Skill Transferability, Migration, and Development:
Evidence from Population Resettlement in Indonesia[*]

Samuel Bazzi[†]
Boston University and BREAD

Arya Gaduh[‡]
University of Arkansas

Alexander D. Rothenberg[§]
RAND Corporation

Maisy Wong[¶]
University of Pennsylvania

November 2015

Abstract

We use a natural experiment in Indonesia to provide causal evidence on the role of location-specific human capital and skill transferability in shaping the spatial distribution of productivity. From 1979–1988, the Transmigration Program relocated two million migrants from rural Java and Bali to new rural settlements in the Outer Islands. Villages assigned migrants from regions with more similar agroclimatic endowments exhibit higher rice productivity and nighttime light intensity one to two decades later. We find some evidence of migrants' adaptation to agroclimatic change. Overall, our results suggest that regional productivity differences may overstate the potential gains from migration.

JEL Classifications: J43, J61, O12, O13, O15, R12

Keywords: Internal Migration, Comparative Advantage, Spatial Labor Allocation, Agricultural Adaptation

---

1 Introduction

Throughout history, soil and climate conditions have shaped migration patterns and influenced the spread of human capital and technology. Steckel (1983) and Diamond (1997) document a striking tendency for migrants and technologies to diffuse east–west rather than north–south in the process of settling the agricultural frontier. Griliches (1957) and Comin et al. (2012) highlight a similar pattern of spatial diffusion within agroclimatic zones. Historical settlement patterns like these can have persistent impacts on today's economic landscape (Ashraf and Galor, 2013; Nunn and Qian, 2011; Putterman and Weil, 2010). They suggest that similarity in agroclimatic conditions is important for the transferability of skills. Yet, we have limited evidence on these relationships because skill transferability is difficult to measure, migrants endogenously sort into places where their skills are transferable, and these spatial diffusion processes are slow and often confounded by time trends.

   This paper uses a remarkable policy experiment in Indonesia to provide causal evidence on the role of location-specific human capital and skill transferability in shaping productivity. Between 1979 and 1988, the Transmigration Program relocated two million voluntary migrants (hereafter, transmigrants) from the Inner Islands of Java and Bali to *newly created* agricultural settlements in the Outer Islands. We develop a novel proxy for skill transferability based on the similarity in agroclimatic conditions between two locations. Using the plausibly exogenous assignment of transmigrants to destination villages, we identify and estimate large causal impacts of location-specific human capital on productivity, suggesting some farming skills may not be easily transferable across space. The exogenous assignment addresses a pervasive identification problem in the study of migration (Roy, 1951), and our measure of agroclimatic distance helps to quantify the importance of skill specificity.

   Our findings are significant for several reasons. First, recent debate questions whether labor is misallocated across space and whether migration can equalize regional productivity differences (Munshi and Rosenzweig, forthcoming; Young, 2013). If some skills are not readily transferable across locations, then spatial productivity gaps may not represent arbitrage opportunities. Second, with growing risks of population displacement from natural disasters, conflict, or climate change, various governments have started planning for resettlement (see IPCC, 2014; de Sherbinin et al., 2011).[1] Extreme weather events are expected to uproot over 60 million people in South Asia alone (Stern, 2007). Third, understanding how abrupt changes in agroclimatic conditions affect productivity is important in light of climate change. Many rainfed, subsistence farmers in developing countries—not unlike Indonesia's transmigrants—may lack the resources to adapt.

   The Transmigration program provides a rich empirical context for studying the relationship between skill transferability and productivity. Designed to alleviate overpopulation concerns in rural Java/Bali and to develop the Outer Islands, the government-run program provided households with free transport to new settlements and two hectare farm plots assigned by lottery. A large spike in global oil prices funded a massive increase in the scale of the program in the late 1970s. However, because of time, information, and logistical constraints, many activities were undertaken on an ad hoc, "plan-as-you-proceed" basis (World Bank, 1988). This gave rise to plausibly exogenous variation in the assignment

---

[1] In addition to weather-induced displacement, 10 million people are displaced annually by infrastructure development (World Bank, 1999), and around 36 million have been displaced by conflict according to the World Bank. Relocation programs are found in many developing countries, including China, India, and Brazil (see Kinsey and Binswanger, 1993).

of transmigrants to new settlements, which we confirm through a battery of identification checks. In practice, the program's unprecedented spatial scope meant that migrants from diverse origins across Java/Bali are observed across a range of agroclimatic conditions in the destinations.

Despite being one of the largest resettlement policies ever implemented, relatively little is known about the economic impact of the Transmigration program due to a lack of data. We collected data from two new sources: a 1998 census of Transmigration villages and planning maps used to identify settlements in the Outer Islands. We combine these sources with granular agroclimatic data, individuals' birth districts and other demographics from the 2000 Population Census, and village-level agricultural activity from a 2002 administrative census.

An important innovation of this study is our proxy for skill transferability. Farming often requires location-specific production methods and associated technical know-how. Our proxy, *agroclimatic similarity*, is higher when the agroclimatic endowments (and hence growing conditions) between migrants' origin and destination regions are more similar. We construct this measure using several sources of geospatial data capturing topography, hydrology, climate, and predetermined soil characteristics from the *Harmonized World Soil Database*.

Our empirical strategy compares Transmigration villages with a high share of migrants from similar origins to observably identical Transmigration villages that have a high share of migrants from dissimilar origins. Using a multi-location Roy model, we show that agroclimatic similarity provides a novel and exogenous measure of comparative advantage.[2] Farmers can transfer their human capital more successfully if destinations more closely resemble their birth locations. Hence, for a given destination, migrants from similar origins have greater comparative advantage relative to migrants from dissimilar origins.

We find that skill transferability has large effects on village-level rice productivity, our primary outcome. Rice was the focal crop of the program and is the primary staple for Indonesia and more than half of the world. It is grown on 144 million farms worldwide (more than for any other crop), and is the crop expected to be most vulnerable to climate change (Mohanty et al., 2013; Peng et al., 1995). Our estimates imply that, on average, an increase in agroclimatic similarity by one standard deviation leads to a 20 percent increase in village-level rice productivity.[3] This translates to an additional 0.5 tons per hectare—an effect size roughly equivalent to twice the productivity gap between farmers with no schooling versus those that have completed junior secondary.

We show further that the productivity gains from skill transferability are larger in adverse growing conditions. Semiparametric regressions reveal a concave relationship, with the steepest productivity losses for villages whose migrants are from the most dissimilar origins. Moreover, consistent with agronomic literature on the complexities of soil management (De Datta, 1981), barriers to transferability appear to be greatest for soil-specific knowledge.

Beyond rice, we also find that agroclimatic similarity is important for other food crops that are sensitive to growing conditions. In contrast, agroclimatic similarity has small and insignificant effects on the productivity of cash crops that have less location-specific farming methods than rice. This result serves as a placebo check, mitigating concerns that our proxy for skill transferability is confounded with

---

[2]The identification problem of endogenous sorting based on unobservable comparative advantage was formalized by Heckman and Honoré (1990) and spans multiple fields in economics. Examples can be found in labor (Dahl, 2002), spatial and urban (Combes et al., 2008), development (Suri, 2011), and trade (Costinot et al., forthcoming).

[3]Our index is scaled between zero and one, with a relatively large standard deviation of 0.14.

unobservable determinants of productivity common across crops.

These findings provide new evidence on barriers to adaptation in response to abrupt changes in agroclimatic conditions. The persistence of effects over two decades is consistent with historical research showing that American farmers faced prolonged difficulties adjusting to sudden agroclimatic change associated with the 1930s Dust Bowl (Hornbeck, 2012) and the early settlement of farmland in new climates along the Western frontier (Olmstead and Rhode, 2011). These barriers to adaptation are particularly salient in developing countries today (e.g., BenYishay and Mobarak, 2014; Conley and Udry, 2010). Our estimates of productivity losses for a major staple crop like rice suggest that imperfect skill transferability may imply adjustment costs not accounted for in existing climate change projections.

We next investigate several adaptation mechanisms and find that crop adjustments and language skills are relatively more important. We use data on ethnolinguistic homelands (from the *Ethnologue* data) to measure the *linguistic similarity* between transmigrants' languages and the indigenous language in nearby Outer Island villages. Linguistic similarity yields significant positive effects on rice productivity, and appears to be more important in places with greater scope for interacting with natives. This is consistent with learning and other productivity-enhancing social interactions.

Language skills are also important for occupational adjustment. A one standard deviation increase in linguistic similarity leads to a 1.8 percentage point (p.p.) greater likelihood of Java/Bali migrants choosing trading and services occupations (relative to mean of 9.9 percent) whereas a one standard deviation increase in agroclimatic similarity leads to a 0.9 p.p. greater likelihood of farming (relative to a mean of 62 percent). These patterns are consistent with occupational sorting based on comparative advantage, but the magnitudes for farming are relatively small.

Turning to crop adjustments, we find that cash crops generate more revenue in villages with low agroclimatic similarity. This is in line with Costinot et al. (forthcoming) who use a simulated trade model to highlight the welfare-enhancing effects of crop adjustment in response to climate change. Additionally, we find limited evidence of selective ex-post migration from settlement areas.

To gauge the overall effects of these adjustments, we show that agroclimatic similarity still has positive effects on the level of economic development in 2010, as proxied by nighttime light intensity (Henderson et al., 2012). Our estimates imply that a one standard deviation increase in agroclimatic similarity leads, on average, to 6–12 percent greater income by 2010. Coupled with the large effects on rice productivity and the evidence on adaptation, these results suggest that the adjustment process was costly and incomplete.

We investigate the policy implications of our results by simulating an alternative assignment process that improves migrants' origin-to-destination match quality. Our simulations approximate an optimal reallocation based on agroclimatic similarity and suggest that planners could have achieved 27 percent higher aggregate rice yields. Next, we estimate the place-based impact of the Transmigration program by comparing Transmigration villages to planned settlement areas that were never assigned transmigrants. These counterfactual, almost-settled villages exist because the program was abruptly halted due to budget cutbacks. Despite an increase in population density and agricultural output associated with extensification, we find small and insignificant effects on productivity, due in part to the persistent effects of agroclimatic similarity in treated villages. These policy exercises demonstrate the importance of matching people (skills) to places (production environments) when designing resettlement schemes.

Our study contributes to the literature on migration and the spatial (mis)allocation of labor in developing countries (Au and Henderson, 2006; Bryan et al., 2014; Gollin et al., 2014). Using survey data for 65 countries, Young (2013) argues that rural-urban wage gaps are explained by efficient geographic sorting rather than barriers to mobility. We focus on rural-to-rural migration, which has been understudied despite its importance in overall flows (see Lucas, 1997; Young, 2013). Our key innovation is to provide causal evidence that complementarities between heterogeneous individuals and heterogeneous places can give rise to persistent spatial productivity differences. In short, skill specificity implies that regional productivity differences may overstate the potential gains from migration.

Our results also complement recent evidence on location-specificity and migrant outcomes. Atkin (forthcoming) and Michalopoulos (2012) show that migrants in India and Africa tend to consume and grow crops that are predominant in their native origins. We provide causal evidence on productivity gains from high quality origin-to-destination matches. Our focus on agriculture is important given that it employs 1.3 billion people globally (World Bank, 2009) and is at the core of ongoing debates about world income inequality (see Caselli, 2005). Our findings shed light on how soil-specificity may contribute to the relatively slower spatial diffusion of technology within agriculture compared to non-agricultural sectors (Schultz, 1975; Rodrik, 2013).

Finally, the reduced form skill transfer elasticity that we estimate parallels research on labor mobility and skill specificity in other contexts. Friedberg (2000), Lubotsky (2007), and Abramitzky et al. (2014) study the speed of economic assimilation of immigrants in Israel and the United States. Poletaev and Robinson (2008) and Gathmann and Schönberg (2010) find sizable productivity losses as workers move between occupations with dissimilar tasks. Giovanni et al. (2015) and Klein and Ventura (2009) calibrate models of migration in which skill transferability across countries is a crucial parameter in assessing the productivity gains to greater labor mobility. However, in many contexts, migrants tend to move to locations where their skills are transferable. The exogenous assignment of the Transmigration program allows us to observe migrants with high *and* low quality matches, providing a unique opportunity to address this sorting bias.

The remainder of the paper proceeds as follows. Section 2 provides background on the Transmigration program. Section 3 describes our data and presents our key proxies for skill transferability and development outcomes. Section 4 develops our theoretical framework and empirical strategy in the context of a multi-sector Roy model. Section 5 presents our main results. Section 6 reports the results of policy exercises. Section 7 concludes.

2   Indonesia's Transmigration Program

Like many countries, the spatial distribution of Indonesia's population has historically been highly skewed. In the 1970s, there were concerns that the Inner Islands of Java and Bali were overpopulated while the *Outer Islands*—Sumatra, Sulawesi, Kalimantan, Maluku, Nusa Tenggara, and Papua—were relatively unsettled. Indonesia's Transmigration program was designed primarily to alleviate these perceived population pressures. The program relocated households from rural areas in Java and Bali to rural areas in the Outer Islands. Planners hoped that the program would increase national food production (especially rice) by moving farmers to unsettled areas, and also promote nation building by integrating

diverse ethnic groups (Kebschull, 1986; MacAndrews, 1978).

Our study focuses on the most intensive period of the program from 1979 to 1988.[4] At that time, the program supported rainfed food crops because Indonesia was the world's largest importer of its primary staple (rice), and annual crops promoted early self-sufficiency. Moreover, farmers in Java/Bali had centuries of experience growing rice (Geertz, 1963). The program targeted entire families for resettlement; participating couples had to be legally married, with the household head between 20 and 40 years of age. In practice, most participants were poor, landless agricultural laborers, and negatively selected (in terms of schooling) relative to the typical outmigrant from rural Java/Bali at the time (Kebschull, 1986).[5]

The Transmigration program was one of the largest resettlement programs of its time and involved complex logistics in both Java/Bali and the Outer Islands. Participating households, who were almost entirely volunteers, would sell their assets and leave for transit camps located in each of the four provinces of Java/Bali. Here, transmigrants would wait to be transported in groups to the Outer Islands. At the same time, in destination areas, program officials identified previously uncleared land reserves that could be developed into settlements, prepared for agricultural use, and connected to the road network. Transmigrants were given free transport to these new settlements, free housing, a two hectare plot of agricultural land allocated by lottery upon arrival, and provisions for the first few growing seasons, including seeds, tools, and food.

2.1 The Assignment Process

The process by which households from Java/Bali were allocated to Transmigration settlements across the Outer Islands is central to our identification strategy. Numerous reports indicate that the process was less rigorous than planners had hoped. For example, Hardjono (1988) observes "(a)s a consequence of the focus on numbers, the land use plans developed during the 1970s were totally abandoned. Transmigrants were placed on whatever land was submitted by provincial governments for settlement purposes."

An array of time, information, and institutional constraints prevented policymakers from systematically assigning transmigrants to destination villages. First, sharp changes in world oil prices strongly affected government revenue, leading to a rapid expansion and sudden contraction of the program. Figure 1 shows large fluctuations in the annual number of transmigrants placed coinciding with the rise and fall of oil prices in the late 1970s and early 1980s. Due to the rapid expansion, a number of program activities were taken from the Directorate General of Transmigration (DGT) and delegated to separate government agencies to speed up the settlement process. Inter-agency coordination problems made it more difficult to carefully match transmigrants (whose information was collected by DGT) to their Outer Island settlements (developed under the Ministry of Public Works).

Second, planners had neither the interest nor the resources to match transmigrants on the basis of

---

[4]The Transmigration program began during the colonial period, but it received a major overhaul during Suharto's third development period, or *Pelita* III (1979-1983). Less than 600,000 people were resettled under the colonial program and post-independence waves (1945-1968) (Hardjono, 1988; Kebschull, 1986). In contrast, the program resettled 1.2 million people in *Pelita* III and initially planned to move 3.75 million people in *Pelita* IV (1984-1989). The total program budget during *Pelita* III and IV was approximately $6.6 billion (in 2000 USD) or roughly $3,330 per person moved (see World Bank, 1982, 1984).

[5]On average, Java/Bali-born individuals who moved to Transmigration villages had 0.5 fewer years of schooling compared to non-migrants from their origin district cohort (based on the 2000 Population Census discussed in Section 3). By contrast, individuals from the same cohorts that moved to urban areas in Java/Bali or to the Outer Islands had 3–4 more years of schooling compared to stayers.

agroclimatic conditions. Many planners believed that Javanese and Balinese farmers had superior farm-
ing skills and could perform better than Outer Islanders in any environment (Dove, 1985). They hoped
that transmigrants would transfer some of their Javanese farming know-how to the Outer Islands. More-
over, matching transmigrants' skills to destinations would have required data on individual farming
skills and up-to-date information on growing conditions in available settlements—details largely un-
available at the time.

Third, the coincidental timing of transmigrants' arrival to the transit camps and the opening of new
settlements in the Outer Islands played a key role in determining where transmigrants were placed
(Hardjono, 1988). Most transmigrants did not wait long at the transit camps and were transported to the
Outer Islands within a few days. With only four provincial transit camps for the 119 origin districts, this
process ensured a mix of origins in each camp and ultimately each settlement. Furthermore, motivated
by the nation-building goals of the program, planners often assigned groups of migrants from each of
the four provinces to a single settlement (Levang, 1995).

Fourth, participants could not choose their destination in the Outer Islands (Levang, 1995). Previ-
ous studies show that just prior to departure, transmigrants were ill-informed about the geographical
location and agricultural systems in the areas where they were sent. In a pre-departure survey of 348
transmigrant families, Kebschull (1986) found that 82 percent knew nothing about the local agroclimatic
conditions, and most transmigrants expected to pursue the same sort of (rice) farming activities they had
been practicing in their origin villages.

All of these factors resulted in significant transmigrant diversity at destinations. In our data de-
scribed below, the median Transmigration village has Java/Bali migrants from 46 sending districts (out
of 119). Using an origin-district Herfindahl index, we find that in the median village, there was only a 12
percent chance that two randomly chosen transmigrants were from the same origin district in Java/Bali.

2.2 External Validity

The Transmigration program provides a laboratory to study the transferability of farming skills across
growing conditions. This is relevant for several reasons. First, our results are particularly informative
for rural-to-rural migration, which comprises population flows that are 1.5 to 2 times greater than those
from rural-to-urban migration (Young, 2013). Given the focus on agriculture, our context is less well
suited to study rural-to-urban migration. However, in Section 5.2, we investigate the importance of
language skills and social interactions, which confer economic benefits in rural settings such as ours as
well as in urban labor markets.

Second, climate change is expected to bring abrupt changes to growing conditions faced by farmers.
This could arise from climate-induced displacements or sudden changes in growing conditions due to
extreme weather events and natural disasters. The International Organization for Migration estimates
that 200 million people may become environmentally-induced migrants by 2050. Like transmigrants,
many farmers vulnerable to climate shocks lack the resources to move or adapt to sudden changes.

Third, our investigation of rice and other food crops is important because research shows that a
large class of annual food crops is particularly sensitive to growing conditions (Glover et al., 2010).[6]

---

[6]Cox et al. (2006) note that "...annual crops require seedbed preparation, precisely timed inputs and management, and good
weather during narrow time windows. With shorter growing seasons and less extensive root systems, annual crops provide

These crops play a central role in sustaining global food security, are planted on almost 70 percent of the world's arable land, and provide close to 70 percent of total calories consumed (Beddow et al., 2010).

Fourth, the program's resettlement of rural households to previously unsettled land offers a unique lens into the historical process of settling the agricultural frontier, which has important implications for the spatial distribution of economic activity today. The scope of resettlement, the remoteness of the new villages, and the common initial conditions for all settlements allow us to isolate the causal impact of skill transferability in a way that has not been feasible in slowly changing historical contexts. These lessons will also be relevant in African countries that hold half of the world's untilled arable soils and are implementing large-scale redistribution policies that have transferred land to many smallholder farmers (FAO, 2010; World Bank, 2013).

Finally, the Transmigration program is one of the world's largest government-sponsored, rural-to-rural resettlement schemes.[7] To date, resettlement has affected millions of households, cost billions of dollars, and is growing in importance as millions are expected to be displaced by extreme weather events, infrastructure development, and conflict.

3 Data: Measuring Skill Transferability and Its Effects

Our main analysis focuses on 814 villages that were created under the Transmigration program. We identify these villages from a newly digitized census of program settlements, produced by the Ministry of Transmigration (MOT) in 1998. Established between 1979 and 1988, these villages received an average of 1,885 migrants in their first year. Figure 2 shows that over half are on the island of Sumatra (482 out of 814 villages), but many are also found on Kalimantan (192) and Sulawesi (128), with smaller numbers in Eastern Indonesia. Below, we first discuss our proxy for skill transferability across locations. We then describe the key outcome variables.

3.1 Agroclimatic Similarity

We construct a novel measure of skill transferability, *agroclimatic similarity*, which captures how similar agroclimatic environments are between migrant origins and destinations. This proxy is similar in spirit to an index developed by Gathmann and Schönberg (2010) to measure the transferability of task-specific human capital across occupations. The ability to directly measure skill transferability across farming environments is an important innovation of our research design. We are able to do so because a wealth of agronomic research has identified and collected data on (predetermined) agroclimatic characteristics vital to farm output.[8]

---

less protection against soil erosion, manage water and nutrients less effectively, store less carbon below ground, and are less resilient to pests and abiotic stresses than are perennial plant communities."

[7]Other examples of rural resettlement schemes include the Polonoroeste program in Brazil that relocated 300,000 migrants between 1981 and 1988 at a cost of US$ 1.6 billion, villagization programs in Ethiopia that relocated 440,000 households between 2003 and 2005, the resettlement of 400,000 individuals in Africa due to dam construction, the resettlement of 4 million migrants in Mozambique between 1977 and 1984, and another 43,000 households that were relocated following floods in the 2000s (Arnall et al., 2013; de Wet, 2000; Hall, 1993; Taye and Mberengwa, 2013; World Bank, 1999). Additional resettlement programs can be found across a range of countries (Bauer et al., 2013; Beaman, 2012; Edin et al., 2003; Glitz, 2012; Sarvimäki et al., 2010).

[8]In a survey of Roy assignment models where workers sort into tasks based on comparative advantage, Autor (2013) notes that "no labor market data equivalent to agronomic data are available for estimating counterfactual task productivities."

We use data from the *Harmonized World Soil Database* (HWSD) and other sources to measure a vector $\mathbf{x}$ of agroclimatic characteristics including elevation, slope, ruggedness, altitude, distance to rivers and the sea coast, rainfall, temperature, and soil texture, drainage, sodicity, acidity, and carbon content. These characteristics, which we measure at a high spatial resolution, are fundamental components of agricultural productivity and exhibit considerable variation across Java/Bali and the Outer Islands (see Appendix Table B.1). All land attributes are either time-invariant or measured before the villages we study were created, and hence do not reflect settler activities. Since land and local climate characteristics change slowly, agroclimatic characteristics measured in the 1970s are still highly predictive of productivity in 2000.

We define the agroclimatic similarity between an individual's origin i and her destination j as:

$$agroclimatic\ similarity_{ij} \equiv \mathcal{A}_{ij} = (-1) \times \mathrm{d}\left(\mathbf{x}_i, \mathbf{x}_j\right) \tag{1}$$

where $\mathrm{d}\left(\mathbf{x}_i, \mathbf{x}_j\right)$ is the agroclimatic distance between locations i and j, using a metric defined on the space of agroclimatic characteristics. We observe origins at the district-level and hence construct the index based on measures of $\mathbf{x}$ in the destinations at that same spatial frequency. We use the sum of absolute deviations as the distance metric, converting each characteristic to z-scores before taking the absolute difference between origins and destinations. Then, $\mathrm{d}\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sum_g |x_{ig} - x_{jg}|$ projects these differences in G dimensions onto the real line. We multiply by (-1) so that larger differences correspond to lower values of agroclimatic similarity. Other distance metrics are considered in robustness checks.

We use $\mathcal{A}_{ij}$ to construct an agroclimatic similarity index for location j by aggregating across i using population weights:

$$agroclimatic\ similarity_j \equiv \mathcal{A}_j = (-1) \times \sum_{i=1}^{I} \pi_{ij}\, \mathrm{d}\left(\mathbf{x}_i, \mathbf{x}_j\right), \tag{2}$$

where π_{ij} is the share of migrants residing in Transmigration village j who were born in district i. To construct the migrant shares, we use the universe of microdata from the 2000 Population Census, which identifies each individual's district of birth and his or her current village of residence. Our main results use π_{ij} terms based on all individuals born in Java/Bali. We use $\mathcal{A}_j$ in our main village-level analysis but occasionally use $\mathcal{A}_{ij}$ for individual-level analyses. Therefore, we refer to $\mathcal{A}_{ij}$ ($\mathcal{A}_j$) as individual- (village-) level agroclimatic similarity.

3.2 Productivity and Development Outcomes

We study the impact of skill transferability on local economic development at the *village* level. We measure agricultural productivity using the triennial administrative census known as *Podes* (or Village Potential). The August 2002 round provides information on agricultural activities, including area planted and total yield for crops grown in 2001-2.

Our main outcome is rice productivity, measured in log yield (tons) per hectare.[9] Rice is the most widely grown crop across our settlements, and the average rice-growing village produced 2.5 tons per

---

[9]We winsorize yields at 20 tons/ha to account for measurement error, but our results are robust to alternative cutoffs.

hectare (Table 1). We also consider a set of secondary food crops, known collectively as *palawija*, which include maize, cassava, groundnuts, sweet potato, and soybeans. After rice, these are the most important food crops for Indonesia.

In addition to food crops, we investigate cash crops, which are mostly perennial tree crops that take 3 to 5 years to mature and can survive and bear fruit for up to 25 years. We observe 28 cash crops, the most important of which are palm oil, rubber, cocoa and coffee (see Appendix Table B.2). Compared to annual food crops, perennials are less sensitive to growing conditions (see footnote 6). Moreover, the critical farming tasks for perennial crops are more uniform across locations.[10] Overall, these differences in farming methods suggest that cash crops require relatively fewer location-specific skills than food crops.

These three groups of crops differ in the relative importance of location-specific skills as well as the crop-specific experience of the transmigrants. Location-specific skills are relatively more important for rice and *palawija* compared to cash crops. In terms of crop experience, almost all transmigrants were rice farmers, some transmigrants also grew certain *palawija* crops, but cash crops were not widely grown in Java/Bali around the time of the program.[11]

We capture broader economic development over the long-run using nighttime light intensity from the National Oceanic and Atmospheric Administration (see Henderson et al., 2012, for details). Light intensity is increasingly used to proxy for income in studies exploiting highly localized identifying variation as we do here (e.g., Hodler and Raschky, 2014; Michalopoulos and Papaioannou, 2014). Olivia and Gibson (2015) show that this proxy works well in capturing subnational variation in income across Indonesia. The level of nighttime light intensity in 2010 serves as our main measure of overall economic development at the village level.

In summary, we have six main data sources: (1) satellite data to capture light intensity, (2) soil attributes (HWSD), (3) temperature and precipitation data (UDel), (4) the 2000 Population Census, (5) the 2002 Village Census *Podes*, and (6) the 1998 Transmigration Census. We also use several auxiliary datasets, including FAO-GAEZ data on potential agricultural yields by crop, a 2004 survey (*Susenas*) that includes household-level rice productivity (but lacks migration data), the 1980 Population Census (to calculate pre-1979 variables), as well as planning maps published in the 1980s to identify planned but unsettled villages (discussed later). We provide further details on these data sources in Online Appendix A.

While the vast spatial scope of the program provides rich variation in agroclimatic attributes, it also poses data constraints. Transmigration villages represent less than five percent of the more than 60,000 villages in Indonesia. As a result, coverage limitations make it difficult to study productivity effects at the individual level (e.g., the *Indonesia Family Life Survey* (IFLS) includes only 50 households with Java/Bali-born migrants in Transmigration villages settled during our study period). Our individual-level dataset with the best coverage, the 2000 Population Census, covers all settlement areas, but does not record productivity outcomes such as wages or agricultural yields. Nevertheless, our granular maps

---

[10]For example, the tapping method and precise timing of harvesting intervals—each important for the productivity of rubber and palm oil, respectively—are mostly standardized (FAO, 1990; Verheye, 2010).

[11]In the late 1970s, less than five percent of farmers in Java/Bali were growing cash crops, according to the 1976 and 1980 (inter-)Census. While some transmigrants may have had prior experience growing cash crops like coffee, cocoa and rubber, none had experience growing palm oil, which was the most widely grown cash crop in Transmigration villages in the early 2000s.

and administrative censuses enable us to measure productivity at the village level.

4 Empirical Framework

This section describes our conceptual framework. We first explain how agroclimatic similarity proxies for skill transferability across locations and serves as a measurable source of comparative advantage. We then derive our key estimating equation and discuss identification.

4.1 Multisector Roy model

Following Dahl (2002), we adapt the classic two-sector Roy (1951) model to a setting in which heterogeneous farmers sort across heterogeneous locations. For now, we assume everyone is a rice farmer and abstract from unobservables to highlight the observable determinants of productivity central to our hypotheses. There is a discrete set of locations, indexed by $j = 1, ..., J$. Because they have distinct farming environments, locations are differentiated by a bundle of characteristics, which we denote using a fixed $(G \times 1)$ vector, $\mathbf{x}_j$. Individual farmers, indexed by i, are born into a birth location, $b(i) \in \{1, ..., J\}$. Hereafter, we denote $\mathbf{x}_{b(i)}$ by $\mathbf{x}_i$ to simplify notation.

To grow rice, farmers have to perform many different tasks, including plowing fields, tilling soil, sowing seeds, watering, applying fertilizer, weeding, managing pests, and harvesting. Crop productivity depends, in turn, on three types of human capital: (i) general (education), (ii) location-specific, and (iii) crop-specific. Our model focuses on the transferability of location-specific human capital, whereby optimal behavior may differ across environments. For example, weeding, seeding and transplanting methods differ across wetland and dryland locations (Vergara, 1992), while the approaches to troubleshooting nutritional disorders depend on soil pH and other characteristics.

Individuals acquire knowledge of how to perform farming tasks that is specific to local growing conditions at their birth locations, captured by $\mathbf{x}_i$. This location-specificity, which captures notions of "latitude-specific" farming skills (Steckel, 1983) and "location-specific amenities" (Huffman and Feridhanusetyawan, 2007), is consistent with local learning models that show how heterogeneous growing conditions can hamper the spatial diffusion of farming knowledge (Foster and Rosenzweig, 2010).

We assume that farmers can only own one unit of land in their location of choice (where they both live and work), and we normalize the output price to one.[12] The (potential) value of output per unit of land owned by farmer i in location j is given by:

$$y_{ij} = \gamma \mathcal{A}_{ij} + \mathbf{x}_j' \boldsymbol{\beta}, \tag{3}$$

where $\mathbf{x}_j' \boldsymbol{\beta}$ maps observable agroclimatic characteristics of location j into productivity, and $\mathcal{A}_{ij}$ is our measure of individual agroclimatic similarity between locations from equation (1).

The key parameter of interest is γ. If skills are perfectly transferable across growing conditions, a migrant's origin does not matter and $\gamma = 0$. Conditional on $\mathbf{x}_j$, a positive γ implies $\mathcal{A}_{ij}$ increases

---

[12]Implicit in this normalization is that transmigrants are price takers, allowing us to ignore possible general equilibrium effects of productivity improvements. This is reasonable because the 814 Transmigration villages in our study are scattered across the Outer Islands (see Figure 2), and because they constitute such a small fraction of all villages in Indonesia.

productivity at the destination, above and beyond the effects of $\mathbf{x}_j$ (land quality) on output. For a given destination, farmers migrating from more similar origins are more productive because it is easier to transfer their farming skills, compared to farmers from dissimilar origins.

When we aggregate across individuals, our model sheds light on the role of comparative advantage in shaping the spatial distribution of productivity. Since higher similarity reflects a better match quality (or greater complementarity) between migrants' skills and local growing conditions, villages assigned a higher share of migrants from agroclimatically similar origins have higher quality-adjusted labor endowments. Such villages therefore have greater comparative advantage in farming than villages assigned a higher share of migrants from dissimilar origins.

A major challenge in estimating γ is the endogenous sorting of farmers across locations. Each farmer i has J potential outcomes, but we only observe the optimal outcome when farmers sort based on comparative advantage (Heckman and Honoré, 1990).[13] It is difficult to find an excludable instrumental variable in this multisector Roy model because location choice and productivity are often confounded (Combes et al., 2011). Furthermore, few instruments are capable of generating a strong first stage for *each* of the J potential locations that also satisfies the exclusion restriction.[14]

Figure 3 illustrates the program-induced variation in agroclimatic similarity, which is key to our identification strategy. We compare the distribution of $\mathcal{A}_j$ across Transmigration villages to other villages in the Outer Islands. We use π weights that include all migrants—both Java/Bali migrants and migrants born in other districts in the Outer Islands. Absent the policy, individuals in non-Transmigration villages appear to sort in a way that increases the agroclimatic similarity between origins and destinations. This sorting shifts the distribution for non-Transmigration villages to the right compared to Transmigration villages. This corroborates our concern about sorting biases and the endogeneity of agroclimatic similarity in typical, non-program villages. Low similarity individuals in Transmigration villages are crucial for our research design because they represent counterfactual outcomes that would be absent were it not for the assignment process of the program.

4.2 Empirical Strategy

We investigate the relationship between productivity and agroclimatic similarity. Our key regression is at the village level:

$$y_j = \gamma \mathcal{A}_j + \mathbf{x}_j' \boldsymbol{\beta} + \omega_j, \tag{4}$$

where γ measures the semi-elasticity of aggregate productivity with respect to average agroclimatic similarity for the village.

Our main regression compares observably identical destination villages with a high share of Java/Bali migrants from similar origins to villages that have a high share of Java/Bali migrants from

---

[13]This can be seen in a stylized two-sector Roy assignment model with two types of farms (e.g., *Lowlands* and *Highlands*) and two types of farmers (born in L and H, respectively). There are four potential outcomes: $y_{LL}, y_{HH}, y_{LH}, y_{HL}$. If farmers born in lowlands have a comparative advantage at growing rice in lowlands (relative to farmers born in highlands), and if farmers sort based on comparative advantage, then we would only observe two of the four outcomes, namely those associated with high similarity: y_{LL}, y_{HH}. In this case of perfect sorting, there is no observed variation in agroclimatic similarity.

[14]For example, Dahl (2002) and Bayer et al. (2011) argue that birth location affects location choice but not productivity. For us, birth location fixed effects are not excludable from the productivity equation because comparative advantage is a function of the proximity between origins and destinations.

dissimilar origins. The key sources of exogenous variation in our village-level index, $\mathcal{A}_j$, include: (i) variation in the absolute differences between predetermined agroclimatic characteristics ($\mathbf{x}$ in destinations versus origins), and (ii) variation in the share of Java/Bali migrants in destination village j who are from origin district i, π_{ij}.

In practice, endogenous location, crop, and occupation choices could undermine the comparison of rice productivity in high and low similarity villages. Ideally, to estimate skill transferability across locations in the agricultural context, we would want (i) to randomly assign farmers from many origins to many destinations (to rule out endogenous location choice), and (ii) to minimize selection biases due to crop and occupational choices.[15]

In addition to the exogenous relocation process discussed in Section 2, our research design has several features that approximate this ideal setting. First, the previously landless transmigrants embarked on the program with the goal of farming, and their newly acquired land tied the first generation movers to farming. Also, rice was grown by virtually all transmigrants prior to departure, and its pervasiveness across program villages makes it a natural focal crop. These features minimize the concern that differences in rice productivity could be driven by compositional differences among individuals who select into rice farming.

Balance Checks. Table 2 reports estimates from separate regressions of agroclimatic similarity on island fixed effects, natural advantages $\mathbf{x}_j$, and 24 different variables capturing (i) potential agricultural productivity based on FAO agronomic data, as well as (ii) measures of district population size, quality of housing and utilities, schooling, literacy, language skills, and sector of work for those living in villages near the Transmigration settlement in 1978. Recall that the Transmigration villages are *new* settlements, and hence there are no pre-1979 outcome measures for these villages.

The results show that agroclimatic similarity is uncorrelated with pre-program correlates of productivity. Out of 24 tests, only one is significant at the 5 percent level, and the difference is negative, which works against our findings. Importantly, agroclimatic similarity is not correlated with potential yields of rice or other major food and cash crops. This rules out first-order concerns about unobserved natural advantages.

To further address concerns about biases due to *rice-specific* natural advantages, we first identify the 100 Transmigration villages with the lowest potential rice productivity. Within those villages, individuals from origin districts in the bottom quintile of potential rice productivity in Java/Bali have significantly higher individual agroclimatic similarity than those coming from origin districts in the top quintile of potential productivity (see Appendix Figure B.1). If our index only proxied for rice-specific natural advantages, then it would be higher for migrants from the top quintile origins who were naturally advantaged to grow rice; instead, we observe the opposite. Overall, the evidence suggests that agroclimatic similarity is balanced across Transmigration villages and is not proxying for unobservable natural advantages prevailing in Java/Bali.

---

[15]We would also need farmers from *many* origins assigned to *many* destinations to estimate the average elasticity for the population. Consider the stylized two-by-two Highland/Lowland example in footnote 13. If we only observed farmers from lowland origins assigned to both types of destinations, we would worry that the elasticity we estimate may not be representative of skill transferability for farmers from highland origins. Similarly, we would be concerned if we only observed farmers from lowland and highland origins assigned to a single destination type.

5 Empirical Results

We first report large average effects of skill transferability on rice productivity. Next, we investigate where the barriers to transferability are most significant and discuss the portability of general agroclimatic skills across crops. We then explore several channels of adaptation and broader impacts of agroclimatic similarity on economic development. Finally, we rule out additional threats to identification, including ex-post sorting.

5.1 Effects of Skill Transferability on Productivity

We begin with estimates of γ for rice productivity in equation (4). Village-level agroclimatic similarity ($\mathcal{A}_j$) is based on the Java/Bali migrant weights, and $\mathbf{x}_j$ includes island fixed effects as well as the full set of predetermined agroclimatic endowments described in Section 3.1.[16] We cluster standard errors using the Conley (1999) GMM approach, allowing for arbitrary spatial correlation in unobservables between all villages within 150 kilometers of one another, but inference is robust to alternative clustering specifications. In all regressions, we rescale the independent variables so that we can read a one standard deviation impact directly from the tables.

The baseline result in Panel A of Table 3 implies that a one standard deviation (0.14) increase in the agroclimatic similarity index leads to a 20 percent increase in rice productivity (column 1). This suggests agroclimatic similarity is an important predictor of cross-sectional differences in rice productivity, translating into a level effect of an additional 0.5 tons per hectare for the average village (relative to a mean of 2.5 tons per hectare, see Table 1). This effect is large, equivalent to twice the productivity gap between rice farmers with no education and those who completed junior secondary school (estimated in auxiliary *Susenas* 2004 data). The magnitude is plausible, especially since our village-level productivity measure aggregates across multiple cropping seasons, and rice farmers in Indonesia report up to three harvest cycles per year.

This key result is robust to several important concerns about identification. First, in column 2, we show that the effect is stable after we drop $\mathbf{x}_j$ controls. This reduces the concern that the difference between high and low similarity villages is driven by comparing naturally advantaged and disadvantaged villages. In fact, the slight drop in the coefficient suggests a *negative* correlation between $\mathcal{A}_j$ and land quality, which works against our findings.

In column 3, we control for productivity differences that may be driven by origin-specific absolute advantages. In particular, we add (i) a π_{ij}-weighted average of predetermined origin controls, including potential rice productivity (i.e., all variables reported in Table 2), (ii) a π_{ij}-weighted average of physical distance to the origins, and (iii) four province-level aggregates of the origin district π_{ij} terms used to construct $\mathcal{A}_j$.[17] Again, the estimate of γ remains unchanged.

In column 4, we add predetermined destination controls (i.e., all variables in Table 2) as well as controls for current demographic characteristics prevailing in each village. Finally, column 5 is our most

---

[16]In addition to $\mathbf{x}_j$, we also control directly for the log of the great circle distance to the closest point in Java/Bali, log total land area, log distance to the subdistrict and district capital, and log distance to the nearest pre-1979 major road. None are material to the results.

[17]Controlling directly for the 119 π_{ij} terms leaves the results unchanged. We retain the parsimonious specification with province shares to preserve degrees of freedom as we introduce additional covariates in subsequent columns.

saturated regression that includes both origin and destination controls (87 in total). In both columns, the estimate of γ falls but is not statistically significantly different from column 1. Overall, these results rule out concerns that agroclimatically similar destinations were initially assigned or subsequently attracted settlers who differ along unobserved dimensions that are correlated with productivity.

To further strengthen our identification, we investigate the impact of agroclimatic similarity on cash crop productivity as a placebo exercise. As discussed in Section 3.2, agroclimatic similarity is not expected to be important for cash crops. Transmigrants were primarily food crop farmers with little prior experience growing cash crops, and farming skills are less location-specific for cash crops. To measure cash crop productivity, we construct a revenue-weighted average of log tons per hectare across crops. We follow Jayachandran (2006) and normalize the productivity of each crop to mean one for comparability. Revenue weights are based on national unit producer prices in 2001-2 from FAO/PriceSTAT. This approach provides a convenient way to aggregate the large number of cash crops grown across Transmigration villages.

Panel B of Table 3 shows that agroclimatic similarity has a small and insignificant effect on cash crop productivity. The point estimates and standard errors, which were estimated using the same specifications as for rice, are small relative to mean productivity of one ton per hectare and a standard deviation of 3.0. Moreover, we can reject the equality of coefficients in Panels A and B. The p-value for this test in our baseline (most demanding) specification in column 1 (5) is 0.001 (0.066).

The insignificant effects of agroclimatic similarity on cash crop productivity address concerns that the effects on rice productivity are driven by unobservables that influence general productivity. These omitted variables include market access, social capital, infrastructure, physical health (e.g., disease resistance) or other determinants of productivity that are common across cash crops and rice. We provide additional identification and robustness checks in Section 5.3.

Where are the Barriers to Transferability? We first show that the large average effects on rice productivity are concentrated among villages in the lower part of the similarity distribution. Following Robinson (1988), we estimate a semiparametric, partially linear version of equation (4),

$$y_j = \alpha + g(\mathcal{A}_j) + \mathbf{x}_j'\boldsymbol{\beta} + \omega_j,$$

where $g(\cdot)$ is a flexible function.

Figure 4 reveals nonlinear effects and a concave shape in which adjustments are increasingly costly the greater the agroclimatic distance to the origins.[18] The steepest effect size is found in the bottom tercile of the index ($\mathcal{A}_j \leq 0.55$) after which the effects of similarity kink and then level off. For these villages in the bottom tercile, a back-of-the-envelope calculation suggests that their low annual rice output produces calories close to subsistence levels. This is consistent with findings from Bryan et al. (2014) that subsistence farmers may underinvest in adaptation because losses from risky experimentation (with high expected return) are particularly costly near subsistence.

Figure 4 also clarifies how our natural experiment provides insights into the importance of sorting. In particular, the density for non-Transmigration villages in Figure 3 coincides with the flatter region

---

[18]We use a bandwidth of 0.05 here. We show in Appendix Figure B.2 that the concave shape is robust to different bandwidths, but the reversals at the tails are not.

in the semiparametric estimate in Figure 4. Without the program-induced skill mismatch, as in most migration settings, our study would lack the "empirical content" to say anything about the productivity implications of sorting based on comparative advantage (Heckman and Honoré, 1990). We would estimate smaller effects of agroclimatic similarity on productivity and mistakenly conclude that skills are transferable across locations precisely because spontaneous migrants tend to sort into destinations where their skills are easily transferable.

The semiparametric estimate provides important policy lessons. First, more careful matching of transmigrants' skills to destination growing conditions may have pushed more villages into the portion of the figure where agroclimatic similarity has relatively small effects. The concave shape suggests that avoiding very bad matches is more important than achieving the best match. Second, greater investments (targeted to low similarity villages) in agricultural extension, retraining programs, and complementary capital inputs may have facilitated greater adaptation and ultimately limited the persistent effects of initial dissimilarity seen in the lower tail of Figure 4. We revisit policy questions in Section 6.

Next, we show in Table 4 that agroclimatic similarity is more important in places with adverse growing conditions. In column 1, we interact agroclimatic similarity with the FAO measure of potential rice productivity.[19] The negative and significant coefficient on the interaction term implies agroclimatic similarity is less important in villages with high potential productivity. The magnitude suggests that, all else equal, planners could have mitigated the rice productivity losses of one standard deviation of dissimilarity by assigning badly matched transmigrants to villages with 1.3 tons/ha greater potential productivity.

Column 2 interacts agroclimatic similarity with indicators for three groups of Transmigration villages with low, medium, and high share of wetland as observed in 2002. This measure, which is uncorrelated with agroclimatic similarity (albeit not predetermined), captures additional variation in potential productivity as well as cultivation systems. The interaction term is positive for all villages but is largest and significant only for the bottom two terciles. One potential explanation is that farmers from Java/Bali were accustomed to wetland cultivation systems and found it difficult to adapt to the dryland environments in the Outer Islands, which require different farming methods (Donner, 1987).

Last, we identify which types of agroclimatic skills face greater barriers to transferability across growing conditions. We decompose our main index into subcomponents by mapping agroclimatic attributes to important tasks in farming. The key steps of production related to agroclimatic attributes include land preparation, water management and soil nutrient management (De Datta, 1981). Accordingly, we decompose our agroclimatic similarity index into three components comprising topographic, water, and soil similarity (see Appendix A.4 for details). Columns 3 to 5 in Table 4 repeat the specification in column 5 of Table 3 with each of the three subcomponents of similarity as the key regressors in place of the agroclimatic similarity index. Column 6 includes all three similarity subindices together.

The coefficient on soil similarity is largest and statistically significant, indicating that knowledge about soil conditions and soil management techniques are the most important.[20] Column 6 shows that a one standard deviation increase in soil dissimilarity causes an average productivity loss of 17 percent,

---

[19]To estimate potential yields, FAO GAEZ uses sophisticated agronomic models with predictions based on some of the topographic and climate data that we use (see Appendix A). We take a weighted average of potential dryland and wetland yields with weights based on the actual share of farmland that is wetland. We also control for potential productivity separately.

[20]Donner (1987, p. 195) highlights the importance of soil management techniques and barriers faced by transmigrants: "proper (soil management) techniques. . . are either unknown to the transmigrants or require too high an investment to be feasible."

implying barriers to the transferability of soil-specific skills. Soil management is complex and nuanced, involving multiple dimensions (soil pH, organic carbon content, sodicity) that determine not only the optimal varieties to employ but also the approach to diagnosing and remedying problems (De Datta, 1981). Moreover, food crops require a constant uptake of nutrients and fertilizer inputs, resulting in continuous changes in soil composition, which contribute to persistent barriers to adaptation.[21]

Skill Portability Across Crops.   To investigate whether agroclimatic skills are specific to rice or general to other crops, we show that agroclimatic similarity also matters for an important group of secondary food crops known as *palawija*. These annual crops discussed in Section 3.2 have similar production systems as rice and have lower entry barriers than cash crops. Because *palawija* crops were not universally grown in Java/Bali before the program, they allow us to estimate the transferability of general agroclimatic skills. We consider the log productivity of a given *palawija* crop using our baseline specification from Table 3. We also report a mean effects specification by creating a summary index for *palawija* crops, following Kling et al. (2007).[22] Figure 5 presents results from the mean effects analysis and also plots estimates for each *palawija* crop. The vertical bars represent the 90 percent confidence interval.

The mean effects estimate shows that a one standard deviation increase in agroclimatic similarity leads to a 7 percent increase in productivity for *palawija* crops (t-statistic ≈ 2). Agroclimatic similarity has a positive effect on productivity for four out of five *palawija* crops. Although estimates are not statistically significant for all crops, we reject the null hypothesis that agroclimatic similarity has zero effect on productivity for all crops (p-value of 0.026). The individual estimates are somewhat noisy because these crops are not as widely grown in destination villages as rice, and hence we lack power. This positive coefficient is consistent with the agronomic literature documenting similarities in the farming methods for rice and *palawija* crops. For example, *palawija* crops share many of the same soil nutrient disorders as rice, such as iron deficiency and salinity. In fact, when we repeat the subindex regressions discussed above for *palawija* crops, soil similarity is also the most important component.

Overall, these results parallel those in the labor literature on general versus specific human capital. The skills associated with agroclimatic similarity can be transferred across locations and applied to different crops for which location-specificity matters.

5.2   Adaptation and Development

So far, the large effects for rice and moderate effects for *palawija* suggest that location-specific skills are important for food crop productivity. These persistent barriers to skill transferability echo related work finding slow adaptation in response to abrupt changes in growing conditions (Hornbeck, 2012; Olmstead and Rhode, 2011). Given these barriers, how might households adapt? We investigate four potential

---

[21]Perry (1985, p. 108) describes "a situation of constant and rapid change in…nutrients…giv(ing) rise to difficulties in determining the correct fertiliser and rates of application to use".

[22]In calculating the mean effect estimate, we follow the supplementary appendix of Kling et al. (2007) and use a seemingly-unrelated regressions (SUR) system to estimate separate effects of similarity on individual *palawija* crops. We form $\tau = K^{-1} \sum_{k=1}^{K} \beta_k / \sigma_k$, where K is the number of *palawija* crops, β_k is the effect of similarity on the productivity of crop k, and σ_k is the standard deviation of crop k productivity in non-Transmigration villages. Standard errors are obtained from the variance-covariance matrix of the SUR system, while maintaing the spatial correlation structure as in our baseline regressions. Although villages differ in terms of which *palawija* crops are grown, this approach allows us to estimate a single mean effect in an unbalanced panel based on all villages and not simply those that grow all crops.

mechanisms: learning and social interactions, occupational switching, crop adjustment, and ex-post migration.

We find that language skills are important for social interactions with natives and for occupational adjustments. We also find evidence of crop adjustment in dissimilar villages. However, there remain sizable long-run differences between high and low similarity villages in nighttime light intensity, a proxy for local income. This suggests that adjustments were costly and perhaps incomplete.

Learning and Social Interactions. Farmers can adapt to new growing conditions by interacting with and learning from others. In Table 5, we investigate the role of social interactions among transmigrants and with native Outer Islanders.

To explore the effects of social interactions among transmigrants, we augment the specification in column 5 of Table 3 (reproduced here in column 1) with two salient and plausibly exogenous measures of diversity within Transmigration villages: the ethnic fractionalization (ELF) index among transmigrants and the Herfindahl index (HI) for origin district population shares. These measures are uncorrelated ($\rho = -0.03$). We include both linear (column 2) and quadratic (column 3) specifications of diversity.

The large effect of agroclimatic similarity remains unchanged in both specifications, but we uncover some interesting nonlinear effects of diversity. The quadratic specification in column 3 implies that greater origin district concentration is positively correlated with rice productivity until around the 70th percentile of HI ≈ 0.2.[23] The inverted-U shape suggests that, up to a point, having fellow transmigrants from the same origin is useful, perhaps due to network and agglomeration effects. However, beyond that point, having too few individuals outside one's own origin group may limit possible productivity-enhancing interactions (e.g., intergroup learning or insurance), rendering concentration a drag on overall productivity.[24]

Next, we provide evidence consistent with social learning from natives. To do so, we use the *Ethnologue* data on language structure and the *World Language Mapping System* (WLMS) data on linguistic homelands to construct a measure of the distance between each of the eight ethnolinguistic groups ℓ indigenous to Java/Bali and each of the nearly 700 ethnolinguistic groups prevailing across the Outer Islands (see Appendix A for details). Our *linguistic similarity* for village j is given by:

$$linguistic\ similarity_j \equiv \mathcal{L}_j = \sum_{\ell=1}^{8} \pi_{\ell j} \left(\frac{branch_{\ell j}}{max\ branch} \right)^{\psi}, \tag{5}$$

where $\pi_{\ell j}$ is the share of immigrants in village j from ethnolinguistic group ℓ in Java/Bali, $branch_{j\ell}$ is the sum of shared language tree branches between ℓ and the language indigenous to village j, $max\ branch = 7$ is the maximum number of shared branches between any Java/Bali language and any native Outer Island language, and ψ is a parameter, set to 0.5 as a baseline following Fearon (2003). Importantly, linguistic similarity is uncorrelated with agroclimatic similarity ($\rho = -0.04$), which is consistent with

---

[23]The turning point is significant at the 6 percent level based on the exact test of Lind and Mehlum (2010).

[24]The concave shape in Figure 4 is robust to controlling quadratically for origin concentration and within-Java/Bali ethnic fractionalization as in column 3. It is possible that villages with extreme dissimilarity may have greater potential for risk diversification through connections to villages with different risk profiles at home in Java/Bali. However, this greater opportunity for risk sharing would bias us against our finding that dissimilarity causes large losses in rice productivity and income.

the exogenous assignment process.

Column 4 of Table 5 shows that a one standard deviation increase in linguistic similarity increases rice productivity by 25.8 percent. In column 5, we find that linguistic similarity is more important in villages with a smaller initial stock of transmigrants and hence more natives. It is in these villages where we would expect a greater scope for interacting with and learning from natives.[25] These results corroborate case studies of Transmigration settlements that discuss the importance of learning from natives (e.g., Donner, 1987). They are also consistent with studies of local learning under heterogeneous growing conditions. For example, Munshi (2004) documents stronger evidence of learning from neighbors in the case of wheat relative to rice because rice varieties are more sensitive to local growing conditions. Similarly, BenYishay and Mobarak (2014) find that farmers are most persuaded by information provided by other farmers who face comparable agricultural conditions.

As with others using these measures of diversity in the literature (e.g., Desmet et al., 2009; Esteban et al., 2012), we view them as reflecting not only ease of communication but also cultural proximity, shared preferences, and the general fluidity of potential interactions between groups. For example, these findings are also consistent with other productivity-enhancing arrangements besides social learning, such as the provision of credit and risk sharing.

Overall, our results highlight the importance of social interactions. However, the robustness of the agroclimatic similarity coefficient to the inclusion of the additional controls in Table 5 suggests that social capital was not a strong enough adaptation mechanism to dampen the effects of agroclimatic similarity.

Occupational Choice.   We next examine the possibility of switching occupations as a form of adaptation. Consider a simple Roy model with two skills, agricultural and language, and two occupations, farming and trading/services. Farming is relatively more intensive in agricultural skills while trading/services is relatively more intensive in language skills. The theory of comparative advantage predicts that individuals assigned to agroclimatically similar villages are more likely to remain as farmers, and those assigned to linguistically similar villages are more likely to switch into trading/services.

We test these predictions in Table 6 using the universe of 2000 Population Census data. We estimate a linear probability model of occupational choice as a function of individual demographic controls, village controls, year-of-settlement fixed effects, and individual agroclimatic and linguistic similarity, which is the term after $\pi_{\ell j}$ in equation (5). Columns 1-3 report estimates for the probability of being a farmer working in either food or cash crop production, while columns 4-6 report the probability of being involved in trading or services. The sample in columns 1 and 4 includes the Java/Bali-born population between the working ages of 15 to 65. Columns 2 and 5 (3 and 6) restrict the sample to individuals who were less (older) than 10 years old in the year of initial settlement.

We find some adjustment in occupation choices, consistent with the theory of comparative advantage. A one standard deviation increase in individual agroclimatic similarity leads to a 0.9 p.p. higher probability of an individual reporting farming as their primary occupation, which is small relative to

---

[25]Although we do not observe the initial native population size, the size of the initial transmigrant population is a good proxy for relative group sizes. Given that program planners accounted for the surrounding native population size when they calculated the carrying capacity, conditional on agroclimatic endowments $\mathbf{x}_j$, a large (small) initial transmigrant population is indicative of a small (large) initial native population.

the sample mean of 62 percent. Meanwhile, a one standard deviation increase in linguistic similarity is associated with 1.8 p.p. higher probability of trading/services, which is large relative to the sample mean of 9.9 percent. Column 5 shows that younger transmigrants in agroclimatically dissimilar places are more likely to become traders, but their response is not statistically different from the older cohorts.

It is important to note that these occupational choices are robust to accounting for the local suitability of the land for agriculture and for rice in particular. The coefficients remain unchanged when controlling for potential rice productivity, which is positively (negatively) correlated with the probability of farming (trading). Moreover, high potential rice yields weaken the strength of occupational sorting, which suggests that fertile land eases the adaptation process (see Appendix Table B.3). This is consistent with the fact that agroclimatic similarity has smaller productivity effects in villages with high potential rice productivity (see Table 4).

Crop Adjustment. Although many low similarity transmigrants remained farmers, crop switching may have been another potentially important margin of adjustment. Table 7 presents evidence of dissimilar villages switching to cash crops. In column 1, a one standard deviation decrease in agroclimatic similarity leads to a 4.3 p.p. increase in the revenue share of cash crops (based on the measure described on p. 14) relative to a mean of 57 percent. Column 2 shows that a one standard deviation increase in agroclimatic similarity leads to a 4.7 p.p. increase in the share of rice.

Like the occupational sorting between farming and trading/services, these results in Table 7 are also consistent with a comparative advantage interpretation. That is, farmers in high similarity villages who have high location-specific human capital allocate more resources towards crops where location-specific human capital is relatively more important (rice). Meanwhile, column 3 indicates that agroclimatic similarity has a small and insignificant effect on the share of farmers whose primary occupation is growing cash crops (according to the 2000 Census).[26] Overall, the patterns in columns 1-3 suggest that switching to cash crops is a potentially important albeit limited margin of adjustment.

Farmers face significant barriers to cash crop adoption. One important barrier is the opportunity cost of land use as farmers can either plant food or cash crops. This cost is significant because farmers have to wait multiple years before the trees mature and bear fruit. Since cash crop productivity is the same across high and low similarity villages (see Table 3), low similarity villages that have low rice productivity have lower opportunity costs and are more likely to switch, consistent with the effect in column 1. Moreover, the adoption of cash crops requires larger initial investments than food crops. Given that villages with low agroclimatic similarity are just above subsistence (see p. 14), the upfront costs of experimentation with cash crops, and their delayed returns, may be prohibitive.

In an attempt to summarize across crops, we show in column 4 that agroclimatic similarity has a small and statistically insignificant effect on revenue-weighted average agricultural productivity across all crops. This is not surprising given that cash crops have a substantially higher potential revenue weight than rice, and agroclimatic similarity has no effect on cash crop productivity. A simple decomposition exercise suggests that the null productivity effect of agroclimatic similarity on cash crops (see Panel B of Table 3) with a high revenue weight of 0.57 offsets the large productivity effect on rice with a lower

---

[26]This null result also holds at the individual-level in regressions similar to those in Table 6. However, it is important to note that this measure of primary occupational choice does not fully capture time allocation and hence labor inputs to cash versus food crop production.

revenue weight of 0.27 and can explain the null result in column 4.[27]

However, there are several reasons why the unobservable ideal weights could be higher for rice and smaller for cash crops. First, 65 percent of farmers grow food crops, which are relatively more important for poor households as a source of livelihood and basic caloric value. Second, there are large differences in fixed and variable input costs of production across rice and cash crops, and there are different time horizons for growing and harvesting different crops. In turn, these likely imply smaller differences between cash and food crops in annual profits, which could be more ideal in capturing welfare.

In Appendix B.1, we provide additional insight into the crop adjustment mechanism. Similar to findings in Michalopoulos (2012), we show that transmigrants' origin region cropping patterns still partially explain destination cropping patterns. Our regression accounts for the predetermined relative suitability for rice as proxied by cropping patterns in nearby, non-Transmigration villages. The effect of origin cropping patterns is only 20 percent as large as neighboring cropping patterns, suggesting some crop adjustments by individual farmers.

(Non-)Selective Migration Patterns. Another way in which farmers may adapt to initial low quality matches is by moving out of their assigned village and perhaps returning to Java/Bali. Several factors explain why this margin of adjustment is less important in our context. First, transmigrants are not as mobile as typical migrants; they volunteered for the program because they were unable to migrate on their own, due to credit, information, or other constraints. These same constraints were arguably just as binding in their new settlements, particularly for the relatively less successful farmers.[28] Second, these migrants were mostly landless farm households who were given land, which may have played a role in tying them to the Transmigration villages. Property rights were only distributed after 5-10 years, and evidence from Mexico suggests that landholdings without certification tend to reduce outmigration (de Janvry et al., 2015). Even with property rights, imperfections in Indonesia's land markets (World Bank, 2008) may have prevented large-scale resales. Finally, aggregate statistics from a 1984 Income Survey of Transmigrants show that 71 percent (11 percent) report higher (equal) income compared after migrating, and even those villages with low similarity achieved productivity levels above subsistence, all of which could also explain why we did not see large return migration flows in the early years.

We confirm that selective out-migration is indeed low. First and foremost, as detailed in Section 5.3, a quasi-gravity regression shows that longer-term, ex-post sorting patterns are uncorrelated with agroclimatic similarity. Also, agroclimatic similarity is uncorrelated with population size and the

---

[27]The decomposition follows from the product rule. Total productivity is calculated as $\omega_R \ln y_R + \omega_C \ln y_C + \mathcal{O}$ where ω_R and $\ln y_R$ are the revenue weights and log productivity for rice, ω_C and $\ln y_C$ are the analogues for cash crops, and $\mathcal{O}$ is the weighted average for all other (food) crops. The effect of a one standard deviation increase in agroclimatic similarity ($\mathcal{A}$) on total productivity is then the sum of the effects for each crop:

$$\frac{\mathrm{d}\omega_R}{\mathrm{d}\mathcal{A}} \ln y_R + \omega_R \frac{\mathrm{d}\ln y_R}{\mathrm{d}\mathcal{A}} + \frac{\mathrm{d}\omega_C}{\mathrm{d}\mathcal{A}} \ln y_C + \omega_C \frac{\mathrm{d}\ln y_C}{\mathrm{d}\mathcal{A}} + \frac{\mathrm{d}\mathcal{O}}{\mathrm{d}\mathcal{A}}.$$

We use estimated effects of agroclimatic similarity on revenue weights and productivity and evaluate this equation using revenue weights and log productivity for the average village. The key is that the revenue weights (calculated using national prices) are low for rice and high for cash crops, so that the large productivity effect for rice (0.2) is weighted down to 0.05 (0.27×0.2) and the small productivity effect for cash crops (0.024) is now relatively higher at 0.014 (0.57×0.024).

[28]Even if poorly matched transmigrants could identify higher similarity locations and had the resources to move, it would still have been extremely difficult to do so given that the farmland in those locations would either have already been occupied or required substantial upfront costs to clear that would have likely been beyond the limited means of the typical transmigrant.

Java/Bali-born migrant share in Transmigration villages in 2000. Finally, the 1998 Transmigration census reports the number of individuals initially placed as well as the population size when a Transmigration village was deemed independent enough that it no longer required official supervision (typically within 5-10 years of placement). We regress the log ratio of these two population sizes on agroclimatic similarity and find small, insignificant effects. If there was selective outmigration from dissimilar villages, these coefficients would be positive and significant.

Light Intensity as a Proxy for Income. Having identified strong effects of agroclimatic similarity on rice productivity, it is important to ask whether the adaptation responses discussed above can undo the adverse effects of dissimilarity on economic development over time. We investigate whether skill trans-ferability has persistent effects on local income using the best available proxy: nighttime light intensity. By 2010, nearly 25 percent of Transmigration villages recorded some nighttime lights.[29]

Table 8 reports positive and statistically significant effects of agroclimatic similarity on light intensity. Although transmigrant farmers adapted in several ways to low quality matches, the results in columns 1–4 imply that such adaptation was costly and may be incomplete. Using the baseline specification from Table 3, column 1 shows that a one standard deviation increase in agroclimatic similarity leads to a 1.6 p.p. increase in the share of village area that has any nighttime light coverage relative to a mean of 8.1 percent. This estimate increases substantially when including the full set of additional controls in column 2. This is perhaps because agroclimatic similarity is negatively correlated with district-level manufacturing intensity and electrification before the Transmigration villages were established (see Table 2). Not controlling for these variables, which are mechanically positively correlated with luminosity, biases us against finding positive effects on light intensity. We fix initial conditions by including year-of-settlement fixed effects in all columns, and since Transmigration villages had no inhabitants or lights prior to the program, this allows us to assign a long-run growth interpretation to the estimates.

Beyond the extensive margin, higher agroclimatic similarity also leads to growth in the intensive margin of light intensity. Columns 3 and 4 demonstrate this using a Poisson estimator with the level of light intensity $\in [0, 57.6]$ as the dependent variable. This strategy is commonly used in the estimation of gravity models for trade flows and is known to perform well despite the relatively large proportion of zeros (Santos Silva and Tenreyro, 2011).[30] Column 3 suggests that a one standard deviation increase in similarity translates to an increase in light intensity of almost 21 percent. Again, the effect increases substantially to 39 percent when adding the full set of controls in column 4. The average marginal effects (AMEs) are only slightly smaller, taking the point estimates $\times$ the mean intensity of 0.75. Using an approach similar to Henderson et al. (2012), Olivia and Gibson (2015) estimate that a one percent increase in light intensity is associated with a 0.4 percent increase in district-level gross GDP. Assuming that this elasticity holds at lower geographic levels, our estimated AMEs imply that a one standard deviation

---

[29]By comparison, 12 percent of the 12.5 km^2 grids making up sub-Saharan Africa had recorded lights in 2007 and 2008 (Michalopoulos and Papaioannou, 2014). The average Transmigration village spans 43 km^2 with 8.1 percent of that area recording any lights.

[30]Results are qualitatively similar using alternative approaches to dealing with zeros such as a square root transformation or using $\ln(c + light\ intensity)$ for some small constant c as in Hodler and Raschky (2014) and Michalopoulos and Papaioannou (2014), but this latter approach is subject to potential biases due to heteroskedasticity (see Santos Silva and Tenreyro, 2006). Moreover, marginal effects are more readily interpretable using the Poisson approach than the square root transformation (Cameron and Trivedi, 2013, pp. 103–4).

increase in agroclimatic similarity increases long-run village-level income by 6.3 to 11.7 percent. These economically significant effects suggest that agroclimatic similarity had persistent impacts on a broad measure of local development.

5.3 Robustness Checks and Other Threats to Identification

Tests for Sorting. Using a quasi-gravity specification, we provide direct evidence in Table 9 that transmigrants did not endogenously sort across Transmigration sites on the basis of agroclimatic similarity. In particular, we examine whether the stock of Java/Bali migrants from origin district i residing in Transmigration village j in 2000 is increasing in agroclimatic similarity ($\mathcal{A}_{ij}$) between i and j:

$$f(migrants_{ij}) = \alpha + \lambda_a \mathcal{A}_{ij} - \lambda_d \ln distance_{ij} + \mathbf{z}_j' \boldsymbol{\zeta} + \tau_i + \upsilon_{ij}, \tag{6}$$

where τ_i are origin fixed effects, and $\mathbf{z}_j$ includes island fixed effects, the year of initial settlement, and the log number of individuals placed in j. Columns 2 and 4 also include all of the predetermined variables in Table 2. We estimate equations for both the extensive margin, $f(migrants_{ij}) := Pr(migrants_{ij} > 0)$, and the intensive margin, $f(migrants_{ij}) := \ln(migrants_{ij})$, of migration flows. In all cases, we two-way cluster standard errors by i and j (Cameron et al., 2011). In Table 9, we parameterize the destination fixed effects using the $\mathbf{x}_j$ as in our village-level regressions in Table 3, but results are unchanged when including village fixed effects instead.

In all specifications, we cannot reject the null hypothesis that $\lambda_a = 0$. Moreover, the estimated λ_a are very small relative to the mean of the given dependent variable. This mitigates concerns that migrants from Java/Bali endogenously sorted into (out of) more (dis)similar sites 12 to 20 years after the initial wave of resettlement. Migrant stocks tend to be somewhat higher in physically closer sites, perhaps due to transport costs (see Section 2.1), which we accounted for by controlling for (origin-weighted) distance. However, agroclimatic "distance" does not exhibit the same hypothesized gravity forces along either the extensive or intensive margin.

The small estimates of λ_a are consistent with a simulation exercise showing that the actual distribution of agroclimatic similarity across Transmigration villages is comparable to the distribution that would have resulted from a purely random assignment (see Appendix B.3). Overall, these results help rule out concerns that farmers are sorting based on unobservable sources of comparative advantage that are spuriously correlated with similarity.

Further Robustness Checks. In Appendix Table B.5, we demonstrate the robustness of our key rice productivity results to (i) confounding program features, (ii) additional controls, and (iii) alternative specifications of the agroclimatic similarity index.

Next, we rule out endogeneity concerns associated with selection into growing rice. First, we show in Appendix Figure B.3 that individual agroclimatic similarity is balanced across education levels. This mitigates concerns that similarity is correlated with ability. Second, we show in Appendix B.2, that the degree of selection needed to explain the productivity effects is implausibly large when compared to the estimated effects of agroclimatic similarity on occupational choices as detailed above. Additionally, we rule out endogeneity concerns associated with the fact that not all villages produce rice (see Ap-

pendix B.3). Following Altonji et al. (2005), we calculate that selection on unobservables would have to be at least 10 times greater than selection on observables to explain the 16.6 percent effect on productivity in column 5 of Table 3. See Appendix B.3 for more details.

Finally, we address possible aggregation bias due to the fact that we are linking village-level productivity to average individual-level similarity.[31] We find similar results if we control for the share of the population from Java/Bali as well as overall log population density. Separately, using *Susenas* data for a small sample of Transmigration villages that includes household-level rice productivity, we find a similar skill transfer elasticity (see Appendix Table B.7).

6 Impact of the Transmigration Program: Policy Exercises

In this section, we demonstrate the aggregate implications of origin-by-destination match quality using two counterfactual policy exercises. First, we use simulations to show that a reallocation of transmigrants to maximize agroclimatic similarity could have led to large increases in total rice output. Second, we use a policy discontinuity and a place-based evaluation strategy to provide the first causal estimates of the average impact of the Transmigration program on local economic development. Ultimately, we argue that the persistent effects of agroclimatic similarity may explain the small average impact of the program on local development. Despite the growing policy relevance of resettlement, there remains a dearth of causal evidence on the medium- to long-run impacts of resettlement programs, especially in developing countries. The findings below fill that gap.

6.1 Optimal Reallocation of Migrants

We attempt here to quantify the aggregate output losses from poorly matching transmigrants' farming skills to local growing conditions. We use the baseline rice productivity results in column 1 of Table 3 and reassign transmigrants to destinations to maximize agroclimatic similarity across all villages, and hence, total rice output. As discussed in Appendix C, this objective is a special case of the generalized assignment problem, a problem in combinatorial optimization that has been shown to be NP-hard in terms of its complexity (Fischer et al., 1986). However, we can approximate the optimal solution using a greedy assignment algorithm, in which similarity is sequentially maximized, village-by-village, subject to a capacity constraint proxied by the number of individuals placed in the initial year.

Using this algorithm, we find that aggregate rice production could have been 27 percent higher if individuals had been assigned in a more optimal manner. While this may not be a global optimum, the solution is computationally feasible and represents an approach to the problem that could be carried out by future resettlement planners. Indeed, this type of agroclimatic assignment mechanism would address an important challenge recognized in the World Bank's Operational Policy on resettlement, namely that individuals are often relocated to places where their skills are not relevant (World Bank, 2001). Our

---

[31]One concern would be if all transmigrants with low (high) agroclimatic similarity had high (low) rice productivity, and the particular mix of each type spuriously led to our estimated effects of average agroclimatic similarity at the village level. This sort of bias is unlikely to arise in practice given that the program fixed land sizes across households and allocated plots by lottery. As a result, within-village land quantity and quality should be orthogonal to individual agroclimatic similarity, and hence average agroclimatic similarity should indeed be informative about aggregate productivity without confounding due to individual-level heterogeneity.

findings provide evidence on the importance of accounting for the costs of skill mismatch in determining the potential returns to resettlement.

6.2 Place-Based Program Evaluation

As noted in Section 2.1, global oil prices collapsed in the mid-1980s, and declining government revenues forced dramatic cutbacks in the Transmigration budget, leading to a significant reduction in the number of sponsored households over the coming years.[32] As a result, numerous selected sites never received any transmigrants. We use this set of planned-but-unsettled (or "almost-treated") villages as counterfactual settlements to estimate the average treatment effects (ATE) of the program on local economic development.

We identify almost-treated villages using the MOT's maps of recommended development areas (RDAs) constructed after large-scale mapping of agricultural viability during the site-selection process. We digitally trace these RDAs and define as almost-treated those 907 non-Transmigration villages that share any area with the RDAs (see Appendix Figure B.4). Our conclusions are similar using other approaches.

We use these almost-treated villages as a control group in the following estimating equation:

$$y_j = \alpha + \theta T_j + \mathbf{x}_j'\boldsymbol{\beta} + \nu_j, \tag{7}$$

where T_j is a treatment indicator equal to one for Transmigration villages and zero for almost-treated villages, and $\mathbf{x}_j$ is the usual vector of predetermined controls from equation (4). The key parameter of interest is the ATE, θ, which measures the impact of being a Transmigration village.

A key concern with assigning θ a causal interpretation is that there are omitted place variables correlated with treatment assignment that influenced both site selection and outcomes. Spatial policies like the Transmigration program often target underdeveloped or distressed areas, which can lead to downward bias in $\hat{\theta}$. We rule out first-order concerns with program placement bias by restricting the sample to treated and almost-treated villages and using a reweighting procedure akin to recent evaluations of place-based policies (Busso et al., 2013; Kline and Moretti, 2014).

In Table 10, column 1 compares Transmigration villages to all other Outer Island villages while columns 2-4 restrict the sample to the set of treated and control villages. Column 2 controls for the predetermined site selection (and agroclimatic) characteristics in $\mathbf{x}_j$. Column 3 implements a double robust approach that additionally reweights control villages according to their estimated odds of treatment. Column 4, our preferred specification, uses the Oaxaca-Blinder reweighting approach of Kline (2011). Comparisons with column 1 show how endogenous program placement biases $\hat{\theta}$ downward.[33]

Table 10 demonstrates that the Transmigration program had mixed effects on local economic development. First, treated villages have substantially higher population density (0.77 log points) than almost-treated villages. We also find a statistically and economically significant effect ($\hat{\theta} \approx 0.47$) on

---

[32]The budget fell from Rp 578 billion in FY 1985-86 to Rp 325 billion in FY 1986-87. In response, the MOT reduced its FY 86/87 targets for settlement on sites already under preparation from 100,000 to 36,000 sponsored households.

[33]All specifications include island fixed effects and cluster standard errors at the district level. Sample sizes vary across outcomes and columns (depending on data availability) but include as many as 31,185 villages in column 1, and 832 treated villages and 668 controls in columns 2-4. We exclude control villages that are within 10 km of Transmigration settlements to minimize bias from spillovers. See Appendix B.4 for further details.

the log of price-weighted total agricultural output. However, we find no significant effects on average agricultural productivity or rice productivity (tons/ha). This is not due to differential selection into rice production. Finally, the program did not have a significant impact on light intensity.

These mixed findings can be explained in light of the persistent consequences of spatial mismatch and incomplete adaptation in Transmigration villages. If the barriers to adaptation in these villages are strong enough and are not binding in control villages, then the Transmigration-funded inputs to production could have been undone after two decades. In other words, the low agroclimatic similarity and incomplete adaptation could have pulled down average productivity in treated villages relative to control villages where the residents have significantly higher agroclimatic similarity. The limited effects on productivity and light intensity in Table 10 are consistent with this interpretation, while the large effects on population density and total agricultural output point to the direct impacts of land extensification due to the program.

7 Conclusion

This paper uses plausibly exogenous variation from a large-scale rural-to-rural resettlement program in Indonesia to identify the causal impact of skill transferability on agricultural productivity. We show that villages that were assigned a higher share of migrants from agroclimatically similar origins in Java/Bali exhibit greater rice productivity compared to villages that were assigned migrants from less similar origins.

Our findings shed new light on the importance of comparative advantage in shaping the spatial distribution of productivity. Our natural experiment suggests that some of the observed spatial productivity gaps may be explained by barriers to transferring skills and ultimately adjusting to new economic environments. The large effects we find for rice and the moderate effects for other food crops are important because food crops are the source of livelihood for many poor farmers. If a large component of their skills is location-specific, then, it is difficult to arbitrage productivity differences across space by migrating. Although transmigrants may have adapted through social interactions or crop adjustments, our findings on nighttime lights indicate costly and perhaps incomplete adjustment over the medium-to long-run period in this study. This suggests regional gaps in agricultural productivity could be persistent. Quantifying the welfare costs of these barriers is an important task for future work.

Our results also have important implications for the design of future resettlement programs. We provide evidence from a simulation exercise suggesting sizable aggregate rice productivity gains from optimally allocating migrants on the basis of agroclimatic similarity. We find the largest barriers to transferability for soil-specific skills and for villages in the bottom tercile of agroclimatic similarity. These highlight where the barriers to adaptation are most significant and point to where government inputs and extension services should be targeted. We also find that both social capital within resettlement areas and interacting with natives are important adaptation mechanisms that should be considered when resettling people to new growing environments.

Our study suggests several directions for future research. While we establish the importance of agroclimatic similarity for agricultural productivity and output, the effects on resilience to weather shocks, risk sharing, and prices, are also relevant. Moreover, an open policy question concerns the causal impact

of program participation on household welfare. Separately, another important objective of the program was nation building, which our natural experiment is well-suited to investigate. Did social cohesion improve between Inner and Outer Islanders, and how is it affected by language and the degree of ethnic diversity? Finally, the empirical approach we develop in this paper can help inform an ongoing debate in American history concerning the role of African-born slaves' location-specific (rice) farming skills in shaping agricultural development in the Americas (see Carney, 2001; Eltis et al., 2010).

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson, "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 2014, *122* (3), 467–717.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Arnall, Alex, David S. G. Thomas, Chasca Twyman, and Diana Liverman, "Flooding, resettlement, and change in livelihoods: evidence from rural Mozambique," *Disasters*, 2013, *37* (3), 468–488.

Ashraf, Quamrul and Oded Galor, "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, 2013, *103* (1), 1–46.

Atkin, David, "The Caloric Costs of Culture: Evidence from Indian Migrants," *American Economic Review*, forthcoming.

Au, Chun-Chung and J. Vernon Henderson, "Are Chinese Cities Too Small?," *The Review of Economic Studies*, 2006, *73*, 549–576.

Autor, David, "The "Task Approach" to Labor Markets: An Overview," *Journal of Labour Market Research*, 2013, *46* (3), 185–199.

Bauer, Thomas K., Sebastian Braun, and Michael Kvasnicka, "The Economic Integration of Forced Migrants: Evidence for Post-War Germany," *The Economic Journal*, 2013, *123* (571), 998–1024.

Bayer, Patrick, Shakeeb Khan, and Christopher Timmins, "Nonparametric Identification and Estimation in a Roy Model with Common Nonpecuniary Returns," *Journal of Business & Economic Statistics*, 2011, *29*, 201–215.

Bazzi, Samuel and Christopher Blattman, "Economic Shocks and Conflict: Evidence from Commodity Prices," *American Economic Journal: Macroeconomics*, 2014, *6*, 1–38.

Beaman, Lori A., "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the US," *The Review of Economic Studies*, 2012, *79*, 128–161.

Beddow, Jason M., Philip G. Pardey, Jawoo Koo, and Stanley Wood, "The Changing Landscape of Global Agriculture," in Julian M. Alston, Bruce A. Babcock, and Philip G. Pardey, eds., *The Shifting Patterns of Agricultural Production and Productivity Worldwide*, Ames, IA: Midwest Agribusiness Trade Research and Information Center, 2010, chapter 2.

BenYishay, Ariel and A. Mushfiq Mobarak, "Social Learning and Communication," *NBER Working Paper w20139*, 2014.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed M. Mobarak, "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 2013, *103*, 897–947.

Cameron, A. Colin and Pravin K. Trivedi, *Regression Analysis of Count Data*, Vol. 53, Cambridge, UK: Cambridge University Press, 2013.

— , **Jonah B. Gelbach, and Douglas L. Miller**, "Robust Inference with Multiway Clustering," *Journal of Business & Economic Statistics*, 2011, *29* (2).

Carney, Judith A., *Black Rice: The African origins of rice cultivation in the Americas*, Cambridge, MA: Harvard University Press, 2001.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," *Handbook of Economic Growth*, 2005, *1*, 679–741.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, 2008, *63* (2), 723–742.

— , — , **and** — , "The Identification of Agglomeration Economies," *Journal of Economic Geography*, 2011, *11*, 253–266.

Comin, Diego A., Mikhail Dmitriev, and Esteban Rossi-Hansberg, "The Spatial Diffusion of Technology," *NBER Working Paper No. w18534*, 2012.

Conley, Timothy G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92*, 1–45.

— **and Christopher R. Udry**, "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*,

2010, *100* (1), 35–69.

Costinot, Arnaud, Dave Donaldson, and Cory Smith, "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields Around the World," *Journal of Political Economy*, forthcoming.

Cox, Thomas S., Jerry D. Glover, David L. Van Tassel, Cindy M. Cox, and Lee R. DeHaan, "Prospects for Developing Perennial Grain Crops," *BioScience*, 2006, *56* (8).

Dahl, Gordon B., "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, 2002, *70*, 2367–2420.

Datta, Surajit K. De, *Principles and Practices of Rice Production*, Hoboken, NJ: John Wiley & Sons, 1981.

de Janvry, Alain, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet, "Delinking Land Rights from Land Use: Certification and Migration in Mexico," *American Economic Review*, 2015, *105* (10).

de Sherbinin, A., M. Castro, F. Gemenne, M.M. Cernea, S. Adamo, P.M. Fearnside, G. Krieger, S. Lahmani, A. Oliver-Smith, A. Pankhurst, T. Scudder, B. Singer, Y. Tan, G. Wannier, P. Boncour, C. Ehrhart, G. Hugo, B. Pandey, and G. Shi, "Preparing for Resettlement Associated with Climate Change," *Science*, 2011, *344*, 456–457.

de Wet, Chris, "The Experience with Dams and Resettlement in Africa," 2000. Consultancy report written for the World Commission on Dams.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Shlomo Weber, "Linguistic Diversity and Redistribution," *Journal of the European Economic Association*, 2009, *7* (6), 1291–1318.

Diamond, Jared M., *Guns, Germs, and Steel*, New York, NY: W. W. Norton & Company, 1997.

Donner, Wolf, *Land Use and Environment in Indonesia*, London: UK: C. Hurst and Company, 1987.

Dove, Michael R., "The Agroecological Mythology of the Javanese and the Political Economy of Indonesia," *Indonesia*, 1985, *39*, 1–36.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic Enclaves and the Economic Success of Immigrants–Evidence from a Natural Experiment," *Quarterly Journal of Economics*, 2003, *118* (1), 329–357.

Eltis, D., P. Morgan, and D. Richardson, "Black, Brown, or White? Color-Coding American Commercial Rice Cultivation with Slave Labor," *The American Historical Review*, 2010, *115* (1), 164–171.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 2012, *102* (4), 1210–1342.

Fearon, James D., "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 2003, *8* (2), 195–222.

Fischer, Marshall L., R. Jaikumar, and Luk N. van Wassenhove, "A Multiplier Adjustment Method for the Generalized Assignment Problem," *Journal of Management Science*, 1986, *32* (9), 1095–1103.

Food and Agriculture Organization of the United Nations (FAO), "The oil palm," in "FAO Economic and Social Development Series: Better Farming Series" 1990, chapter 24.

— , "Africa's Changing Landscape: Securing Land Access for the Rural Poor," 2010. Accessed online on 15 Oct 2015, available at http://www.cpahq.org/cpahq/cpadocs/FAO

Foster, Andrew D. and Mark R. Rosenzweig, "Microeconomics of Technology Adoption," *Annual Review of Economics*, 2010, *2* (1), 395–424.

Friedberg, Rachel M., "You Can't Take It With You? Immigrant Assimilation and the Portability of Human Capital," *Journal of Labor Economics*, 2000, *18*, 221–251.

Gathmann, Christina and Uta Schönberg, "How General Is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 2010, *28* (1), 1–49.

Geertz, Clifford, *Agricultural Involution: The Processes of Ecological Change in Indonesia*, Vol. 11, Berkeley, CA: University of California Press, 1963.

Giovanni, Julian, Andrei A. Levchenko, and Francesc Ortega, "A Global View of Cross-Border Migration," *Journal of the European Economic Association*, 2015, *13* (1), 168–202.

Glitz, Albrecht, "The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany," *Journal of Labor Economics*, 2012, *30* (1), 175–213.

Glover, J. D., J. P. Reganold, L. W. Bell, J. Borevitz, E. C. Brummer, E. S. Buckler, C. M. Cox, T. S. Cox, T. E. Crews, S. W. Culman, L. R. DeHaan, D. Eriksson, B. S. Gill, J. Holland, F. Hu, B. S. Hulke, A. M. H. Ibrahim, W. Jackson, S. S. Jones, S. C. Murray, A. H. Paterson, E. Ploschuk, E. J. Sacks, S. Snapp, D. Tao, D. L. Van Tassel, L. J. Wade, D. L. Wyse, and Y. Xu, "Increased Food and Ecosystem Security via Perennial Grains," *Science*, 2010, *328*.

Gollin, Douglas, David Lagakos, and Michael E. Waugh, "The Agricultural Productivity Gap in Developing Countries," *Quarterly Journal of Economics*, 2014, *129* (2), 939–993.

Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 1957, *25* (4), 501–522.

Hall, Anthony, "Making People Matter: Development and the Environment in Brazilian Amazonia," *International Journal of Contemporary Sociology*, 1993, *30* (1), 63–80.

Hardjono, Joan, "The Indonesian Transmigration Program in Historical Perspective," *International Migration*, 1988, *26* (4), 427–439.

Heckman, James J and Bo E. Honoré, "The Empirical Content of the Roy Model," *Econometrica*, 1990, *58* (5), 1121–1149.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, *102* (2), 994–1028.

Hodler, Roland and Paul A. Raschky, "Regional Favoritism," *Quarterly Journal of Economics*, 2014, *129* (2), 995–1033.

Hornbeck, Richard, "The Enduring Impact of the American Dust Bowl: Short-and Long-Run Adjustments to Environmental Catastrophe," *American Economic Review*, 2012, *102* (4), 1477–1507.

Huffman, Wallace E. and Tubagus Feridhanusetyawan, "Migration, Fixed Costs, and Location-Specific Amenities: A Hazard Analysis for a Panel of Males," *American Journal of Agricultural Economics*, 2007, *89* (2), 368–382.

Intergovernmental Panel on Climate Change (IPCC), "Summary for policymakers," in C.B. Field, V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White, eds., *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, UK; New York, NY, USA: Cambridge University Press, 2014, pp. 1–32.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), 538–575.

Kebschull, Dietrich, *Transmigration in Indonesia: An Empirical Analysis of Motivation, Expectations and Experiences*, Hamburg, Germany: Transaction Publishers, 1986.

Kinsey, Bill H. and Hans P. Binswanger, "Characteristics and Performance of Resettlement Programs: A Review," *World Development*, 1993, *21*, 1477–1494.

Klein, Paul and Gustavo Ventura, "Productivity differences and the dynamic effects of labor movements," *Journal of Monetary Economics*, 2009, *56* (8), 1059–1073.

Kline, Patrick, "Oaxaca-Blinder as a Reweighting Estimator," *American Economic Review: Papers & Proceedings*, 2011, *101* (3), 532–537.

Kline, Patrick M. and Enrico Moretti, "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 2014, *129*, 275–331.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Levang, Patrice, "Tanah sabrang (la terre d'en face) la transmigration en indonesie: permanence d'une politique agraire contrainte." PhD dissertation, Ensa Montpellier 1995.

Lind, Jo Thori and Halvor Mehlum, "With or Without U? The Appropriate Test for a U-Shaped Relationship," *Oxford Bulletin of Economics and Statistics*, 2010, *72* (1), 109–118.

Lubotsky, Darren, "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 2007, *115* (5), 820–867.

Lucas, Robert E. B., "Internal Migration in Developing Countries," in M. R. Rosenzweig and Ol Stark, eds., *Handbook of Population and Family Economics*, Vol. 1, Amsterdam, Netherlands: North Holland Publishing Company, 1997, pp. 721–798.

MacAndrews, Colin, "Transmigration in Indonesia: Prospects and Problems," *Asian Survey*, 1978, *18*, 458–472.

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 2012, *102*, 1508–1539.

__ **and Elias Papaioannou**, "National Institutions and Subnational Development in Africa," *Quarterly Journal of Economics*, 2014, *129* (1), 151–213.

Mohanty, S., R. Wassmann, A. Nelson, P. Moya, and S.V.K. Jagadish, "Rice and Climate Change: Significance for Food Security and Vulnerability," *International Rice Research Institute*, 2013.

Munshi, Kaivan, "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73* (1), 185–213.

__ **and Mark Rosenzweig**, "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, forthcoming.

Nunn, Nathan and Nancy Qian, "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment," *Quarterly Journal of Economics*, 2011, *126* (2), 593–650.

Olivia, Susan and John Gibson, "Value Added from Space? Empirical Evidence from Indonesia," *Unpublished Manuscript*, 2015.

Olmstead, Alan L. and Paul W. Rhode, "Responding to Climatic Challenges: Lessons from US Agricultural Development," in "The Economics of Climate Change: Adaptations Past and Present," National Bureau of Economic Research, Inc, 2011, pp. 169–194.

Peng, Shaobing, Keith T. Ingram, Heinz-Ulrich Neue, and Lewis Ziska, *Climate Change and Rice*, Berlin Heidelberg, Germany: Springer, 1995.

Perry, Douglas H., "The Economics of Transmigrant Farming," *Bulletin of Indonesian Economic Studies*, 1985, *21* (3), 104–117.

Poletaev, Maxim and Chris Robinson, "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000," *Journal of Labor Economics*, 2008, *26* (3), 387–420.

Putterman, Louis and David N. Weil, "Post-1500 Population Flows and The Long-Run Determinants of Economic Growth and Inequality," *Quarterly Journal of Economics*, 2010, *125* (4), 1627–1682.

Robins, James M., Andrea Rotnitzky, and Lue Ping Zhao, "Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data," *Journal of the American Statistical Association*, 1995, *90*, 106–121.

Robinson, Peter M., "Root-N-consistent Semiparametric Regression," *Econometrica*, 1988, *56* (4), 931–954.

Rodrik, Dani, "Unconditional Convergence in Manufacturing," *Quarterly Journal of Economics*, 2013, *128* (1), 165–204.

Roy, Andrew Donald, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3*, 135–146.

Santos Silva, J.M.C. and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

__ **and** __ , "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator," *Economics Letters*, 2011, *112* (2), 220–222.

Sarvimäki, Matti, Roope Uusitalo, and Markus Jäntti, "Long-term Effects of Forced Migration," *Unpublished Manuscript*, 2010.

Schultz, Theodore W., "The Value of the Ability to Deal with Disequilibria," *Journal of Economic Literature*, September 1975, *13* (3), 827–46.

Steckel, Richard H., "The Economic Foundations of East-west Migration during the 19th Century," *Explorations in Economic History*, 1983, *20* (1), 14–36.

Stern, Nicholas Herbert, *The Economics Of Climate Change: The Stern Review*, Cambridge, UK: Cambridge University Press, 2007.

Suri, Tavneet, "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 2011, *79* (1), 159–209.

Taye, Masresha and Ignatius Mberengwa, "Resettlement: A Way to Achieve Food Security? A Case Study of Chewaka Resettlement Scheme, Oromia National Regional State, Ethiopia," *Journal of Sustainable Development in Africa*, 2013, *15* (1), 141–154.

Verardi, Vincenzo and Nicolas Debarsy, "Robinson's Square Root of N Consistent Semiparametric Regression Estimator in Stata," *Stata Journal*, 2012, *12* (4), 726–735.

Vergara, Benito S., *A Farmer's Primer on Growing Rice: Revised Edition*, Manila: International Rice Research Institute, 1992.

Verheye, Willy, "Growth and Production of Oil Palm," in "Encyclopedia of Life Support Systems (EOLSS): Soils, Plant Growth, and Crop Production, Volume II," UNESCO-EOLSS Publishers, 2010.

World Bank, *World Development Report, 1982: Agriculture and Economic Development*, Washington, DC: International Bank for Reconstruction and Development, 1982.

— , *World Development Report, 1984: Population Change and Development*, Washington, DC: International Bank for Reconstruction and Development, 1984.

— , "Indonesia: The Transmigration Program in Perspective," A World Bank Country Study, World Bank 1988.

— , *Resettlement and Development: The Bankwide Review of Projects Involving Involuntary Resettlement, 1986–1993*, Washington, DC: International Bank for Reconstruction and Development, 1999.

— , "Operational Policies 4.12 - Involuntary Resettlement," in "The World Bank Operational Manual" 2001.

— , *Doing Business 2009: Country Profile for Indonesia*, Washington, DC: World Bank, 2008.

— , *Indonesia Country Profile*, Washington, DC: World Bank, 2009.

— , "Africa's Food Markets Could Create One Trillion Dollar Opportunity by 2030," 2013. Accessed online on 18 Sep 2015, available at http://www.worldbank.org/en/news/press-release/2013/03/04/africas-food-markets-could-create-one-trillion-dollar-opportunity-2030.

Young, Alwyn, "Inequality, the Urban-Rural Gap, and Migration," *Quarterly Journal of Economics*, 2013, *128* (4), 1727–1785.

Figures

Figure 1: Transmigration Flows and Oil Prices
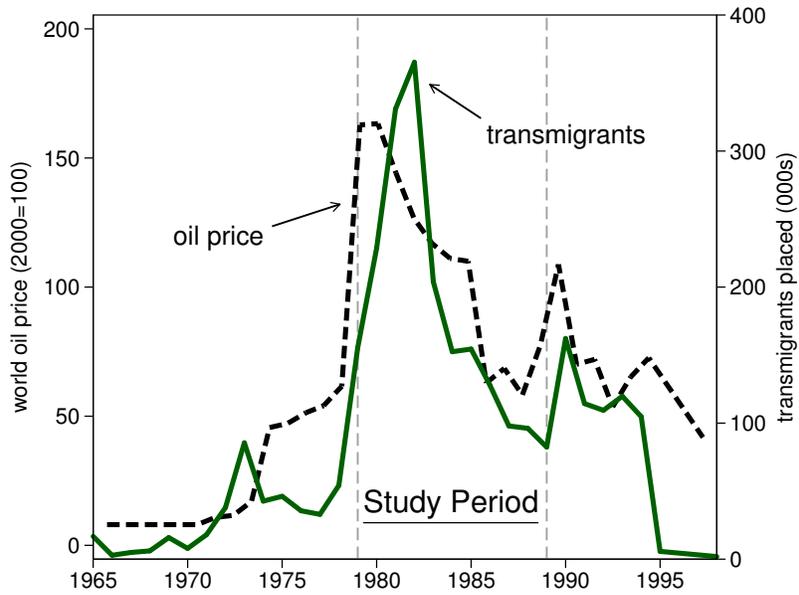
Figure 2: Map of Transmigration Villages
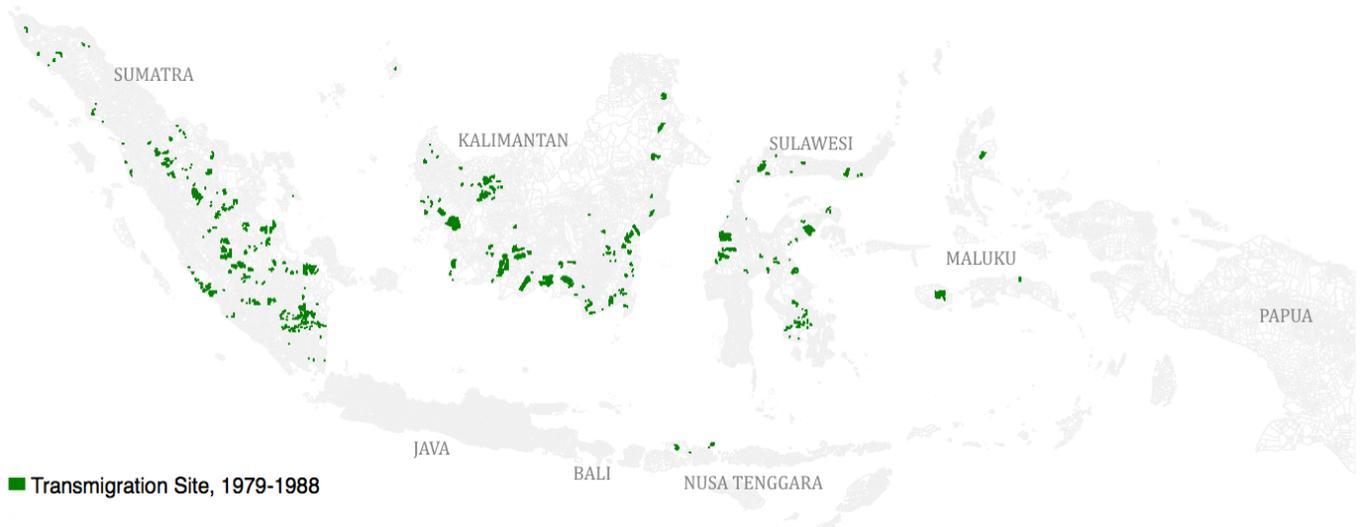
Notes: The figure shows all Transmigration villages outside of Papua settled in 1979–1988 based on our digitization and mapping of the Transmigration Villages in the 1998 MOT Census.

Figure 3: Agroclimatic Similarity: Transmigration vs. Other Outer Islands Villages
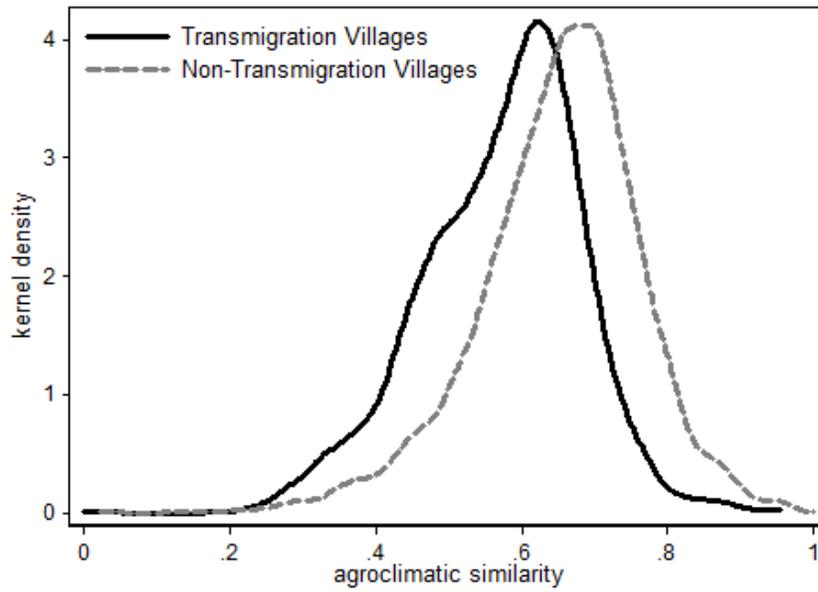
Notes: Kernel densities of village-level agroclimatic similarity computed over all immigrants in the village separately for Transmigration settlements and all other Outer Islands villages. The agroclimatic similarity indices for village j, $\mathcal{A}_j$, are constructed according to equation (2) with π_{ij} being the share of the immigrant population in j from each origin district i excluding $i = j$. All indices are standardized to lie on the unit interval.

Figure 4: Semiparametric Effects: Agroclimatic Similarity and Rice Productivity
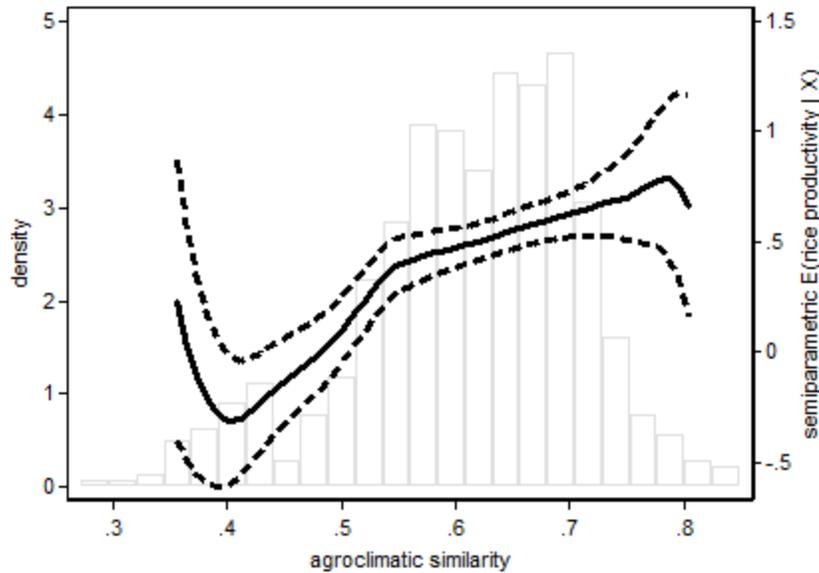
Notes: This is based on semiparametric Robinson (1988) extensions of the parametric specification in column 1 of Table 3 relating agroclimatic similarity to log rice productivity. The dashed lines correspond to 90% confidence intervals based on clustering of standard errors at the district level. The local linear regressions use an Epanechnikov kernel and a bandwidth of 0.05 and are estimated using the `semipar` command due to Verardi and Debarsy (2012). The histogram captures the distribution of standardized agroclimatic similarity. The top 5 and bottom 5 villages are trimmed for presentational purposes.

Figure 5: Effects of Agroclimatic Similarity on *Palawija* Productivity
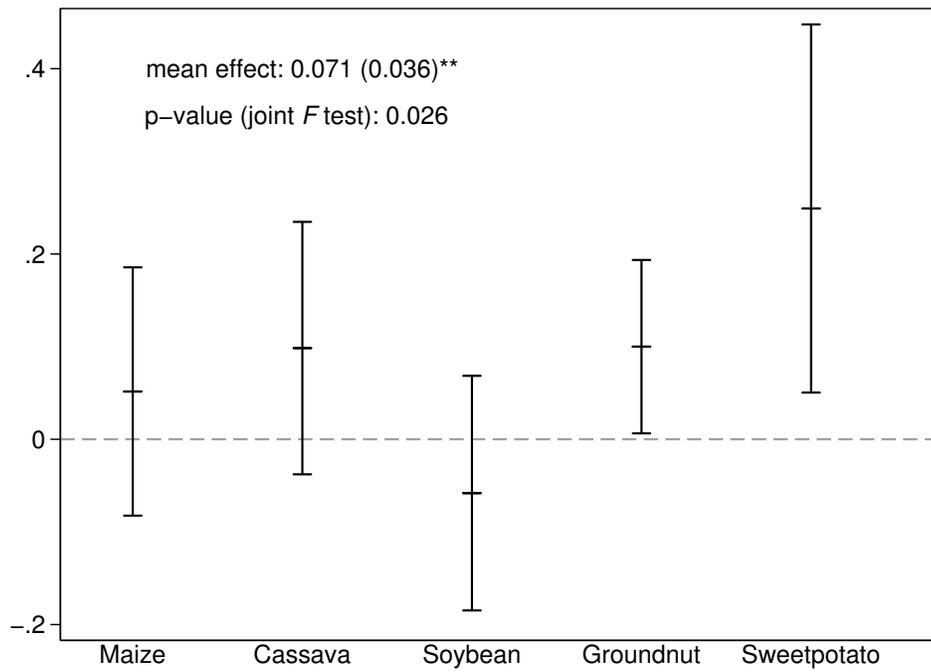
Notes: 90 percent confidence interval based on the specification from column 1 of Table 3 relating agroclimatic similarity to productivity of each *palawija* crop listed on the x-axis. We estimate the equations simultaneously while clustering the standard errors based on Conley (1999) with a bandwidth of 150km. The p-value is for the joint test that the coefficient on agroclimatic similarity equals zero for all five crops. The mean effects estimate is based on the procedure in Kling et al. (2007) (see footnote 22).

Tables

Table 1: Summary Statistics: Transmigration Villages

	Mean	Std. Deviation	No. of Villages
Demographic Characteristics			
total population (2000)	2,041	(1,283)	814
population per square km (2000)	140	(651)	814
Java/Bali-born population share	0.39	(0.19)	814
Transmigrant ethnicity population share	0.69	(0.29)	814
average years of schooling	4.00	(0.90)	814
Economic Characteristics			
farming employment share	0.69	(0.24)	814
any rice production in village	0.74	(0.44)	814
rice productivity (tons/ha)	2.52	(2.81)	600
cash crop productivity (tons/ha)	0.99	(3.00)	695
total agricultural productivity (tons/ha)	1.00	(2.65)	770
village area with any lights, 2010	0.08	(0.22)	814
light intensity, 2010	0.75	(3.07)	814
Similarity			
$\mathcal{A}_j$: agroclimatic similarity index $\in [0,1]$	0.67	(0.14)	814
$\mathcal{L}_j$: linguistic similarity index $\in [0,1]$	0.59	(0.07)	814

Notes: This table reports summary statistics for Transmigration villages. The similarity indices have been standardized to lie between zero and one. All agricultural outcomes are as observed in the 2001-2 growing season. Rice output per hectare has been winsorized above 20 tons/ha. Cash crop and total agricultural productivity are each winsorized at the fourth maximum order statistic to account for three extreme outliers. All results in the paper are robust to alternative cutoffs or not winsorizing at all. The number of villages differs for rice and total agricultural productivity as a result of missing or zero production of the given crops. See Appendix A for details on data sources and construction.

Table 2: Agroclimatic Similarity and Predetermined Development Proxies (Destinations)

Dependent Variable	agroclimatic similarity
wetland rice potential yield (ton/Ha)	0.030
	(0.030)
dryland rice potential yield (ton/Ha)	0.046
	(0.049)
cocoa potential yield (ton/Ha)	-0.063
	(0.079)
coffee potential yield (ton/Ha)	-0.105
	(0.102)
palmoil potential yield (ton/Ha)	0.008
	(0.022)
cassava potential yield (ton/Ha)	-0.005
	(0.030)
maize potential yield (ton/Ha)	-0.070
	(0.051)
log district population, 1978	-0.028
	(0.017)
own electricity (% district pop.)	-0.170
	(0.091)*
own piped water (% district pop.)	0.001
	(0.124)
own sewer (% district pop.)	-0.187
	(0.187)
use modern fuel source (% district pop.)	-1.366
	(1.419)
own modern roofing (% district pop.)	0.060
	(0.061)
own radio (% district pop.)	-0.027
	(0.196)
own TV (% district pop.)	-0.257
	(0.142)*
speak Indonesian at home (% district pop.)	-0.153
	(0.118)
literate (% district pop.)	-0.078
	(0.167)
average years of schooling in district	0.011
	(0.019)
agricultural sector (% district pop.)	0.125
	(0.079)
mining sector (% district pop.)	-0.202
	(0.505)
manufacturing sector (% district pop.)	-0.986
	(0.414)**
trading sector (% district pop.)	-0.393
	(0.265)
services sector (% district pop.)	-0.055
	(0.134)
wage worker (% district pop.)	-0.192
	(0.150)

Notes: */**/*** denotes significance at the 10/5/1 percent level. Each cell corresponds to a regression of agroclimatic similarity on the given variable in the row, island fixed effects, and the predetermined village-level control variables described in the text. Potential yields are obtained from FAO-GAEZ. The variables beginning with "log district population, 1978" are (i) based on data from the 1980 Population Census (available on IPUMS International), (ii) measured at the district level based on 1980 district boundaries, (iii) computed using the sampling weights needed to recover district-level population summary statistics, and (iv) restricted to the population in each district that did not arrive as immigrants in 1979 or earlier in 1980 (i.e., the still living population residing in the district in 1978). Standard errors in parentheses are clustered at the (1980) district level for the Census variables and allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999) for the potential yield variables.

Table 3: Effects of Agroclimatic Similarity on Rice Productivity

	(1)	(2)	(3)	(4)	(5)
			Panel A: **Rice Productivity**		
agroclimatic similarity	0.204	0.182	0.210	0.151	0.166
	(0.064)***	(0.045)***	(0.075)***	(0.057)***	(0.068)**
Number of Villages	600	600	600	600	600
R^2	0.149	0.032	0.178	0.281	0.318
			Panel B: **Cash Crop Productivity (Placebo Test)**		
agroclimatic similarity	0.024	-0.015	0.044	-0.006	-0.014
	(0.049)	(0.044)	(0.067)	(0.099)	(0.076)
Number of Villages	695	695	695	695	695
R^2	0.054	0.009	0.099	0.133	0.173
Island Fixed Effects	Yes	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	No	Yes	Yes	Yes
Origin Province Migrant Shares	No	No	Yes	No	Yes
Log Weighted Avg. Distance to Origins	No	No	Yes	No	Yes
Weighted Avg. Predetermined Controls (Table 2), Origins	No	No	Yes	No	Yes
Predetermined Controls (Table 2), Destinations	No	No	No	Yes	Yes
Demographics and Schooling	No	No	No	Yes	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. The dependent variable in Panel A is log rice output per hectare with a mean of 2.5 tons/ha, and in Panel B is the revenue-weighted log cash crop productivity with a mean of 1.0 tons/ha. The latter is calculated using crop-specific revenue-weights for 28 cash crops, primary among which are palm oil, rubber, cocoa, and coffee (see Appendix A). Agroclimatic similarity is normalized to have mean zero and a standard deviation of one. All regressions include island fixed effects and except in column 2 also include predetermined village-level control variables described in the text. "Origin Province Migrant Shares" are four variables capturing the share of the Java/Bali-born population hailing from the given province. "Log Weighted Avg. Distance to Origins" is the weighted log great circle distance between j and all Java/Bali districts i with weights equal to the share of the Java/Bali-born population from i. "Predetermined Controls, Destinations" are all of the variables reported in Table 2, and "Weighted Avg. Predetermined..." are those same variables observed in the origins i weighted by the share of j born in i. "Demographics and Schooling" are age, gender, and schooling shares for each of the Java/Bali-born and Outer Islands-born populations residing in j. Standard errors in parentheses allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999).

Table 4: Heterogeneous Effects of Agroclimatic Similarity on Rice Productivity

	(1)	(2)	(3)	(4)	(5)	(6)
agroclimatic similarity	0.424 (0.112)***					
$\cdots \times$ log potential rice yield	-0.536 (0.175)***					
$\cdots \times$ tercile 1 wetland share $\in [0, 0.16]$		0.355 (0.079)***				
$\cdots \times$ tercile 2 wetland share $\in (0.16, 0.66]$		0.141 (0.059)**				
$\cdots \times$ tercile 3 wetland share $\in (0.66, 1.0]$		0.059 (0.120)				
topographic similarity			0.070 (0.071)			0.033 (0.078)
water condition similarity				0.041 (0.071)		0.000 (0.089)
soil content similarity					0.188 (0.079)**	0.172 (0.091)*
Number of Villages	599	600	600	600	600	600
R^2	0.327	0.340	0.315	0.314	0.318	0.318
Island Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	Yes	Yes	Yes	Yes	Yes
Added Controls in Column 5 of Table 3	Yes	Yes	Yes	Yes	Yes	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. The dependent variable in all specifications is log rice output per hectare. All similarity regressors are normalized to have mean zero and a standard deviation of one. Log potential rice productivity is based on the FAO-GAEZ measure described in the text. We lose one observation relative to baseline after taking logs. Retaining this village and using potential productivity in levels or adding a small constant inside the logarithm does not affect the results. "Wetland share" denotes the fraction of agricultural land that is wetland in 2002. Topographic similarity is based on elevation, ruggedness, and slope. Water similarity is based on soil drainage, rainfall, temperature, and distance to river. Soil similarity is based on soil texture, distance to coast, carbon content, sodicity, and topsoil pH. Standard errors in parentheses allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999).

Table 5: Agroclimatic Similarity, Social Interactions, and Rice Productivity

	(1)	(2)	(3)	(4)	(5)
agroclimatic similarity	0.166	0.156	0.143	0.150	0.146
	(0.068)**	(0.064)**	(0.068)**	(0.061)**	(0.061)**
Herfindahl Index, Java/Bali origin district shares		0.039	0.243		
		(0.061)	(0.109)**		
Herfindahl Index squared			-1.570		
			(0.986)		
within-Java/Bali ethnic fractionalization		-0.032	-0.030		
		(0.053)	(0.148)		
within-Java/Bali ethnic fractionalization squared			0.007		
			(1.606)		
linguistic similarity				0.258	0.214
				(0.088)***	(0.099)**
linguistic similarity × small initial cohort					0.084
					(0.036)**
Number of Villages	600	600	600	600	600
R^2	0.318	0.320	0.322	0.330	0.325
Island Fixed Effects	Yes	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	Yes	Yes	Yes	Yes
Added Controls in Column 5 of Table 3	Yes	Yes	Yes	Yes	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. The dependent variable in all specifications is log rice output per hectare. All continuous regressors are normalized to have mean zero and a standard deviation of one. Linguistic similarity is defined in equation (5). "Within-Java/Bali ethnic fractionalization" equals $1 - \sum_{e=1}^{8} (N_{ej}/N_j)^2$ where N_{ej} is the number of individuals in 2000 from transmigrant ethnic group e, and N_j is the total transmigrant ethnic population in village j. The Herfindahl index equals $\sum_{i=1}^{I} (N_{ij}/N_j)^2$ where N_{ij} is the number of Java/Bali-born migrants from district i and N_j is the number of Java/Bali-born migrants. "Small initial cohort" in column 5 is an indicator equal to one if the village received below the median number of transmigrants placed in the initial year of settlement. Standard errors in parentheses allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999).

Table 6: Occupational Sorting within Transmigration Villages

| | Pr(Occupation = …) | | | | | |
| | Farming | | | Trading/Services | | |
	All (1)	Young (2)	Old (3)	All (4)	Young (5)	Old (6)
individual agroclimatic similarity	0.0090	0.0119	0.0079	-0.0037	-0.0050	-0.0032
	(0.0052)*	(0.0057)**	(0.0053)	(0.0027)	(0.0028)*	(0.0027)
individual linguistic similarity	-0.0139	-0.0153	-0.0134	0.0175	0.0154	0.0183
	(0.0161)	(0.0179)	(0.0155)	(0.0067)**	(0.0068)**	(0.0067)***
Number of Individuals	566,956	175,546	391,410	566,956	175,546	391,410
Dependent Variable Mean	0.622	0.489	0.682	0.099	0.089	0.103
Island Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year of Settlement Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. This table estimates the linear probability that a Java/Bali-born individual living in a Transmigration village as recorded in the 2000 Population Census works in farming (columns 1-3) or trading/services (columns 4-6). Columns 1 and 4 include all Java/Bali-born individuals between the ages of 15 and 65. Columns 2 and 5 restrict to individuals who were less than 10 years old at the time of the initial settlement in their village. Columns 3 and 6 restrict to individuals aged 10 years and greater at the time of the initial resettlement. Both similarity measures are normalized to have mean zero and a standard deviation of one. All regressions include: (i) fixed effects for the year of settlement, (ii) predetermined village-level controls used in previous tables, and (iii) individual-level controls, including age interacted with a male dummy, married dummy, indicators for seven schooling levels, Java/Bali indigenous ethnic group dummy, immigrant from Java/Bali within the last five years, immigrant from another Outer Islands province within the last five years, immigrant from district within the same (Outer Islands) province within the last five years, and indicators for seven religious groups. Results are similar omitting the individual-level controls. Standard errors are clustered at the district level.

Table 7: Agroclimatic Similarity and Crop Adjustments

| | revenue weight on | | share of cash crop farmers (3) | total agric. productivity (4) |
	cash crops (1)	rice (2)		
agroclimatic similarity	-0.043	0.047	0.001	0.014
	(0.021)**	(0.017)***	(0.022)	(0.079)
Number of Villages	770	770	770	770
R^2	0.156	0.161	0.229	0.086
Dep. Var. Mean (Levels)	0.572	0.273	0.348	0.996
Island Fixed Effects	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	Yes	Yes	Yes
Added Controls in Column 5 of Table 3	Yes	Yes	Yes	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. Agroclimatic similarity is normalized to have mean zero and a standard deviation of one. The controls are as in Column 5 of Table 3. The sample of villages is restricted to those with agricultural output data in *Podes* 2002. The dependent variable in columns 1 (2) is the share of cash crops (rice) in total potential revenue based on the approach described in Section 3.2. The dependent variable in column 3 is the share of farmers whose primary occupation is farming cash crops in the 2000 Population Census, which has separate occupational entries for food and cash crop farming. The dependent variable in column 4 is the measure of revenue-weighted agricultural productivity building on that same approach and normalizing the mean tons/ha to be one across all crops for comparability (results are similar without weighting). Standard errors in parentheses allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999).

Table 8: Agroclimatic Similarity and Nighttime Lights in 2010

| | coverage | | intensity | |
	(1)	(2)	(3)	(4)
agroclimatic similarity	0.016	0.043	0.205	0.391
	(0.007)**	(0.008)***	(0.050)***	(0.044)***
Number of Villages	814	814	814	814
R^2	0.253	0.230	0.114	0.367
Dep. Var. Mean	0.081	0.081	0.751	0.751
Estimator	OLS	OLS	Poisson	Poisson
Island Fixed Effects	Yes	Yes	Yes	Yes
Predetermined Village Controls ($\mathbf{x}_j$)	Yes	Yes	Yes	Yes
Added Controls in Column 5 of Table 3	No	Yes	No	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. Agroclimatic similarity is normalized to have mean zero and a standard deviation of one. The controls are as in Table 3 with the addition of indicators for the year the village was established. The dependent variables are the two measures of nighttime lights capturing, respectively, the fraction of the village with any light *coverage* and the average *intensity* of nighttime lights. The coefficients in columns 3-4 are based on Poisson pseudo maximum likelihood, and the average marginal effects simply equal the coefficient × the mean of the dependent variable. Standard errors in parentheses allow for unrestricted spatial correlation between all villages within 150 kilometers of each other (Conley, 1999).

Table 9: Quasi-Gravity Regression of Migration from Java/Bali to the Outer Islands

| | $\mathbf{Pr}(migrants_{ij} > 0)$ | | $\ln(migrants_{ij})$ | |
	(1)	(2)	(3)	(4)
agroclimatic similarity	0.0027	0.0015	-0.0004	0.0001
	(0.0066)	(0.0069)	(0.0200)	(0.0220)
$(-1)\times$ log distance	0.1262	0.1272	0.1287	0.2036
	(0.0192)***	(0.0238)***	(0.0597)**	(0.0753)***
Observations	96,866	96,866	37,446	37,446
Dep. Var. Mean (Levels)	.39	.39	16.8	16.8
Birth District (Java/Bali) Fixed Effects	Yes	Yes	Yes	Yes
Island Fixed Effects	Yes	Yes	Yes	Yes
Year of Settlement Fixed Effects	Yes	Yes	Yes	Yes
Individuals Placed in Year of Settlement	Yes	Yes	Yes	Yes
Predetermined Controls (Table 2), Destinations	No	Yes	No	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. This table regresses the stock of migrants from origin district i in Java/Bali residing in Outer Islands village j in the year 2000 on the agroclimatic similarity between i and j and the inverse log great circle distance between i and j. The unit of observation is an origin district i (of which there are 119) by destination Transmigration village j. The dependent variable in columns 1-2 is an indicator equal to one if there are migrants from i in j. The dependent variable in columns 3-4 is the log number of migrants from i in j. All specifications include birth district fixed effects, destination island fixed effects, the log number of transmigrants placed in the initial year of settlement, and indicators for the year of settlement. Columns 2 and 4 additionally control for the predetermined district-level variables reported in Table 2. Results are similar using destination district or village fixed effects. Standard errors are two-way clustered by birth district and destination village.

Table 10: Average Treatment Effects of the Transmigration Program

Dependent Variable	(1)	(2)	(3)	(4)
log population density	-0.390	0.556	0.799	0.769
	(0.118)***	(0.132)***	(0.220)***	(0.170)***
any rice production	-0.041	-0.094	-0.027	-0.029
	(0.036)	(0.035)***	(0.059)	(0.060)
log rice productivity	-0.316	-0.241	-0.035	-0.166
	(0.099)***	(0.134)*	(0.175)	(0.218)
log revenue-weighted avg. agricultural productivity	-0.051	-0.193	0.023	0.134
	(0.083)	(0.136)	(0.172)	(0.142)
log revenue-weighted total agricultural output	0.641	0.170	0.410	0.472
	(0.134)***	(0.186)	(0.247)*	(0.258)*
percent any light coverage, 2010	-0.187	0.008	0.018	0.009
	(0.030)***	(0.017)	(0.033)	(0.025)
Treatment/Control Only	No	Yes	Yes	Yes
Geographic Controls	No	Yes	Yes	Yes
Reweighting	No	No	Yes	Yes
Blinder-Oaxaca	No	No	No	Yes

Notes: */**/*** denotes significance at the 10/5/1 percent level. Each cell reports the coefficient from a regression of the given dependent variable on an indicator for whether the village is a Transmigration village. Column 1 comprises all Outer Islands villages (with non-missing data). Column 2 restricts to our quasi-experimental design including only Transmigration and control/RDA sites and conditions on the predetermined village-level characteristics that explain (sequential) site selection. Column 3 is a double robust specification (Robins et al., 1995) that (i) reweights controls by normalized $\widehat{\kappa} = \widehat{P}/(1-\widehat{P})$ where $\widehat{P}$ is the estimated probability that the village is a Transmigration settlement and (ii) controls for the predetermined village-level characteristics. Column 4 is a control function specification based on a Blinder-Oaxaca decomposition developed in Kline (2011). All specifications include island fixed effects. Sample sizes vary across outcomes (depending on data availability) and columns but include as many 31,185 villages in column (1), and 814 treated villages and 668 controls in columns 2-4. Standard errors are clustered by district in parentheses and are estimated using a block bootstrap in column 3 to account for the generated $\widehat{\kappa}$ weights.