r,1} p_{r,1} / \sum_r n_{r,1}$ and $\sum_r n_{r,1} p_{r,1}^{\text{PI}} / \sum_r n_{r,1}$, respectively. The difference isolates the response to perfect knowledge of urban productivity within the fitted 2007 environment, conditional on the structural assumptions and fixed barriers.

We evaluate the counterfactual using the period utility $u(c)$. We treat the fitted migration wedge as a reduction in consumption after migration and integrate over the unselected skill distribution

and private signals. Because consumption at $t = 0$ is A_r in both scenarios and urban incumbents are unaffected, we report the gain for rural individuals at $t = 1$. For the baseline and perfect-information scenarios, let K^{II} and K^{PI} denote the expected values of $(c_1^i)^{-a}$ across rural individuals, weighted by district population $n_{r,1}$. The reported consumption-equivalent gain is

$$g = \left(\frac{K^{\text{II}}}{K^{\text{PI}}}\right)^{1/a} - 1.$$

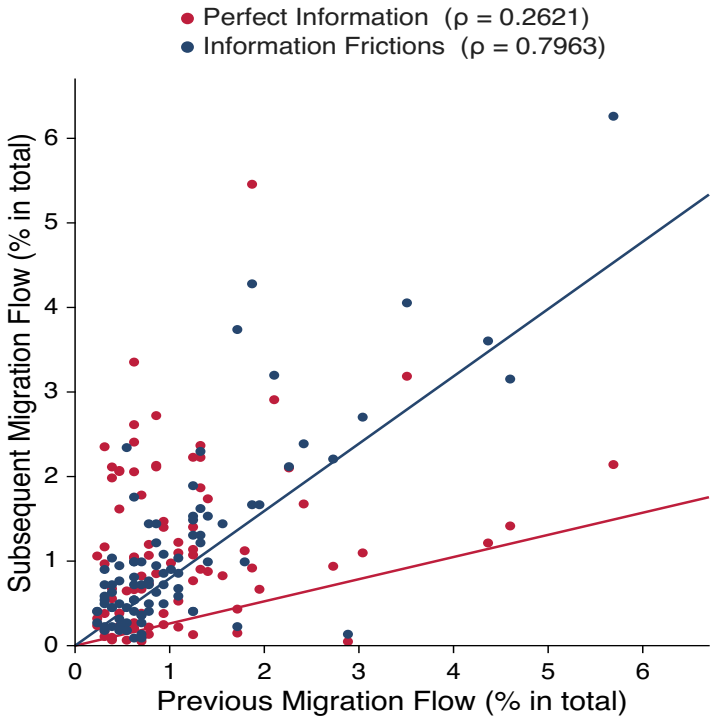
This is a conditional counterfactual evaluated at the calibrated realization of urban productivity, not a fully ex ante welfare calculation integrating over alternative realizations of aggregate urban productivity. It measures the gain for those who face the migration decision, not total welfare. Because the residual wedges may be negative and are treated as consumption wedges, their welfare interpretation warrants caution.

We display the results of this counterfactual in the second column of Table 5. Removing information frictions leads to a substantial increase in the economy-wide migration rate from 10.5 percent to 14.7 percent. More importantly, information frictions drive the substantial persistence in migration. With perfect information, the correlation of district-level migration flows falls from 0.80 to 0.26. See Figure 8. The associated consumption-equivalent welfare gain for rural individuals at $t = 1$ is 2.47 percent. GDP per capita also sees an increase of 1.4 percent, as workers shift to the higher productivity urban sector.

4.5 Eliminating All Migration Costs

The second counterfactual is presented in the third column of Table 5. In this quantitative experiment, in addition to removing information frictions (perfect information case), we also eliminate all other residual migration costs. Practically this means that in addition to $v_r = 0$, we also set $\xi_r = 0$ for all districts r. Removing all labor mobility barriers raises the migration rate to 32.3 percent. Not all rural workers move, because those with low urban skills s^i earn more in their rural district than in the urban sector. Comparing the second and third columns shows that eliminating information frictions alone accounts for 19.35% of the increase in migration obtained when all migration frictions are eliminated. While the welfare and productivity gains from removing all migration costs are considerably higher, a comparison of the second and third column suggests that information frictions account for a sizable portion of overall migration costs, of about 1/5. Importantly, the persistence of district-level migration flows over time is tightly linked to the poor

Figure 8: Persistence in Migration Flows - The Role of Information



Notes: The x-axis represents the initial period migration flow from a district, while the y-axis the subsequent period migration flow. Migration flows by district are measured relative to the total number of migrants in each period. The "Information Frictions" case is the benchmark estimated model. The "Perfect Information" corresponds to the counterfactual in which all information frictions are eliminated, keeping residual migration costs.

quality of information about urban earnings opportunities.

4.6 Belief Errors and the Risk Penalty

We separate the effect of perfect information into correcting signal errors and removing the risk penalty $av_{r,1}/2$. This penalty captures workers' reluctance to migrate when urban earnings are uncertain. Removing it makes migration more attractive even if average perceived earnings remain unchanged.

Using the IFLS calibration, we first set public and private signals equal to true log urban productivity, keeping baseline posterior weights and the risk penalty fixed. Separately, we remove the penalty while keeping beliefs unchanged. Applying both changes reproduces perfect information. Migration barriers and all other parameters remain fixed. These are accounting adjustments to the decision rule; an actual improvement in signal precision would change beliefs and uncertainty

together.

Table 6 shows that, accounting for interactions, signal corrections explain 85.87% of the migration increase and 97.76% of the rural welfare gain. Removing the risk penalty explains the remainder.

Table 6: Belief Errors and the Risk Penalty

Panel A: Diagnostic experiments

Scenario	Migration rate (%)	Persistence	GDP per capita change (%)	Rural period-2 welfare gain (%)
Baseline	10.53	0.7963	–	–
Correct signal errors only	14.06	0.2851	+1.29	+2.45
Remove risk penalty only	11.05	0.7833	+0.15	+0.09
Perfect information	14.74	0.2621	+1.42	+2.47

Panel B: Shares of the baseline-to-perfect-information change (%)

Outcome	Signal corrections	Removing the uncertainty penalty
Migration increase	85.87	14.13
Persistence reduction	96.63	3.37
GDP per capita gain	89.88	10.12
Rural period-2 welfare gain	97.76	2.24

Notes: GDP and welfare changes are relative to the baseline; welfare is the rural-origin period-2 consumption-equivalent gain. Panel B averages the contributions of public-signal correction, private-signal correction, and penalty removal across all six orders. Baseline posterior weights and migration barriers remain fixed.

5 Validation of the Information Mechanism

We now turn to the evidence from Kenya to assess the role of information in our theory. The Kenyan data contain what the Indonesian data lack: direct measures of rural workers' beliefs about urban earnings. We use them in two ways. First, we use the survey and the randomized information treatment of Baseler (2023) to examine what rural workers know about urban earnings, what they learn from previous migrants, and how they respond when they receive new information. Second, we estimate the model on Kenyan villages, once with the indirect inference strategy we use for Indonesia and once with reported beliefs, and compare their implications.

The key mechanism of the model in Section 3 has four links: uncertainty, local learning from previous migrants, belief revision and new migration. Rural workers are uncertain about urban

earnings because they do not observe urban productivity A_u. They learn about A_u from previous migrants of their own location, in addition to their prior and a private signal. They revise their beliefs as this information arrives, so workers in locations with more previous migrants hold beliefs closer to true urban earnings. These beliefs determine who migrates, and the new migrants become a source of information for later workers from the same location, which makes migration flows persistent. We examine these links using survey beliefs and the information experiment, and compare the model estimated indirectly with the model disciplined by reported beliefs.

Uncertainty about urban earnings and heterogeneity in beliefs. In the survey data in Kenya individuals residing in villages have imprecise beliefs, systematically under-estimating earnings in cities, on average by 40%. The mis-perception about urban opportunities signifies the presence and severity of uncertainty. While rural residents under-estimate urban earnings on average, there is considerable dispersion in those beliefs both across villages as well as within villages.

To estimate the dispersion of beliefs across and within villages, we regress beliefs of individual i, in village v, about earnings in city c on village fixed effects,

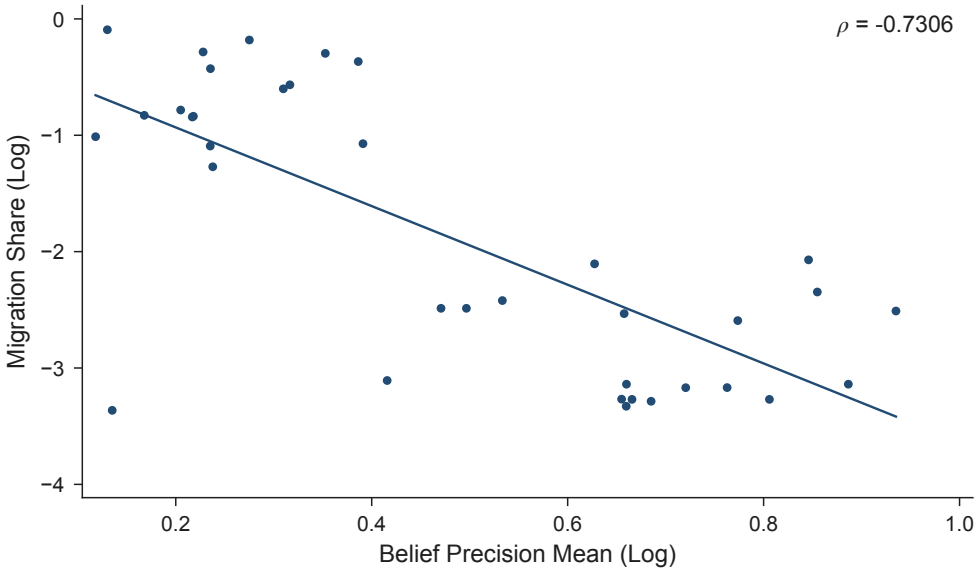
$$B_{ivc} = a_v + \varepsilon_{ivc}$$

To obtain the across village dispersion, we extract the estimated village fixed effects $\widehat{a}_v$ and compute the standard deviation. To obtain the pooled within village dispersion of beliefs we compute the standard deviation over the residuals $\widehat{\varepsilon}_{ivc}$. We estimate an across village dispersion of beliefs of 0.11, and within village of 0.83. This motivates the spatial dispersion in beliefs in the model, and the heterogeneity in beliefs even within rural districts.

Local learning from previous migrants. In the pooled sample of beliefs about Nairobi, Kisumu and Eldoret, we calculate the correlation between the log migration share and the mean log ratio of actual to perceived urban wages. Figure 9 shows that the correlation is -0.7306. Table 7 shows that this pattern also holds for individual respondents. We regress the log ratio on the log migration share and, separately, on an indicator for having a migrant family member. Having a migrant family member is associated with 0.228 log points less underestimation (column 2), so perceived earnings relative to actual earnings are 25.6% higher. For a respondent who perceives urban earnings 40% below actual earnings, this lowers the underestimation to about 25%, an error about 38% smaller. A 10% higher previous migration share is associated with 1.7% higher perceived earnings relative to actual earnings (column 1). Because these estimates include village fixed effects, they are not driven by differences between villages, and the controls account for differences in the education and

earnings of household heads. These patterns are consistent with more accurate beliefs where rural workers have more contact with previous migrants. In the model, more previous migrants make the public signal more precise, which moves beliefs toward true urban earnings.

Figure 9: Previous migration and underestimation of urban earnings in Kenya



Notes: The figure pools beliefs about Nairobi, Kisumu and Eldoret in the data of Baseler (2023). The horizontal axis is the mean log ratio of actual to perceived urban earnings, so larger values mean that respondents underestimate urban earnings more. The vertical axis is the log migration share. The line is the linear fit.

Belief update from new information. In the RCT individuals that are treated with new information about urban wages, revise their beliefs by 20% relative to the control group that receive no new information.

Migration responses to new information. The RCT indicates that beliefs directly affect migration: individuals that receive new information migrate 40% more to Nairobi relative to the control group. We can capture the effect of information on migration, through the elasticity of migration flows with respect to belief revision. Based on the experimental evidence we have that,

$$\frac{\%\Delta(\text{Migration})}{\%\Delta(\text{Beliefs})} = \frac{40\%}{20\%} = 2$$

Table 7: Previous migrants and underestimation of urban earnings within villages

	(1)	(2)
Log migration share	-0.178^{***}	
	(0.015)	
Migrant family member		-0.228^{***}
		(0.041)
Household-head controls	Yes	Yes
Village fixed effects	Yes	Yes
Observations	2,522	3,686
R^2	0.128	0.065

Notes. The dependent variable is the log ratio of actual to perceived urban earnings, so a negative coefficient means smaller underestimation. Migrant family member equals one if the respondent has a migrant family member. Household-head controls are the age, education and earnings of the male and female household heads, and an indicator for households without a male head. Standard errors in parentheses. *** $p < 0.01$. Data from Baseler (2023).

To compare the magnitude of this response in our framework, we raise $m^i_{r,1}$ by $\log(1.20)$ in the model estimated with IFLS data, keeping $v_{r,1}$ and $\xi_{r,1}$ fixed. In this case, the migration rate rises from 10.53% to 14.09%. This is an increase of 3.56 percentage points, or 33.79% relative to the benchmark, and implies an elasticity of 1.69. This response is not targeted. The model estimated for Indonesia therefore delivers a response of a similar magnitude to the response in the Kenyan experiment.

The Kenyan data also allow us to test our estimation strategy. In Indonesia, we infer information frictions indirectly, because beliefs are not observed. The Kenyan data contain both the migration and earnings moments used in Section 4.2 and reported beliefs, so we can estimate the model both ways and compare the results. We use 15 villages in the data of Baseler (2023). In the indirect inference columns of Table 8 we apply the same indirect inference strategy as in Indonesia. We match the cross-village correlation between mean log migrant earnings in 2017 and migrant counts in 2019, which is 0.271, close to the 0.242 we target in Indonesia. We use the dispersion of migrant wage gaps within each village to discipline the private signals.

In the reported-belief columns we use reported beliefs to discipline information frictions. We set σ_0^2 to the pooled variance of reported log beliefs and $\sigma_{\eta_r}^2$ to their variance within village r. To discipline μ_0, we choose the prior mean so that the mean posterior belief in 2017, after workers update using their private signals, equals the mean reported belief. In both calibrations we divide the variance of selected migrant earnings, V_r^m, by the number of previous migrants in 2017, $M_{r,0}$.

Table 8: Estimation of information and learning parameters in Kenya

		Indirect inference		Reported beliefs	
Parameter	Description	Target	Value	Target	Value
σ_0^2	Prior variance	Correlation of 2017 mean log migrant earnings and 2019 migrant counts, 0.2709	29.2054	Pooled variance of reported log beliefs	0.7040
μ_0	Prior mean	Prior centered on true log urban productivity	0.0000	Mean reported log belief equals the model's mean posterior in 2017	−0.5137
σ_s^2	Urban skill variance	Pooled variance of log migrant earnings, accounting for selection	0.8507	Pooled variance of log migrant earnings, accounting for selection	0.8507
μ_s	Mean urban skill	Pooled mean log migrant earnings, accounting for selection	−0.3259	Pooled mean log migrant earnings, accounting for selection	−0.3259
$\sigma_{e_r}^2$	Public-signal noise variance	Selected earnings variance divided by the 2017 migrant count	0.0241–0.1279	Selected earnings variance divided by the 2017 migrant count	0.0241–0.1279
$\sigma_{\eta_r}^2$	Private-signal noise variance	Within-village dispersion of previous migrants' rural-urban wage gaps	0.3268–15.7244	Within-village variance of reported log beliefs	0.3768–1.1169

Notes. Both calibrations use 15 villages from Baseler (2023). The ranges give village minima and maxima. The skill mean μ_s is in log units relative to mean log rural earnings.

We use the selection correction from the Indonesian calibration to recover urban skill moments in both calibrations. For Kenya we report all log earnings relative to mean log rural earnings. This changes only the units of μ_s and $\log A_r$: the signal z_r and the migration decision depend on differences in log earnings, so no result is affected. The two calibrations imply identical skill moments and public-signal noise variances. The two calibrations differ instead in the uncertainty workers face before they observe previous migrants. The prior variance is 29.21 with indirect inference and 0.70

with reported beliefs, and the private-signal variances are more dispersed with indirect inference. Reported beliefs also imply a prior mean 0.51 log points below true urban productivity, consistent with the underestimation documented above, whereas indirect inference centers the prior on the truth. The reported-belief calibration therefore allows us to test how well our estimation strategy quantifies the role of information frictions relative to disciplining information frictions directly with beliefs.

Table 9 reports the results of the same counterfactual experiments for Kenya. Both calibrations reproduce baseline migration and persistence by construction, so the comparison concerns the counterfactuals, which are not targeted. Moving to perfect information raises migration from 10.80% to 13.08% when we use indirect inference and to 13.46% when we use reported beliefs, increases of 2.28 and 2.66 percentage points. In both cases persistence falls from 0.63 to 0.27, and the welfare gain for rural individuals at $t = 1$ is 2.42 and 2.13 percent. Eliminating all frictions raises migration to 36.37% in both, and persistence falls to 0.15 in both.

Table 9: Effects of information frictions in Kenya

	Information frictions disciplined by	
Outcome	Indirect inference	Reported beliefs
Baseline		
Aggregate Migration Rate (%)	10.80	10.80
Migration Persistence	0.6310	0.6310
Perfect Information		
Aggregate Migration Rate (%)	13.08	13.46
Migration Persistence	0.2721	0.2726
Rural Period-2 Welfare Gain (%)	+2.42	+2.13
+ *No Migration Costs*		
Aggregate Migration Rate (%)	36.37	36.37
Migration Persistence	0.1465	0.1465

Notes. Both calibrations reproduce the baseline migration rate and persistence by construction. Migration persistence is the cross-village correlation between 2017 and 2019 migrant counts. Welfare changes are relative to the fitted baseline.

Although indirect inference centers the prior on the truth while reported beliefs place it 0.51 log points below, and the estimated information parameters differ, the two calibrations project very similar aggregate migration and persistence under perfect information and when all barriers are

removed. The reason is that by 2019 workers have observed previous migrants. In every village and under both calibrations, the public signal receives at least 66 percent of the posterior weight, and its precision is disciplined by the same migrant counts and selected earnings variance. With reported beliefs, the prior receives at most 14 percent of the posterior weight, so its lower mean reduces posterior means by at most about 0.07 log points. The posterior uncertainty that governs migration decisions is therefore similar in the two calibrations, although the prior variance is not. Thus, disciplining information frictions directly with reported beliefs yields nearly the same effects of information on migration, persistence and welfare as our indirect inference strategy.

We next raise beliefs by 20% in the two Kenyan calibrations, as we did for Indonesia. This increases migration by 37.67% with indirect inference and by 37.63% with reported beliefs, implying elasticities of 1.88 (Table 10). Neither calibration uses the experiment, yet both elasticities are very close to the experimental elasticity of 2.00. The model therefore reproduces the migration response to the information treatment whether information frictions are disciplined indirectly or with reported beliefs. This supports the use of indirect inference in Indonesia, where beliefs are not observed.

Table 10: The effect of belief changes on migration

	Kenya RCT	IFLS model	Kenya model	
			Indirect inference	Reported beliefs
Belief change (%)	20	20	20	20
Migration change relative to baseline (%)	40	33.79	37.67	37.63
Elasticity of migration with respect to beliefs	2.00	1.69	1.88	1.88

Notes. The Kenya model columns use the two calibrations in Table 8. The experimental comparison is based on Baseler (2023).

6 Conclusions

We demonstrate that uncertainty about urban earnings opportunities can be an important source of migration costs and spatial misallocation in developing countries. Empirically, we show that districts in Indonesia that have more and earlier access to internet, are more likely to see their residents migrate to urban centers. This is true even when the force that instigates the migration

is a natural disaster. We also show that prior migrant average incomes and subsequent migration flows from the same districts are positively correlated, and that migration flows are highly persistent over time.

To rationalize these observations we develop a theory of information frictions in rural–urban migration. Our theory features decentralized social learning where rural workers learn, albeit imprecisely, from those of their own district that have migrated to the city before them. In the face of uncertainty about urban earnings this localized information transmission can lead to spatial differences in patterns of migration. We identify information frictions from the correlation of previous migrant earnings and migration flows, as well as the dispersion of rural-urban wage gaps within districts of previous migrants. We allow for residual barriers to labor mobility that encompass all other migration frictions. In the estimated model, perfect information raises aggregate migration by about 40 percent, from 10.5 to 14.7 percent, and lowers persistence from 0.80 to 0.26. This sharp decline shows that information frictions are an important source of persistent migration patterns across districts. Persistence remains positive under perfect information because other migration barriers remain in place. Removing information frictions yields a second-period consumption-equivalent gain of 2.47 percent for the rural-origin population. The migration increase is about one-fifth of that obtained when all barriers are removed. Reported beliefs from Kenya validate the local learning mechanism and the model's quantitative results. Calibrations using reported beliefs and indirect inference yield similar migration, persistence and welfare effects, and both closely match the untargeted migration response to the randomized information treatment. This agreement provides a credible basis for applying our framework in other settings where migration and earnings moments are available but beliefs are unobserved. The approach therefore extends the study of information frictions beyond surveys that directly elicit beliefs. These findings have a direct policy implication. Information interventions that improve rural workers' knowledge of urban earnings opportunities can increase rural-to-urban migration by reducing uncertainty and correcting underestimation of the returns to moving.

References

Adamopoulos, T. (2025). Spatial integration and agricultural productivity: Quantifying the impact of new roads. *American Economic Journal: Macroeconomics*, 17(1):343–378.

Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., and Wei, X. (2024). Land security and mobility frictions. *The Quarterly Journal of Economics*, 139(3):1941–1987.

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2022). Misallocation, selection, and productivity: A quantitative analysis with panel data from China. *Econometrica*, 90(3):1261–1282.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–1697.

Adamopoulos, T. and Restuccia, D. (2022). Geography and agricultural productivity: Cross-country evidence from micro plot-level data. *The Review of Economic Studies*, 89(4):1629–1653.

Adamopoulos, T. and Shahidi, Z. (2025). Spatial patterns of internal migration across countries. Unpublished manuscript, York University.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Alvarez, J. A. (2020). The agricultural wage gap: Evidence from Brazilian micro-data. *American Economic Journal: Macroeconomics*, 12(1):153–173.

Alvarez-Cuadrado, F., Amodio, F., and Poschke, M. (2019). Selection and absolute advantage in farming and entrepreneurship. Technical report, IZA Discussion Papers.

Angeletos, G.-M. and La'O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.

Asher, S. and Novosad, P. (2020). Rural roads and local economic development. *American Economic Review*, 110(3):797–823.

Barnett-Howell, Z., Baseler, T., Ginn, T., and Gordeev, S. (2026). Reaching the novice or nudging the expert? Information delivery and the returns to migration. Working paper.

Baseler, T. (2023). Hidden income and the perceived returns to migration. *American Economic Journal: Applied Economics*, 15(4):321–352.

Beegle, K., De Weerdt, J., and Dercon, S. (2011). Migration and economic mobility in Tanzania: Evidence from a tracking survey. *Review of Economics and Statistics*, 93(3):1010–1033.

Boltz, M., Marazyan, K., and Villar, P. (2019). Income hiding and informal redistribution: A lab-in-the-field experiment in Senegal. *Journal of Development Economics*, 137:78–92.

Boltz, M., Marazyan, K., and Villar, P. (2020). Is informal redistribution costly? evidence from a lab-in-the-field experiment in Senegal. *The World Bank Economic Review*, 34(Supplement_1):S72–S78.

Brooks, W. and Donovan, K. (2020). Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua. *Econometrica*, 88(5):1965–1997.

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

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the China trade shock. *Econometrica*, 87(3):741–835.

Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of Economic Growth*, 1:679–741.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–2678.

David, J. M., Hopenhayn, H. A., and Venkateswaran, V. (2016). Information, misallocation, and aggregate productivity. *The Quarterly Journal of Economics*, 131(2):943–1005.

De Janvry, A., Emerick, K., Gonzalez-Navarro, M., and Sadoulet, E. (2015). Delinking land rights from land use: Certification and migration in Mexico. *American Economic Review*, 105(10):3125–3149.

Eckert, F. and Peters, M. (2022). Spatial structural change. Technical report, National Bureau of Economic Research.

Fogli, A. and Veldkamp, L. (2011). Nature or nurture? Learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138.

Gollin, D., Lagakos, D., Ma, X., and Mandi, S. (2025). The dynamics of agricultural productivity gaps: An open economy perspective. Unpublished manuscript, Boston University.

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

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Gröger, A. and Zylberberg, Y. (2016). Internal labor migration as a shock coping strategy: Evidence from a typhoon. *American Economic Journal: Applied Economics*, 8(2):123–153.

Hamory, J., Kleemans, M., Li, N. Y., and Miguel, E. (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, 19(3):1522–1555.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60(1):126–142.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. *Handbook of Economic Growth*, 2:855–941.

Herrendorf, B. and Schoellman, T. (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, 10(2):1–23.

Jakiela, P. and Ozier, O. (2016). Does Africa need a rotten kin theorem? Experimental evidence from village economies. *The Review of Economic Studies*, 83(1):231–268.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives*, 34(3):174–192.

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

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–980.

Meghir, C., Mobarak, A. M., Mommaerts, C., and Morten, M. (2022). Migration and informal insurance: Evidence from a randomized controlled trial and a structural model. *The Review of Economic Studies*, 89(1):452–480.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village India. *Journal of Political Economy*, 127(1):1–46.

Morten, M. and Oliveira, J. (2024). The effects of roads on trade and migration: Evidence from a planned capital city. *American Economic Journal: Applied Economics*, 16(2):389–421.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

Pulido, J. and Swiecki, T. (2021). Barriers to mobility or sorting? Sources and aggregate implications of income gaps across sectors and locations in Indonesia. *Unpublished Working Paper, Vancouver School of Economics*.

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.

Squires, M. (2024). Kinship taxation as an impediment to growth: Experimental evidence from Kenyan microenterprises. *The Economic Journal*, 134(662):2558–2579.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, 109(5):1843–1872.

Veldkamp, L. L. (2006). Information markets and the comovement of asset prices. *The Review of Economic Studies*, 73(3):823–845.

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