NBER WORKING PAPER SERIES

# THE AGGREGATE PRODUCTIVITY EFFECTS OF INTERNAL MIGRATION: EVIDENCE FROM INDONESIA

Gharad Bryan
Melanie Morten

Working Paper 23540
http://www.nber.org/papers/w23540

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2017

We thank Paco Buera, Rebecca Diamond, Dave Donaldson, Greg Fischer, Pete Klenow, David Lagakos, Torsten Perrson, Steve Redding, Daniel Sturm, Adam Sziedl, Michael Waugh, Ben Olken and seminar audiences at the Stanford Institute for Theoretical Economics, Paris School of Economics, Namur, Helsinki, Berkeley, Yale, LSE, Columbia, the University of Chicago, Toronto, NBER SI 2015, Harvard/MIT, and the CEPR/PODER conference for helpful comments and suggestions. Anita Bhide, Allan Hsiao and Jay Lee provided outstanding research assistance. Any errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia
Gharad Bryan and Melanie Morten
NBER Working Paper No. 23540
June 2017
JEL No. J61,O18,O53,R12,R23

**ABSTRACT**

We estimate the aggregate productivity gains from reducing barriers to internal labor migration in Indonesia, accounting for worker selection and spatial differences in human capital. We distinguish between movement costs, which mean workers will only move if they expect higher wages, and amenity differences, which mean some locations must pay more to attract workers. We find modest but important aggregate impacts. We estimate a 22% increase in labor productivity from removing all barriers. Reducing migration costs to the US level, a high mobility benchmark, leads to an 8% productivity boost. These figures hides substantial heterogeneity. The origin population that benefits most sees an 104% increase in average earnings from a complete barrier removal, or a 37% increase from moving to the US benchmark.

Gharad Bryan
London School of Economics
Houghton Street
London WC2A 2AE
United Kingdom
g.t.bryan@lse.ac.uk

Melanie Morten
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
memorten@stanford.edu

# 1 Introduction

Recent evidence suggests that a policy of encouraging internal labor migration could have large productivity effects in developing countries. On the macro side, Gollin et al. (2014) show that non-agricultural (urban) workers produce 4 times more than their agricultural (rural) counterparts. On the micro side Bryan et al. (2014) show a 33% increase in consumption from experimentally induced seasonal migration. Neither of these results, however, is definitive. The experimental estimates apply to seasonal migration, and a specific part of Bangladesh. The macro estimates don't account for selection on unobservables (Young 2013), and only apply to movement between rural and urban areas.

This paper uses micro data from Indonesia to quantify the aggregate effect of increasing mobility. Two observations motivate our approach. First, migration could increase productivity if it: a) allows individuals to sort into a location in which they are personally more productive (sorting); b) allows more people to live in more productive locations (agglomeration); or c) both.[1] Second, in the absence of constraints or frictions people will maximize their production; therefore, a policy that encourages migration will have no effect on output if there are no existing constraints on mobility.[2]

We build a model in which workers have idiosyncratic location-specific productivity and, in which locations differ in their overall productivity. This setup allows for both sorting and agglomeration effects. Into this framework, we incorporate two kinds of mobility constraints. Movement costs exist if workers must be paid higher wages to induce them to work away from home. Compensating wage differentials exist if workers must be paid higher wages to work in low amenity locations. The result is a general equilibrium Roy model in which workers sort across locations that have heterogeneous amenities and productivities.[3] The model is similar to that used by Hsieh et al. (2016); our approach

[1] Selection is known to be important in migration (e.g. Borjas 1987; Young 2013; Lagakos and Waugh 2013). Much of the literature treats selection as an econometric problem. In our model, sorting is both something that must be accounted for, and also a source of gains from labor market integration.

[2] Our view is that if one cannot find evidence of constraints, then productivity gaps found in the macro literature (e.g. Gollin et al. 2014) are likely driven by selection. Our approach is similar to Lagakos and Waugh (2013) who observe that the food problem constrains movement across sectors. Like us, they consider how this constraint affects aggregate productivity in the presence of worker selection.

[3] We also allow for workers from some locations to have lower human capital. With limited movement, this heterogeneity can drive spatial wage gaps.

also has close connections to the seminal work of Hsieh and Klenow (2009).[4] We use this structural framework to quantify the change in aggregate productivity that would result from removing movement costs and/or equalizing amenity differentials. Like Hsieh and Klenow (2009) and Caselli (2005), we do not consider specific policies, but rather try to quantify the potential impacts of a set of policy options. [5]

Before turning to our structural analysis, we document five motivational facts, which suggest both that movement costs and compensating differentials exist, and that selection is important in the data. Indonesian data that records location of birth, current location and earnings allows us to demonstrate these facts. In the case of movement costs, we first show that a gravity relationship holds in the data. A 10% reduction in the distance between two locations leads to a 7% increase in the proportion of migrants that flow between the two locations.[6] We also show that people who live farther from their location of birth have higher wages. A doubling of distance leads to a 3% increase in average wages, suggesting that people need to be compensated to induce them to move away from home. In running these regressions, we think of distance as a proxy for movement costs, which may not capture all policy relevant constraints. For compensating differentials, we show that workers in observably low amenity locations receive higher wages. Again, our measure of amenities is, at best, a proxy. Furthermore, selection effects, appear to be an important in the data: the greater the share of people born in origin $o$ that move to destination $d$, the lower their average wage. The elasticity of average wage with respect to share is approximately $-0.04$. Importantly, because our model is one in which movement costs reduce migration and lead to selection we show that there is almost no

[4]Our framework also has much in common with recent quantitative models of economic geography such as Allen and Arkolakis (2014), Redding (2016) and Desmet et al. (2016). The framework is similar that used in work by Tombe and Zhu (2015). Relative to that paper, we use more detailed micro data which enables us to directly estimate the extent of selection, and we are interested in a different set of questions.

[5]Policies to reduce movement costs might include road construction (see, e.g., Morten and Oliveira 2016), religious tolerance building, language training, migration subsidies (see, e.g. Bryan et al. 2014), or telecommunications improvements. Policies to reduce compensating differentials might include industrial pollution abatement, hospital building, or removal of building height restrictions (See e.g. Harari 2016). This latter policy would reduce rents. High rents enter our calculation as low amenities which reduce labor movement.)

[6]We consider all movement costs to be relative to a person's location of birth, which we refer to as their origin. So, a person who is born in Melbourne and moves for 10 years to New York and then to London pays a cost of living in London that is relative to Melbourne, not New York.

effect of distance on average wages once the proportion of the origin population at the destination is controlled for. All of these effects are predicted by our model.

To estimate the potential effects of policy, we turn to our structural model. When estimating the model, we treat both movement costs and amenity differentials as non-parametric objects to be inferred from the data. This reflects our view that amenities are hard to measure and distance is unlikely to capture all policy relevant dimensions of movement costs. Our model allows for straightforward quantification of the effects of reducing movement costs, or amenity-driven wage differentials.[7] The intuition is straightforward. We first generate counterfactual population distributions by estimating where people would live if we removed their empirical tendency to stay at their place of birth and their tendency to avoid some locations that have high measured productivity.[8] Next, we ask how productivity would change if people moved as suggested by our counterfactuals. Our model of selection implies that each additional migrant will earn less than the last; to account for this we need to understand how wages change as workers move. Since selection, in our model, is relative to location of birth, it is average wages of people from a given origin who live in a given destination that matter. Our unique data, which captures both location of birth and current location of work, combined with an IV strategy inspired by our model, allows us to estimate the relevant elasticity.[9]

Our results suggest moderate aggregate gains, but important heterogeneity. Removing all frictions is predicted to increase aggregate productivity by 22%. These gains are modest relative to the potential gains suggested by studies such as Gollin et al. (2014), but are inline with what one may expect from other microeconomic studies. For the people born in some locations, however, the results are much larger, with predicted gains peaking at 104%. Because complete barrier removal may be impossible, we also compute the gains from moving to a US level of movement costs, which we see as a high-mobility

[7] While the intuition presented here is accurate, the analysis requires functional form assumptions. For example, we rely heavily on the assumption that individual productivities are drawn from a Fréchet distribution. In this we follow the seminal work of Eaton and Kortum (2002) and also Hsieh et al. (2016).

[8] As explained in Section 4 productivity is estimated as a fixed location effect that remains after adjusting for the selection of workers in to specific locations.

[9] We also account for the potential endogeneity of location productivity (due to agglomeration and price effects) and location amenity (due to congestion effects) by using consensus estimates from the literature.

benchmark.[10] We predict an aggregate productivity boost of 8%, with the origin that gains most seeing a 37% increase. We conclude that unless dynamic effects are large, encouraging migration is unlikely to have very large productivity effects or the sort estimated, for example, by Hsieh and Klenow (2009). Targeted policies may, however, have big impacts on the lives of some communities.[11]

Our paper differs from existing approaches in three ways. First, we consider region-to-region rather than rural-to-urban movement. Since Lewis (1954) and Harris and Todaro (1970) the development and migration literature has been dominated by rural to urban studies. In our setting this is potentially inappropriate. Figure 1 shows kernel density plots of the log of the average monthly wage, calculated at the regency level and broken down by rural/urban status.[12] The figure highlights that while there is large variation between regencies, there is little variation across rural-urban. Table 1 shows that the majority of migration also occurs within category, rather than across category: between 75 and 85% of migration out of urban areas is to another urban area, and between 25 and 30% of migration out of a rural area is to another rural area. Focusing only on rural-urban migration misses the within rural and the within urban migrations.[13]

Second, we focus on counterfactual estimates that predict the effect of removing constraints. While we can learn much from work documenting returns to past migration,[14]

[10] This counterfactual keeps Indonesian amenity difference constant, and just lowers movement costs to the US level. Inline with the general argument in Hsieh and Moretti (2017) we find that the US has greater amenity dispersion than Indonesia.

[11] This last observation is in line with Lagakos et al. (2015) who find small aggregate welfare effects, but that migration subsidies may be well-targeted towards households that are particularly poor. Relative to that paper, we focus on permanent migration rather than seasonal migration, in a setting where we observe positive, rather than negative, sorting.

[12] Our data records an individual's current regency of residence, and whether their location is considered urban or rural. The data also records their regency of birth, but does not record whether their birth location was urban or not. To fill this data gap we code regencies that have greater than median rural population share as rural, and the remaining regencies as urban. Appendix Figure 1 shows that the same patterns hold if rural/urban is defined on the individual level.

[13] We find broadly similar back-and-forth migration rates as Young (2013) who uses DHS data to estimate rates of urban and rural migration. The DHS data code a respondent's origin as one of capital/other city/town/countryside but does not give the actual location of birth. Young (2013)'s estimate for males is that 63% of urbanites migrate, and of these, 41% move to a rural area. For 1976, we find out migration rates of 54%, with 20% migrating to a rural area. Relative to the DHS data our data gives the actual location of birth so we are able to define rural and urban migrants based on the characteristics of that specific location.

[14] Recent work by Kleemans and Magruder (2017); Hicks et al. (2017); Beegle et al. (2011); Garlick et al. (2016) provide important estimates of the returns to, and impact of, past migrations in Indonesia, Kenya, Tanzania and South Africa.

there are challenges moving from these estimates to predictions of future returns. On one hand, selection effects mean future migrants may earn less than past migrants; on the other hand, migration policies work by reducing constraints, and so will tend to encourage migration where past movement was minimal. Because of this, past returns may contain little information on the likely effects of future policies. For our analysis we directly estimate the impact of removing constraints. Our only use of past migration is to estimate the strength of selection effects. While this approach is similar to macroeconomic estimates based on productivity gaps (e.g. Gollin et al. 2014), it accounts for selection effects that are likely to be important.

Finally, we take account of GE effects. First, by incorporating sorting, we allow for aggregate productivity gains in the absence of large net populations flows. Second, we calibrate agglomeration, congestion and price elasticities using consensus estimates, and we then assess how our results depend on these parameters.

Our results are limited in three ways. First, we look only at static gains, leaving examination of dynamic effects for future work.[15] Second, we look only at productivity and only at gains. We do not consider welfare effects of migration and we do not consider the costs of policy. A full consideration of costs is difficult and can be avoided if benefits are small. Third, we do not consider specific policies, but rather provide estimates of the total gains that may be available. Our approach is similar, therefore, to the development accounting and macro misallocation literatures (Caselli, 2005; Hsieh and Klenow, 2009).

The paper starts by laying out five motivational facts. These facts strongly suggest that spatial labor markets in Indonesia are characterized by costs of movement, compensating wage differentials and selection on productivity. The facts imply the possibility of productivity gains from increased movement. We then provide a simple two-location example that explains how we quantify the possible gains. We follow this by briefly describing out our formal model, discussing identification and estimation, and demonstrating that our structurally estimated parameters correlate sensibly with real world proxy measures. Finally, we present results from counterfactual exercises.

[15]There are several potential sources of dynamic gains. For example, migration costs may be endogenous (Carrington et al., 1996), firm openings may depend on the pool of available migrant labor, or both.

# 2 Data, Motivation and Two Location Example

## 2.1 Data

Our approach has specific data requirements. In our view, people will only migrate if their earnings increase enough to compensate them for living away from home (which we take to be their location of birth). We therefore need data that records an individual's location of birth, current location of work, and earnings. Our interest in aggregate returns implies that data has to be geographically representative. Because we want to non-parametrically estimate movement costs, the data set must be large enough that it records flows between all pairs of locations. Data of this kind is available in very few locations, and Indonesia is the unique country that meets these specifications and has location recorded at a level below the equivalent of a state.

Our Indonesian data come from the 1995 SUPAS (Intercensal Population Survey) and from the 2011 and 2012 SUSENAS (National Socioeconomic Survey). These data sets record, for a large representative set of people, location of birth (from now on origin $o$), current work location $d$ (which could be the same as the origin) and monthly earnings (which we refer to as the wage).[16]

We also use data from the United States, both to show that our migration facts hold more generally, and to generate a suitable counterfactual for a high-mobility economy. We use the 1990 5% Census sample and the 2010 American Community Survey, as these dates overlap most closely with our Indonesia dates.

We restrict the sample to be male head-of-households between 15 and 65 years old. Summary stats for the Indonesian and the United States sample are given in Appendix Table 1. In the US, we have location of birth and work recorded at the level of the state; in Indonesia, we have this for either regency or province.[17] Because of the census nature

[16] In all the surveys except the 2011 SUSENAS income information is only recorded for those who are employed by someone else. This means that we are missing income information for those who are self employed which includes those who work in agriculture. To understand whether this includes bias we supplement this census data occasionally with data from the Indonesian Family Life Survey (IFLS) from 1993, 1997, 2000 and 2007. This data has a much smaller sample and also by design covers only 13 out of 25 provinces, but does collect much more detailed information on incomes. Summary statistics for the IFLS data are given in Appendix Table 2.

[17] Regency is a second level administrated subdivision below a province and above a district. For all

of our data, our measure of migration is permanent migration.[18] All wage variables are reported in monthly terms.

We use the 2005 and 2011 Village Potential Statistics (PODES) datasets to get measures of amenity. From the village-level data, we collapse to regency levels, using population weights.

## 2.2 Five Empirical Facts About Migration

From our data, we can calculate the proportion of people from each origin $o$ that move to each destination $d$, which we denote $\pi_{do}$ as well as the average wage within origin destination pair, $\overline{wage}_{do}$. Using this data, we document five empirical facts about migration. We present these five facts at the regency level. For the later estimation of the model, we aggregate regencies into provinces.[19]

**Fact 1** (Gravity: Movement Costs Affects Location Choice)**.** *Controlling for origin and destination fixed effects, the share of people born in o that move to d is decreasing in the distance between o and d.*

To document Fact 1, we run a regression

$$\ln \pi_{dot} = \delta_{dt} + \delta_{ot} + \beta \ln dist_{do} + \epsilon_{dot}$$

surveys, we drop the provinces of Papua and West Papua. We generate a set of regencies which have maintained constant geographical boundaries between 1995 and 2010. This primarily involves merging together regencies that were divided in 2001. This leaves us with a sample of 281 regencies, where the average regency population surveyed in 2011 is 3700 people. Later, for the structural estimates we aggregate regencies up to the level of province, of which there are 25.

[18] Substantial amount of migration is temporary, rather than permanent (Bryan et al. 2014; Morten 2017). Further, we are not able to track people across census waves, and so do not have information on people who have moved multiple times, or who have moved and returned home. To ascertain the scope of these issues we look at detailed migration histories collected in the IFLS. We find that close that multiple and return migration is not a large issue in our context. Conditional on moving out of the birth regency, 49% of all migrants make only 1 migration; 27% make two moves (of which, 30% of these second moves are to return home); and only 12% make 4 or more moves (tables available upon request). These numbers are broadly similar to those for the US: the average male migrant makes 1.98 moves and 50.2% of movers move home (Kennan and Walker, 2011).

[19] We concentrate here on the Indonesian data. Appendix Table 3 shows that the first four facts also hold in the US (we do not document Fact 5 in the US as we do not have amenity data at the state level). The Indonesian results are also robust to aggregating to the province level (Appendix Table 4) and using the IFLS data (Appendix Table 5).

where $\delta_{dt}$ and $\delta_{ot}$ are destination-year and origin-year fixed effects respectively and $dist_{do}$ is the straight distance between regency $o$ and regency $d$.[20] The destination effect controls for any productivity or amenity differences across destinations, and the origin effect controls for the benefits of other possible locations from the perspective of those living at the origin (this term is similar to the multilateral resistance term in the trade literature.)

We interpret distance as a proxy for migration costs. The results are shown in Table 2 Column 1. We estimate that the elasticity of $\pi_{do}$ with respect to $dist_{do}$ is negative, strongly significant, and sizeable. A 10% increase in distance leads to a 7% reduction in the proportion migrating. These results strongly suggest that there are costs of moving people across space.

**Fact 2** (Movement Costs Create Productivity Wedges)**.** *Controlling for origin and destination fixed effects, the average wage of people born in destination o and living in destination d is increasing in the distance between o and d.*

To establish Fact 2, we run the regression

$$\ln \overline{wage}_{dot} = \delta_{dt} + \delta_{ot} + \beta \ln dist_{dot} + \epsilon_{dot}.$$

The results are shown in Tables 2 Column 2. We estimate that the elasticity of the average wage with respect to distance is positive, strongly significant, and sizeable. A doubling of the distance between origin and destination leads to a 3% increase in the average wages. These impacts can be very large. For example, the straight line distance from Denpasar to Jakarta on the western tip of Java is about 1000km. On the other hand, the distance from Denpasar to Banyuwangi on the eastern tip of Java is about 100km. Our estimates suggest that the average wage of migrants from Denpasar to Jakarta will be 30% more than those to Banyuwangi.

As we explain in more detail in our two location example below, this fact suggests that movement costs reduce productivity. To easily illustrate this, consider two locations $d$ and $d'$ that are identical except that $d$ is closer to $o$. Fact 2 implies that those who choose

[20] $dist_{do}$ is the straight line distance, in kilometers, between the centroid of regency $o$ and the centroid of regency $d$. We have experimented with movement time, generated using Dykstra's algorithm and assumptions about the time cost of different types of travel. This does not materially affect the results.

to move to $d'$ have higher average wages than those who choose to move to $d$. Under the hypothesis that the two destinations are identical, that workers are rational and are paid their marginal products, the only way that those in $d'$ can have higher wages is if migration costs dissuaded some positive productivity movers, who would have earned less.

**Fact 3** (Selection)**.** *Controlling for origin and destination fixed effects, the elasticity of average wages with respect to origin population share is negative.*

Fact 3 is documented by running the regression

$$\ln \overline{wage}_{dot} = \delta_{ot} + \delta_{dt} + \beta \ln \pi_{dt} + \epsilon_{dot}. \tag{1}$$

Estimates from this regression are presented in Table 2 Column 3. Our estimates, which are strongly statistically significant, show that the elasticity of average wages is negative. In Indonesia, a doubling of the share of people that migrate to a particular destination leads to a 4% decrease in average wages. This fact is subject to a potential endogeneity concern: any shock to productivity in destination $d$ that differentially affects people from different origins $o$ will tend to also alter $\pi_{do}$. Below, we use our full theoretical model to motivate an instrument to correct for this. Instrumentation changes the quantitative results, but does not alter the qualitative fact.

**Fact 4** (Movement Costs Reduce Productivity by Reducing Selection)**.** *The elasticity of average wage to distance drops to almost zero after controlling for the fraction of origin population that migrates.*

We document Fact 4 by running the regression

$$\ln \overline{wage}_{dot} = \delta_{ot} + \delta_{dt} + \beta \ln \pi_{dt} + \gamma \ln dist_{do} + \epsilon_{dot}. \tag{2}$$

Results are presented in Table 2 Column 4. The coefficient on $\ln \pi_{dt}$ changes little when the distance control is added, but the magnitude of the estimated distance effect, while still positive and statistically signficant, drops relative to the results in Column 2, falling to an economically insignificant size.

Facts 3 and 4 together suggest a framework where increasing movement costs, proxied here by distance, leads to a reduction in the proportion of people that move (Fact 1). This, in turn, leads to an increase in wages (Fact 2), but these wage effects are generated by a selection effect created by a reduced proportion moving (Facts 3 and 4). This is consistent with our discussion of Fact 2, where we assume that workers are paid their marginal productivity, so once destination and origin fixed effects are controlled for wage differences reflect selection. Importantly, Fact 4 suggests that our structural approach of estimating the impact of reducing movement costs using the elasticity of wage with respect to proportion moving will capture most of the effects of removing movement cost.

**Fact 5** (Compensating Wage Differentials)**.** *Controlling for origin fixed effects, locations with higher amenities have lower wages.*

To document Fact 5 we run the regression

$$\ln \overline{wage}_{dot} = \delta_{ot} + \delta_{dot} + \beta amen_{dt} + \epsilon_{dot}$$

where $amen_{dt}$ is measured amenity in destination $d$ at time $t$. To determine amenity, we take six different measures of amenity from the Indonesian PODES survey and convert to a single measure by taking the first principle component. We then standardize to give this variable a zero mean and unit standard deviation. The results are shown in Table 2 Column 5. Our estimates imply that a one standard deviation increase in amenities leads to a 2.3% decrease in average wages. This is direct evidence that firms pay a compensating wage differential to attract workers to low amenity locations. Importantly, there is little endogeneity concern with the sign of this result. While one may be concerned that higher wage locations can afford higher amenities, this result goes in the opposite direction.

## 2.3 An Example with Two Locations

In this section we briefly discuss a two-location version of our model. We highlight the mechanisms through which migration costs and amenity differentials reduces productivity. We also show how we estimate the productivity impacts of policies that reduce migration frictions. Because of the simplicity of the two-location model, we can give an intuitive graphical analysis.

We think of each work place, or destination $d$, as being characterized by a productivity $w_d$ and amenity $\alpha_d$. We also assume that each location produces different goods and that people's productivities depend on their location. In particular, we assume that the wage of person $i$ living in destination $d$ is $w_d s_{id}$, where $s_{id}$ is the skill level of person $i$ for location $d$. Total utility for person $i$, from location $o$ who decides to live and work in destination $d$, is then $\alpha_d w_d s_{id}(1-\tau_{do})$, where $\tau_{do}$ is the cost that a person born in origin $o$ pays to live in destination $d$. We refer to $\tau_{do}$ either as a movement cost or migration cost. We assume that $\tau_{do} \in [0,1]$, $\tau_{oo} = 0$ and $\tau_{do} = \tau_{oo}$. In our empirical work we will back out $\alpha_d$ and $\tau_{do}$ as residuals, and so this way of writing the utility function normalizes the measure of amenities and movement costs relative to wages.

Figure 2 shows the distribution of skill ($s_{id}$) across two locations, which we call $A$ and $B$; the figure is drawn from the perspective of people born in location $B$. If there were no frictions, people would live where their earnings, $w_d s_{id}$, are highest. As drawn, location $A$ has the higher productivity, and all those above the ray $OE$, which has slope $w_B/w_A$, should move to location $A$ (that is those in regions $I$, $II$, and $III$ should migrate). If the two locations had equal productivity, those above the 45 degree line (in areas $I$ and $II$) should move to maximize productivity.

With movement costs, people from $B$ must be compensated for their move to $A$. This means that earnings in $A$ are effectively less valuable, and only those above the line $OC$, which has slope $w_B/w_A(1-\tau_{AB})$ will choose to move. We can divide those born in location $B$ into four groups. Those below ray $OE$ (the blue dots in region $IV$) should not move, because their returns are highest to stay in B, and they do not. Those above $OE$ and below the 45 degree line (the red dots in region $III$) should move, because $A$ has higher productivity than $B$. The higher productivity in $A$ compensated these people for the fact that their comparative advantage lies in $B$. With movement costs, these people do not move. Those above the 45 degree line and below ray $OC$ (the yellow dots in region $II$) should move, for two reasons. First, they have a comparative advantage in location $A$.

Second, $A$ is a more productive location. Consider person $x$: she loses productivity equal to the distance $xy$ because she has a comparative advantage in $A$ but does not move, and an additional amount $yz$ because $A$ is more productive. These two channels mean that movement costs reduce productivity by reducing sorting, and by reducing agglomeration in high-productivity locations. Finally, those above $OC$ in region $I$ should move and they always do.

Fact 2 and its interpretation can be seen in this diagram. As movement costs increase, fewer people move to $A$ and the wages of those that move increase. This increase occurs because some people who would have been more productive in $A$ now choose to stay in $B$.[21]

Amenities also move worker locations away from the productivity-maximizing allocation. With amenities, but no movement costs, people now maximize $\alpha_d w_d s_{id}$. The effect can be understood in the same diagram. With no movement costs and $B$ having higher relative amenity, the ray $OC$ would have slope $\frac{\alpha_B w_B}{\alpha_A w_A}$. The same effects – a lack of sorting and too little agglomeration – are present and so long as the level of amenity in $A$ differs from the level of amenity in $B$ productivity will not be maximized. The main difference between amenity differentials and movement costs is that movement costs will reduce migration relative to home, while amenity differentials reduce the number of people living in one location relative to the other.

It is worth noting that selection plays two roles in our model. On one hand, worker heterogeneity and selection are a source of gains. Movement costs, which stop workers from moving to their location of comparative advantage, reduce productivity. On the other hand, selection limits the potential gains from moving more workers to high-productivity locations. In the absence of selection on productivity, all workers who move will have the same wage, and so aggregate impacts of removing amenity differentials can be larger.

Our empirical task is to estimate the gain in productivity that would come from allocating people to their productivity maximizing location. This problem can be separated

[21] This fact depends on the properties of skills distribution. In the language of Lagakos and Waugh (2013), comparative and absolute advantage must be aligned. Our Facts 2 and 3 imply that this is the case in our setting. See also Adao (2016) for a discussion.

into two parts. First, we estimate the movement response. This is equivalent to estimating how many people lie in the triangle $OCE$ (those colored yellow and red). This is conceptually straightforward. In the case where there are no productivity differences between locations, the productivity maximizing choice is that half the people from $B$ will stay in $B$ and half will live in $A$. Second, we estimate how this movement will affect the average wages of the four groups in our data: those from $A$ that move to $B$, those from $B$ that live in $A$ and those that stay in $A$ or $B$. Functional form assumptions laid out below imply that average wages for these groups are a constant elasticity function of the fraction of the origin population that live in the destination. This elasticity is estimable given our data, which records origin and destination, and is shown in Fact 3 above.[22] Because our data records the proportion of people from each origin who live in each destination $\pi_{do}$, and counterfactual population distributions can be expressed in the same way, this elasticity is sufficient to estimate the counterfactual aggregate productivity. In the next two sections, we lay out how these ideas extend to more than two locations, how to account for heterogeneous location productivities, and how we incorporate general equilibrium effects.

# 3 Model

In this section we present a simple general equilibrium model of migration. The model is an extension of the labor sorting model in Hsieh et al. (2016), which itself draws on Eaton and Kortum (2002). The model also has similarities with recent work on quantitative economic geography, particularly Allen and Arkolakis (2014). The economy consists of $N$ locations. Each location can be an origin "$o$" or destination "$d$" for a given migrant. Workers are born in a particular origin, draw a skill for each destination, and sort across destinations according to wages, amenity and migration costs. Wages and amenities are endogenous and adjust to ensure equilibrium. We first discuss how work-

[22] Our data records origin, current location, and, crucially, current earnings at the individual level. Because we have individual level wage data we can therefore estimate the sorting parameter. This additional piece of information differs out work from the estimation of Tombe and Zhu (2015), who cannot estimate this elasticity directly given their data.

ers choose where to live and work taking wages and amenities as given, and then turn to production and general equilibrium determination of wages and amenities.

## 3.1 Utility and Sorting

$L_o$ individuals are born in each origin $o$. Each person $i$ receives a skill draw $s_{id}$ for each possible work destination $d \in N$. The individual also receives a skill draw for her location of origin. Skill is drawn from a multivariate Fréchet distribution,

$$F(s_1, \ldots, s_N) = \exp\left\{ - \left[ \sum_{d=1}^{N} s_d^{-\frac{\tilde{\theta}}{1-\rho}} \right]^{1-\rho} \right\},$$

which does not depend on the location of birth.[23] Here, $\tilde{\theta}$ measures the extent of skill dispersion or the importance of comparative advantage. As $\tilde{\theta}$ decreases, there is a greater difference between skills across locations. $\rho$ measures the correlation in skills across locations. As $\rho$ increases, individuals with a high draw in destination $d$ are also likely to have a high draw for destination $d'$. The interpretation is that each different location has a different set of required skills. To the extent that $\tilde{\theta}$ is estimated to be high, locations do not differ greatly in their skill requirements. We allow for correlation between skill draws to allow for general talent, and the case in which talent is unidimensional is a limiting case as $\rho \to 1$. Throughout it is useful to work with $\theta = \tilde{\theta}/(1-\rho)$ rather than $\tilde{\theta}$.

Innate skills are combined with schooling in the location of origin to become human capital. Location $d$ human capital for individual $i$ born in location $o$ is given by

$$h_{ido} = s_{id} q_o.$$

Throughout, we refer to $q_o$ as the quality of schooling in $o$, but it likely reflects a broader set of factors that contribute to human capital. The wage per effective unit of labor in destination $d$ for someone from origin $o$ is given by $w_d \epsilon_{do}^{w}$ where $w_d$ is destination $d$ productivity, and $\epsilon_d^{w}$ is a mean 1 log normally distributed error which captures any reason

[23] Barring strong evidence of selection and heritability of skills, the assumption that inherent skill does not depend on location of birth seems a reasonable starting point.

why people from origin $o$ may be more productive in destination $d$. The wage for individual $i$ form origin $o$ is therefore

$$\text{wage}_{ido} = w_d \epsilon_{do} h_{ido} = w_d \epsilon_{do} s_{id} q_o.$$

Utility for individual $i$ from origin $o$ living in destination $d$ is given by

$$U_{ido} = \alpha_d \epsilon^{\alpha}_{do} (1 - \tau_{do}) w_d \epsilon^{w}_{do} s_{id} q_o \equiv \bar{w}_{do} s_{id}. \tag{3}$$

The term $w_d \epsilon_{do} q_o s_{id}$ captures consumption, which is equal to the wage. $\alpha_d$ measures the amenity of location $d$ and captures the need for compensating differentials. Moving to a location with half the amenity level would be compensated by a doubling of earnings. Amenities could include natural beauty, the availability of services, or rental rates.[24] The term $\epsilon^{\alpha}_{do}$ is assumed to be mean zero and log-normally distributed; it captures differences in amenity that depend on location of origin.[25] $\tau_{do}$ captures the utility cost of living away from home (the origin $o$), and we refer to it as a moving cost. We assume that $\tau_{oo} = 0$ so that moving away from home to a destination $d$ would require an individual to be compensated with $(1 - \tau_{do})$ times the income. For example, compared to consumption at the origin $o$, the same level of consumption at destination $d$ may be less pleasurable as it is not undertaken with family and friends. We assume throughout that movement costs are symmetric, so that $\tau_{do} = \tau_{od}$. With this background, known results regarding the Fréchet distribution imply the following results.

First, let $\pi_{do}$ be the portion of people from origin $o$ that choose to work in desination $d$. We have

$$\pi_{do} = \frac{\tilde{w}^{\theta}_{do}}{\sum_{j=1}^{N} \tilde{w}^{\theta}_{jo}} \tag{4}$$

where $\tilde{w}_{do} = w_d \epsilon^{w}_{do} \alpha_d \epsilon^{\alpha}_{do} (1 - \tau_{do})$. Here $\tilde{w}_{do}$ measures the attractiveness of location $d$

[24] Much work in the tradition of Rosen (1979) and Roback (1982) separate out rents from other amenities. This is important when considering welfare, because rents are transfers from one person to another which cannot be considered a loss. Given our emphasis on productivity the difference between rents and amenities is not important.

[25] The error terms $\epsilon^{\alpha}_{do}$ and $\epsilon^{w}_{do}$ mean the model does not perfectly fit the data, and allow us to speak meaningfully about endogeneity issues.

for someone from $o$. Equation (4) is the key sorting equation, and it asserts that sorting depends on relative returns, relative amenities and relative movement costs; it does not depend on the quality of human capital formation in the origin, $q_o$. That sorting does not depend on $q_o$ is key to our exercise: we wish to distinguish between human capital or schooling effects which lead to higher production and human capital effects which are a barrier to migration. Barriers to migration coming from differences in human capital are, to the extent they are symmetric, captured in $\tau_{do}$.[26] To the extent that human capital differences are a barrier to migration but are not symmetric, they will be captured in $\epsilon_{do}$ and will not form part of our counterfactuals.

Second, we can use this characterization to determine the average skill of workers from $o$ working in $d$ by noting that

$$E(s_d \mid choose\ d) = \pi_{do}^{-\frac{1}{\theta}} \bar{\Gamma}, \tag{5}$$

where $\bar{\Gamma} = \Gamma\left(1 - \frac{1}{\theta(1-\rho)}\right)$ and $\Gamma(\cdot)$ is the Gamma function. This equation implies that the more people from $o$ that move to $d$, the lower is their average skill. This is intuitive as it implies that there is less selection: the marginal migrant is drawn from further down the left tail of the talent distribution. Finally, we can work out the average wage in a particular location for people from a given origin:

$$\overline{wage}_{do} = w_d \epsilon_{do} q_o E(s_d \mid choose\ \ d) = w_d \epsilon_{do} q_o \pi_{do}^{-\frac{1}{\theta}} \bar{\Gamma}. \tag{6}$$

Equations (4) and (6) are our main estimating equations. Taking logs of these two equations also shows that the model is consistent with the give reduced form facts discussed earlier. Fact 1, gravity, is an estimate of equation (4) where distance is substituted for moving cost. Fact 2 comes from (6) with $\pi_{do}$ substituted from equation (4). Facts 3 - 5 come directly from (6).

One important implication of our modeling choices is worth noting. When we observe large average wages gaps between locations or sectors, it is tempting to think that there will be large productivity gains to moving people. Our model highlights *two* reasons why

[26] See, for example, Bazzi et al. (2016) for evidence of this kind of human capital differences.

this may not be the case. First, the gaps may reflect selection, as in Young (2013). Second those in low productivity locations may simply have low human capital in total, captured by low $q_o$ in our model. In our empirical work we will non-parametrically estimate $q_o$, allowing for unobservable heterogeneity in the quality of human capital production.

### 3.2 Production and General Equilibrium

A large number of firms in each location produce a non-differentiated product according to a linear production technology. Profits for firm $j$ in location $d$ are given by

$$\Pi_{jd} = p_d A_d h_{jd} - w_{jd} h_{jd}$$

where $A_d$ is labor productivity in location $d$, $p_d$ is the price, which firms take as given, $w_{jd}$ is the wage paid by firm $j$, and $h_{jd}$ is the total amount of human capital employed by firm $j$. Firms compete for laborers by setting wages $w_{jd}$, which implies that in equilibrium $w_{jd} = w_d$ and $\Pi_{jd} = 0 \; \forall j$ and so

$$w_d = p_d A_d.$$

Total economy wide production is given by the CES aggregate

$$Y = \left( \sum_{d=1}^{N} y_d^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where $y_d$ is the total production in location $d$, and $\sigma$ captures the degree of substitutability between products produced by different locations. [27]

Output of good $d$ depends on the amount of human capital in location $d$ according to the function

$$y_d = A_d H_d$$

[27]If $\sigma \to \infty$ all products are perfect substitutes, so the case in which all locations produce the same good is a limit case of our model. An alternative specification would be to allow for locations to produce goods that are perfectly substitutable with a decreasing returns to scale production function. Hsieh and Moretti (2017) show that the two approaches are isomorphic.

where $H_d$ is the total human capital (or effective labor units) available at location $d$ and

$$A_d = \bar{A}_d H_d^{\gamma}$$

is the productivity of location $d$. In this formulation, $\bar{A}_d$ can be thought of as intrinsic productivity – an exogenous parameter, which may change over time. For example, New York may presently have high productivity due to its proximity to a port, but this may have been even more important 100 years ago. Current labor productivity, $A_d$ depends on intrinsic productivity and the total amount of human capital in location $d$ with $\gamma$ parameterizing the extent of human capital spillovers, or productive agglomeration externalities.

Finally, amenity is also endogenously determined. We assume

$$\alpha_d = \overline{\alpha}_d \hat{L}_d^{\lambda}$$

where $\overline{\alpha}_d$ is baseline amenity; for example, natural beauty, $\lambda$ is a measure of congestion effects and likely to be less than zero, and $\hat{L}_d$ is the (endogenously determined) population of location $d$.

Before turning to identification, we note one key characteristic of the model. Dividing through (4) and (6) it is easy to show

$$\frac{\overline{wage}_{do}}{\overline{wage}_{d'o}} = \left(\frac{\alpha_{d'}}{\alpha_d}\right)\left(\frac{1-\tau_{d'o}}{1-\tau_{do}}\right).$$

Hence, within origin, there are no wage gaps without frictions. We, therefore, rule out the kind of behavior discussed in Young (2013), where selection alone drives wage gaps. Our model is somewhere between the work of Young (2013), in which selection is the sole driver of average wage differences, and the work of Gollin et al. (2014), where raw wage gaps are used to infer potential gains from movement.

# 4 Identification and Estimation

In this section, we discuss how we identify and estimate the exogenous parameters of the model $\{\theta,\ \rho,\ q_o,\ w_d,\ \alpha_d,\ \tau_{do}\}$. We also note that we do not need to take a stand on the general equilibrium parameters ($\gamma$, $\lambda$ and $\sigma$) for identification. We make several normalizations. First, as noted above, we assume that $\tau_{oo} = 0$ and $\tau_{do} = \tau_{od}$: movement costs are symmetric, and it is costless to live at home. Second, we normalize $\alpha_1 = 1$: because we do not observe utility levels, the only variation we have to identify $\alpha$ comes from people's relative preferences for locations. Third, we normalize $q_1 = 1$: we identify only relative qualities of human capital generation. This normalizes productivity $w_d$ as well: the wage $w_d$ is what would be earned by someone living at location $d$ who was born in location 1 and who has a skill draw of 1. This means that any on average improvements in human capital generation would be captured in productivities, $w$, and changes in $q$ capture changes in the spatial allocation of human capital production possibilities.

## 4.1 Identification of model parameters

### 4.1.1 Frechet parameters: $\{\theta, \rho\}$

Taking the log of equation (6), we have

$$\ln(\overline{wage}_{do}) = \underbrace{\ln(\overline{\Gamma}) + \ln(w_d)}_{\text{Destination fixed effect}} - \frac{1}{\theta}\ln(\pi_{do}) + \underbrace{\ln(q_o)}_{\text{Origin fixed effect}} + \ln\epsilon^{w}_{do}. \tag{7}$$

That is, after controlling for origin and destination fixed effects, the elasticity of the average wage with respect to the proportion of migrants identifies the Fréchet parameter $\theta$. Intuitively, if people are very similar (or destinations differ little in their skill needs), then $\theta$ is high, so the marginal migrant is not greatly less skilled than the previous migrant, and average wage will change little with movement. However, if dispersion in talent is large (or there are large differences in the skill needs in different destinations), then the marginal migrant is much less skilled than the previous migrant, and so their wage is significantly lower.

Inspection of equation (4) shows that the error term $\epsilon^{w}_{do}$ also enters the definition of $\pi_{do}$. This is intuitive; any random variation that means wages for those from origin $o$ are relatively high in destination $d$ will encourage migration between the two locations. This correlation between the error term and the regressor $\pi_{do}$ creates an endogeneity problem that will lead us to underestimate the extent of selection by overestimating $\theta$. We address this concern with an instrumental variables strategy motivated by our model.

We wish to isolate the variation in $\pi_{do}$ that is driven by variation in the relative amenity of $d$, and productivity in other locations $\neg d$. The proportion of people from other origins $\neg o$ that migrate to destination $d$ is affected by these factors, but not by the random error $\epsilon_{do}$. The set of migration proportions $\{\pi_{d\neg o}\}$ are therefore valid instruments for $\pi_{do}$, although the first stage relationship between $\ln \pi_{d\neg o}$ and $\ln \pi_{do}$ is non-linear. Therefore, we follow the advice of Angrist and Pischke (2009) and instrument $\ln \pi_{do}$ with the fitted value from a "zero stage" regression in which $\ln \pi_{do}$ is regressed on a polynomial in $\ln \pi_{d\neg o}$. Monte Carlo estimates based on a roughly calibrated version of our model confirm that this strategy leads to unbiased estimates and suggests that there are few efficiency gains to increasing the polynomial beyond a quadratic.[28]

To separate comparative and absolute advantage, we use a property of the Fréchet distribution which implies:

$$\frac{\mathrm{var}(w_{do})}{(\overline{wage}_{do}))^2} = \frac{\Gamma\left(1 - \frac{2}{\theta(1-\rho)}\right)}{\left(\Gamma\left(1 - \frac{1}{\theta(1-\rho)}\right)\right)^2} - 1. \tag{8}$$

Using data on the distribution of wages, combined with the $\theta$ identified as above, this equation identifies $\rho$, the parameter defining absolute advantage. Intuitively, if there is little correlation in skill types, so that everyone has some destination in which they excel, then the within destination origin pair wage variance will be low. If, in contrast, $\rho$ is high, then people of many different skill levels will find the same location to be best and so the variance in observed wages will be high relative to the mean.

[28] Results of the Monte Carlo simulations are available from the authors on request.

#### 4.1.2 Location Characteristic Affecting the Wage: $\{w_d, q_o\}$

Considering again equation (7), with the estimates of $\rho$ and $\theta$ in hand, we can identify $w_d$ from the destination fixed effect by noting $\bar{\Gamma} = \Gamma\left(1 - \frac{1}{\theta(1-\rho)}\right)$ which is identified. We identify $w_d$ in levels using the normalization that $q_1 = 1$. Intuitively, after controlling for selection through $\pi_{do}$ and the quality of human capital through $q_o$ any differences in wages between locations must be driven by differences in productivity. The quality of the human capital environment $q_o$ can be similarly determined. After controlling for productivity differences at the destination as well as selection, any differences in wages earned by people from different origins must be accounted for by the relative quality of human capital formation opportunities.

#### 4.1.3 Characteristics Affecting Movement: $\{\tau_{do}, \alpha_d\}$

Taking the log of (4) gives a gravity equation

$$\ln(\pi_{do}) = \theta \ln(w_d) + \theta \ln(\alpha_d) + \theta \ln(1 - \tau_{do}) - \underbrace{\ln\left(\sum_j \tilde{w}_{jo}^{\theta}\right)}_{\text{origin fixed effect}} + \theta(\ln \epsilon^w_{do} + \ln \epsilon^\alpha_{do}). \quad (9)$$

This equation allows us to identify movement costs. Intuitively, low movement could be caused by amenity difference, productivity differences, or movement costs. Among these, movement costs are the only force that leads both people from $o$ to be unlikely to move to $d$ *and* people from $d$ to be unlikely to move to $o$. This intuition is confirmed by rearranging the gravity equation to give:

$$(\ln \pi_{do} - \ln \pi_{oo}) + (\ln \pi_{od} - \ln \pi_{dd}) = 2\theta \ln(1 - \tau_{do}) + \eta_{do},$$

where $\eta_{do}$ is a zero mean shock specific to the locations $d$ and $o$.[29] We see that movement costs are high when people tend to stay at home, and given an estimate of $\theta$ (identified as above), we can use differences in movement relative to staying at home to identify $\tau_{do}$.

The gravity equation also allows for identification of relative amenities. The multi-

[29] $\eta_{do} = \theta(\ln \epsilon^w_{do} + \ln \epsilon^\alpha_{do} - \ln \epsilon^w_{oo} - \ln \epsilon^\alpha_{oo} + \ln \epsilon^w_{od} + \ln \epsilon^\alpha_{od} - \ln \epsilon^\alpha_{dd} - \ln \epsilon^w_{dd})$.

lateral resistance term, $\ln\left(\sum_j \tilde{w}_{jo}^{\theta}\right)$, is correlated with the error, but can be removed by differencing the equation. Given this, and having identified $w_d$, $\theta$, and $\tau_{do}$, the only unknown in (9) is $\alpha_d$. Amenities are things that lead to skewed movement in a particular direction, but which do not cause changes in wages, after controlling for selection. We can only identify amenities up to a normalization because of the origin fixed effect in the equation.

## 4.2 Estimation

We estimate the model using Poisson pseudo-maximum likelihood (PPML). The PPML model has several advantages for estimating migration flows. First, because it estimates the level of migration, rather than the log, it can rationalize zero observed migration flows between locations. This is important because in our context, as in most studies of migration and trade flows, zero observed flows are common (Silva and Tenreyro, 2006). Second, the PPML model respects the general equilibrium adding-up constraints implicit in the model (Fally, 2015).

Our two estimating equations, Equations (7) and (9), are:

$$\ln(\overline{wage}_{do}) = \ln(\overline{\Gamma}) + \ln(w_d) - \frac{1}{\theta}\ln(\pi_{do}) + \ln(q_o) + \ln \epsilon_{do}^{w} \tag{7'}$$

$$\ln(\pi_{do}) = \theta \ln(w_d) + \theta \ln(\alpha_d) + \theta \ln(1-\tau_{do}) - \ln\left(\sum_j \tilde{w}_{jo}^{\theta}\right) + \theta(\ln \epsilon_{do}^{\alpha} + \ln \epsilon_{do}^{w}). \tag{9'}$$

The identification assumption to estimate Equations (7′) and (9′) by PPML is that the (level) error terms are mean one and are uncorrelated with the (exponentiated) regressors. As discussed above, we assume that the errors are mean one, and we deal with correlation with the regressors through IV and differencing strategies.

We proceed as follows. We first employ an IV procedure to estimate $\theta$. We then take this estimate of $\theta$ and estimate the system of three equations (Equations 7′, 9′ and 8) using GMM. In implementing the procedure, we drop observations with less than five observed migrants from the wage data. Although our estimation method rationalizes the presence of zero observed migration between any two locations, we are concerned about small sample sizes affecting the precision of wage estimates. We bootstrap this entire procedure to generate standard errors for our estimated values of $\theta$ and $\rho$.

# 5 Estimation Results

This section presents our parameter estimates. Our main goal is to show that our structurally estimated parameters correlate with proxy measures, and so they appear to measure something real. We show estimates for both Indonesia and the US. We use our US model to estimate US level movement costs to generate a counterfactual for a high-mobility economy. It, therefore, matters whether or not our model does a good job of estimating these parameters in the US. Our preferred estimates of migration cost use no structure other than symmetry. We show that this nonparametric estimate correlates with observable characteristics such as distance. As a robustness check, we re-estimate the model and repeat all the counterfactual analysis, imposing that migration costs are a linear function of (log) distance; we report these estimates in the Appendix Table 6.[30] Recall that for these estimates, we do not need to take a stand on the value of the GE parameters $\{\sigma, \gamma, \lambda\}$.

## 5.1 Fréchet Parameters

Table 3 presents estimates of the distributional parameters for both Indonesia and the US. The skill distribution is summarized by the value $\theta$, which combines both the dispersion and correlation factor. A higher value of $\theta$ means that there is less scope for comparative advantage, either because skill is less dispersed or because the within-person correlation is higher. We estimate a value of $\theta$ of 28 for the US and 13 for Indonesia.[31] Figure 3 shows random draws from the estimated distributions for Indonesia and the US, where each

[30]We display the distributions of all estimated parameters are displayed in Appendix Figures 2 and 3.

[31]The US has slightly more dispersion of skill, with $\tilde{\theta}$ equal to 2.7 for the US and 3.2 for Indonesia. However, the US has a much higher correlation parameter, 0.9 compared with 0.7. The high correlation factor dominates and leads to a much lower total dispersion of skill.

axis is the productivity level for location 1 or 2. The figure shows that the distribution is more dispersed in Indonesia than the US.

## 5.2 Migration costs

We estimate substantial migration costs. Table 3 reports the mean value of $\tau_{do}$, which is 0.56 for Indonesia and 0.22 for the US. On average, migrants in Indonesia must be compensated with a 56% higher income, while Americans require a 22% gain. In this sense, the US is a high-mobility country according to our estimates.

Migration costs, for both the United States and Indonesia, correlate with distance. Figure 5 plots estimated migration cost $\tau_{do}$ against the (log) of distance. Particularly striking is the much lower correlation between distance and movement costs in the US. The elasticity of cost to distance is 3% in the US, compared to 15% in Indonesia. Several mechanisms are possible causes. It may be that transportation is cheaper in the US. Alternatively, it may be that people in the US are more welcoming of migrants from physically distant communities, or that the US is more culturally homogenous.

Measured movement costs also correlate with measures of social distance. Using census data, we construct indices of religious and linguistic similarity.[32] Figure 4 plots the relationship between these indices and movement costs, after controlling for distance. There is no correlation between migration costs and religion, but migration costs are statistically significantly correlated with linguistic similarity.

## 5.3 Amenities

Estimated amenities correlate with measured amenities. The left panel of Figure 6 shows that estimated amenities are negatively correlated with the (standardized) first principal component of pollution amenities. The right panel shows that estimated amenities corre-

[32] The index is constructed by calculating the probability that a person selected at random from the origin will have the same characteristic (religion or language) as a person selected at random from the destination. For example, if the origin is 50% Hindu and 50% Muslim, and the destination is 100% Hindu, then the religious similarity index would be 0.5. If the destination was also 50% Hindu and 50% Muslim, then the index would also be 0.5.

late positively with the first principal component of health and market access amenities.[33]

### 5.4 Quality of human capital formation

Figure 7 shows that estimate of $q_o$ (educational quality) correlate with average educational attainment. This is to be expected if people choose to receive more schooling where there are higher returns to schooling.

## 6 Counterfactuals

We now turn to the policy question we posed at the start of the paper: would there be productivity gains from removing mobility constraints?

To produce counterfactuals, we need to take a stance on the GE parameters. We set these using estimates from the literature, and then evaluate the robustness of our findings to different choices.

A large literature estimates the agglomeration parameter ($\gamma$). The literature is reviewed in Rosenthal and Strange (2004) and Combes and Gobillon (2015). Recent consensus estimates suggest a $\gamma$ of between 0.01 and 0.02 for the developed world, although some studies (e.g. Greenstone et al. 2010) suggest much higher numbers. Estimates for developing countries are more sparse and suggest a $\gamma$ up to 1. We present our main estimate for $\gamma = 0.05$, but also consider robustness for numbers between 0 and 0.08. We expect that spatial integration will have a greater impact when $\gamma$ is high.

A much smaller literature attempts to estimate the congestion parameter $\lambda$. On one hand, the work in Albouy (2012) could be seen as suggesting that $\lambda = 0$ in the US. In contrast, work by Combes and Gobillon (2015) suggests a $\lambda$ of around $-0.04$. We take 0

[33] Appendix Table 7 correlates the estimated amenities with observed amenities one-by-one. Each entry in the table is the regression coefficient from separate regression of estimated amenities on amenities. As we only have 25 estimated parameters we do not expect individual signs to necessarily be statistically significant, but we note the general pattern in these results: overall, measures of pollution are negatively correlated with amenities; measures of health outbreaks such as malaria, tuberculosis and vomiting and also negatively correlated with amenities, although access to health care facilities seems also be to negatively correlated, village lighting and commercial banks are positively correlated and we see a mixed pattern for natural disasters such as flooding and earthquakes.

as our starting point and consider values between 0 and $-0.08$. As $\lambda$ decreases (and congestion becomes more important), we expect that reducing frictions will have a smaller impact. It will be hard to move people in to productive areas, even if movement costs are low.

Accurate estimates the elasticity of substitution across regions are also hard to obtain. Allen and Arkolakis (2014) use a figure of 8, and we follow them in our main results. We also consider values between 2 and 8. We expect that as $\sigma$ increases, there will be larger benefits to spatial integration: a high elasticity of substitution means that the products from different locations become less substitutable, and so there are larger costs to low production of some goods.

## 6.1 Reducing migration costs

The first policy we consider is removing migration costs. On a practical basis, this might be achieved by a set of policies depending on the second and the results of future research. Examples of policies that exist include migration subsidies (Bryan et al. (2014)), migrant welcome centres, language training, and road building (Morten and Oliveira (2016)). To estimate possible impacts, we scale our estimated costs by a reduction factor $\kappa$, yielding $\widetilde{(1-\tau)} = (1-\tau)^{1-\kappa}$, with $\kappa \in [0,1]$. When $\kappa = 0$ this corresponds to the baseline case we estimated. When $\kappa = 1$ this corresponds to removing migration costs entirely.[34]

We find modest gains. We predict an 8% output gain from reducing migration costs to the US level, and a 12% gain from reducing migration costs to zero. The US is usually considered the archetype of a spatially mobile economy, so the 8% figure is probably the maximum attainable. These results are illustrated for a range of values of $\kappa$ in Figure 8.

These modest gains hide substantial heterogeneity across origin populations. While the average increase from eliminating all migration costs is 12%, the effect ranges from -12% to 79%.[35] That is, the people born in some provinces may see a 79% increase in their

[34]The average value of $\tau_{us} = 0.22$ and the average value of $\tau_{ind} = 0.56$, so the policy experiment of lowering migration costs in Indonesia to the US level is equivalent to considering $\kappa = 0.3$.

[35]Recall there is no restriction that reducing migration costs will lead to increases in output. Reducing migration costs may lead people to migrate away from high-productivity-low-amenity locations towards low-productivity-high-amenity ones. This is indeed what we see in these counterfactuals.

average wage $\sum_d \overline{wage}_{do}$. For a move to the US benchmark, the gains range from -5% to 37%. The distribution of gains from complete removal is depicted in Panel A of Figure 9 and the US benchmark is presented in 10.

As noted above, selection plays two roles in our model. On one hand, skill heterogeneity implies that there are gains from sorting. The greater the heterogeneity, the greater the return to sorting. On the other hand, if each additional migrant earns less than the last, selection will strongly reduce predicted gains from agglomeration. These two opposing mechanisms mean that ignoring selection could lead to us to either over – or under – estimate policy gains. To understand the importance of selection, we recompute productivity changes, shutting down the selection margin.[36] Sorting is the main source of output gains from removing migration costs. Column 1 in Table 4 shows that all estimated gains come from improving worker sorting (we estimate a 12% gain with sorting compared to a 3% loss without sorting). Ignoring selection would lead us to underestimate the gains from removing movement costs.

## 6.2 Reducing amenity dispersion

The second counterfactual we consider reduces the dispersion of amenities. Again, this corresponds to the aggregate impacts of a set of possible policies. For example, encouraging home building in high-demand locations (Harari 2016 and Hsieh and Moretti 2017), reducing pollution in cities and providing equal access to schooling and hospitals. Amenities are estimated to scale. As with movement costs, we rescale amenities by a reduction factor $\kappa$, yielding $\frac{\widetilde{\alpha_i}}{\alpha_1} = \left(\frac{\alpha_i}{\alpha_1}\right)^{1-\kappa}$, with $\kappa \in [0, 1]$. When $\kappa = 0$ this corresponds to the baseline case we estimated. When $\kappa = 1$ this corresponds to equalizing amenities across all locations.[37]

Here we do not compute a US benchmark, this is for two reasons. First, we believe that it is plausible to have zero amenity differentials; there is no obvious reason why

[36] We do this by setting the endogenous component of human capital equal to 1. This maps to a model where people are migrating based on preference shocks, such as is considered in Allen and Arkolakis (2014); Redding (2016).

[37] We find little difference in the underlying distribution of amenities between Indonesia and the US, as shown in Appendix Figures 2 and 3.

some locations have to have fewer services and more pollution. Second, inline with the general argument in Hsieh and Moretti (2017), we find that the US has greater amenity dispersion than Indonesia. Hsieh and Moretti argue that that restrictive housing policies lead to high rents in some very productive locations; this would show up in our estimates as high amenity dispersion. We find that equalizing amenities would lead to an increase in output of 12.6%. These gains are illustrated in panel B of Figure 8. As with migration costs, we find substantial heterogeneity. Some origin-locations receive wage gains of up to 88%.

As above, we ask how these results are affected by selection. We find, in Column 2 of Table 4 that, in contrast to migration costs, removing the selection margin has very little effect on predicted gains. As noted above, by ignoring selection, we overestimate the gains from agglomeration.

## 6.3 Reducing both migration costs and amenities differentials

Finally, we consider eliminating both barriers – migration costs and compensating differentials – simultaneously. Doing so leads to a 21.5% output gain. The effect of reducing both barriers is slightly smaller than the sum of their independent effects, suggesting the policies are very mild substitutes. Under the policy of reducing all barriers to mobility, the origin that benefits the most would face wage increases of 104%. For this combined policy, accounting for selection is also important. Table 4 shows that if we do not account for selection, we understate gains by 40%.

## 6.4 GE effects

The main results use our baseline parameters for the agglomeration, congestion, and substitution. We undertake robustness over these parameters. Results are reported in Appendix Tables 8 through 10. As expected, when agglomeration is high, congestion forces are low, and the elasticity of substitution is high, the gains to removing barriers to mobility increase. For the experiment of reducing both migration costs and amenities, our baseline estimate was an increase in output of 21.5%. The range of results in Appendix

Table 10 is from 13.4% to 23.8%.

# 7 Conclusion

Large spatial wage gaps and recent experimental evidence suggest there may be important productivity gains from encouraging internal migration in developing countries. We estimate the size of the aggregate gains in Indonesia. Our approach entails using movement data to identify constraints on migration, and considering how removing these constraints would affect locational choices and wages, taking in to account selection and GE effects. Aggregate output gains are small but important, on the order of 20%. These estimates hide a great deal of heterogeneity, with some more constrained areas seeing gains of over 100%. Failure to account for selection would lead to an underestimate of the gains; accounting for selection both reduces estimated gains to agglomerating workers in one location, and allows for larger gains through improved sorting. We find that the latter effect dominates.

Future research could aim to deepen our understanding of the mechanisms through which migration affects productivity. Theoretical and macroeconomic research could concentrate on the dynamic effects of encouraging migration. Microeconomic experimental evidence on the extent and nature of selection among internal migrants, as well as the strength of comparative advantage effects, would also add to our understanding. Experimental research along these lines is currently taking place as part of the broad research agenda motivated by Bryan et al. (2014) and related work, including this project.

## References


Adao, Rodrigo, "Worker Heterogeneity, Wage Inequality, and International Trade: Theory and Evidence from Brazil," 2016.

Albouy, David, "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas," 2012.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

Angrist, J and Jorn-Steffen Pischke, *Mostly Harmless Econometrics*, Princeton University Press, 2009.

Bazzi, Samuel, Arya Gaduh, Alexander Rothenberg, and Maisy Wong, "Skill Transferability , Migration , and Development : Evidence from Population Resettlement in Indonesia ," *American Economic Review*, 2016, *106* (9), 2658–2698.

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

Borjas, G J, "Self-selection and the earnings of immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

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

Carrington, W, E Detragiache, and Tara Vishwanath, "Migration with endogenous moving costs," *The American Economic Review*, 1996, *86* (4), 909–930.

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

Combes, Pierre-Philippe and Laurent Gobillon, "The Empirics of Agglomeration Economies," in Gilles Duranton and Will Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 2015, pp. 247–348.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-hansberg, "The Geography of Development," *Journal of Political Economy (forthcoming)*, 2016, pp. 1–54.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Fally, Thibault, "Structural Gravity and Fixed Effects," *Journal of International Economics*, 2015, (September), 1–26.

Garlick, Julia, Murrary Leibbrandt, and James Levinsohn, "Individual Migration and Household Incomes," *NBER Working Paper*, 2016.

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

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010.

Harari, Maria, "Women's Inheritance Rights and Bargaining Power: Evidence from Kenya," 2016.

Harris, John and Michael Todaro, "Migration, Unemployment and Develpment: A Two-Sector Analysis," *The American Economic Review*, 1970, *60* (1), 126–142.

Hicks, Joan Hamory, Marieke Kleemans, Nicholas Li, and Edward Miguel, "Reevaluating Agricultural Productivity Gaps with Longitudional Microdata," 2017.

Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation," 2017.

_ and Peter J Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, pp. 1–60.

_ , Erik Hurst, Charles Jones, and P Klenow, "The Allocation of Talent and US Economic Growth," 2016.

Kennan, J and J Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251.

Kleemans, Marieke and Jeremy Magruder, "Labor markets changes in response to immigration: evidence from internal migration driven by weather shocks," *The Economic Journal (forthcoming)*, 2017.

Lagakos, David and Michael E Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, apr 2013, *103* (2), 948–980.

_ , Mushfiq Mobarak, and Michael E. Waugh, "The Welfare Effects of Encouraging Rural-Urban Migration," 2015.

Lewis, W A, "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 1954, *22* (2), 139–191.

Morten, Melanie, "Temporary Migration and Endogenous Risk Sharing in Village India," 2017.

_ and Jaqueline Oliveira, "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," 2016.

Redding, S J, "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 2016, *101*, 148–167.

Roback, J, "Wages, Rents, and the Quality of Life," *The Journal of Political Economy*, 1982, *90* (6), 1257–1278.

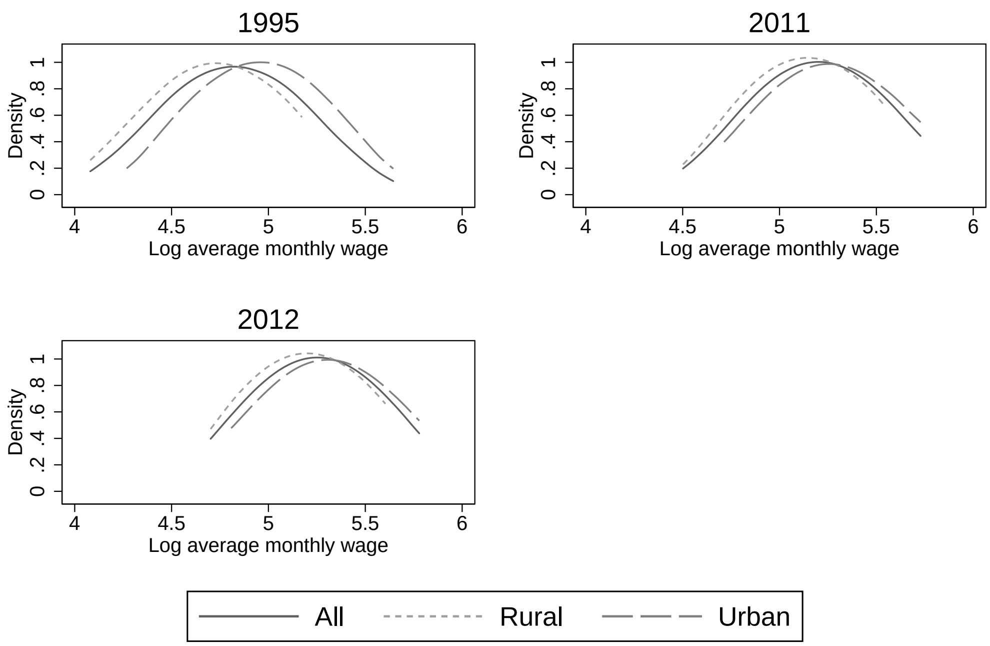
Rosen, Sherwin, "Wage-Based Indexes of Urban Quality of Life," in P Mieszkowski and M Straszheim, eds., *Current Issues in Urban Economics1*, 1979.

Rosenthal, Stuart S. and William C. Strange, "Evidence on the Nature and Sources of Agglomeration Economies," in "Handbook of Regional and Urban Economics," Vol. 4 2004, pp. 2119–2171.

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

Tombe, Trevor and Xiadong Zhu, "Trade Liberalization, Internal Migration and Regional Income Differences: Evidence from China," 2015.

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




Unit of observation is the regency.
Regency is defined as either rural or urban to match the national share of rural.
Source: 1995 SUPAS, 2011 SUSENAS; 2012 SUSENAS.

Figure 1: Distribution of wages: regency level

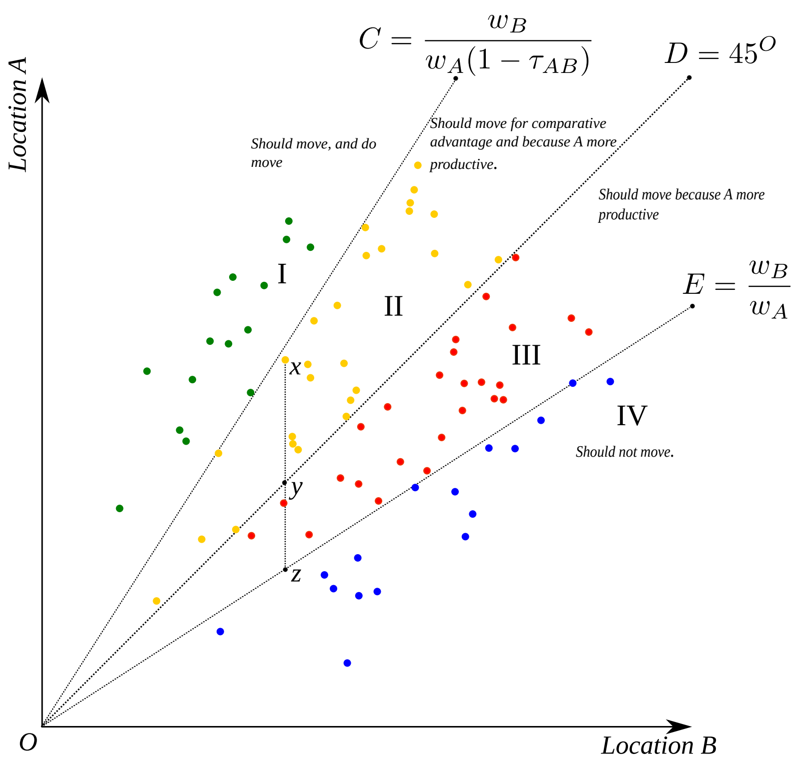



Figure 2: Productivity and Location Choices of People Born in B

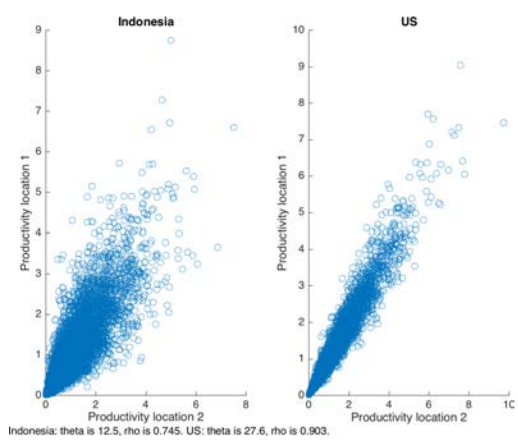



Figure 3: Simulated Frechet Distribution

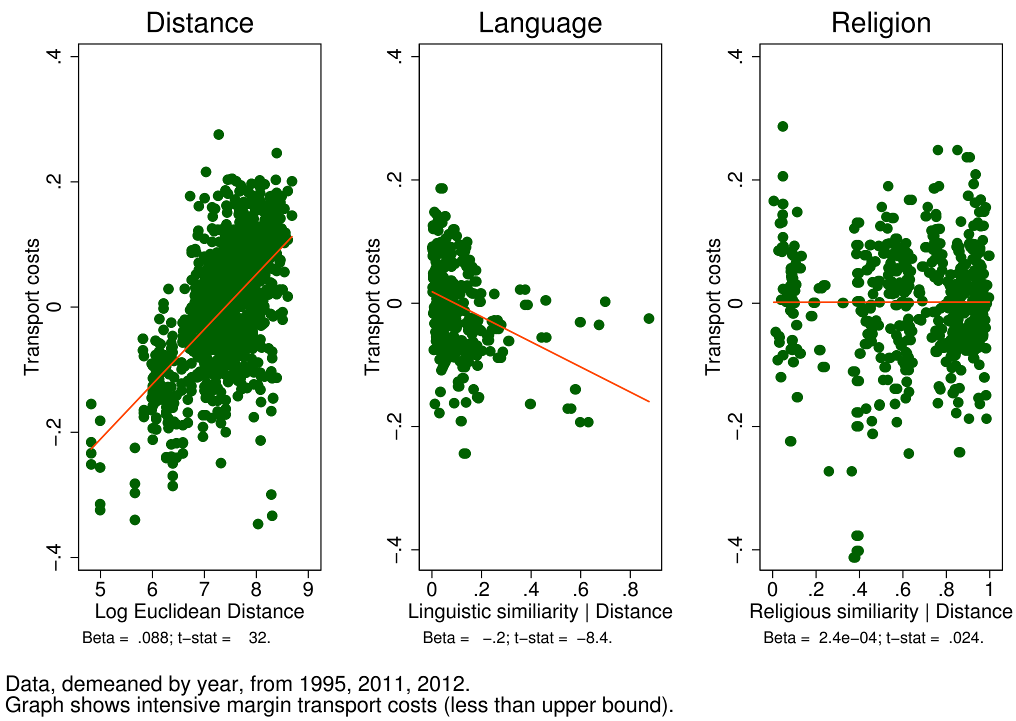



Figure 4: Correlates of iceberg costs in Indonesia

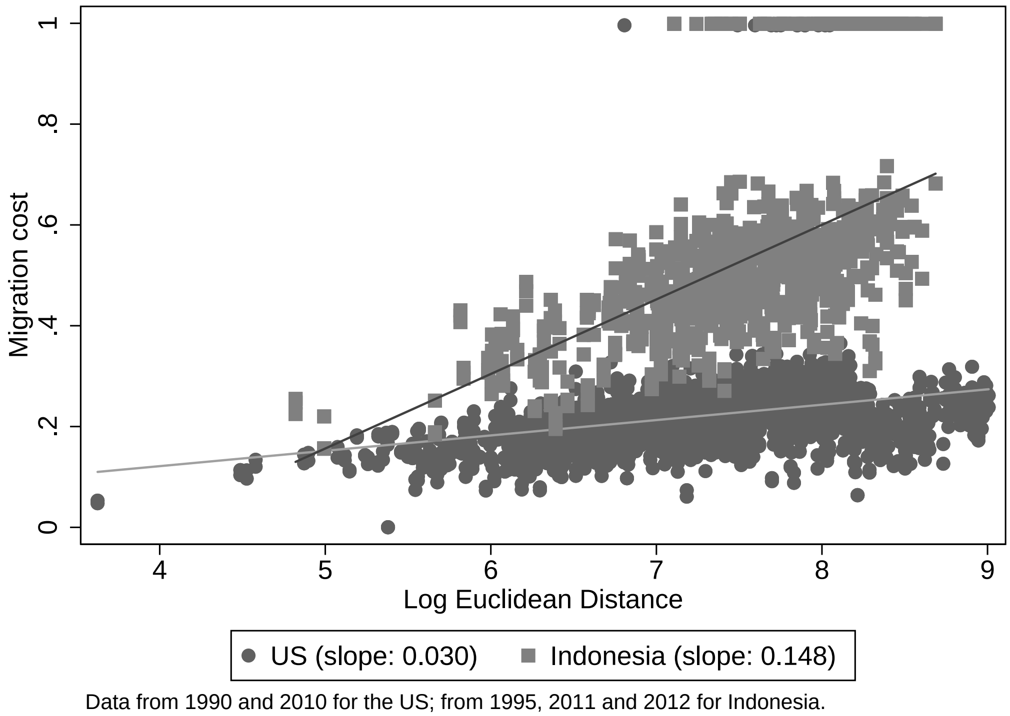



Figure 5: Relationship between iceberg costs and distance in Indonesia and the United States

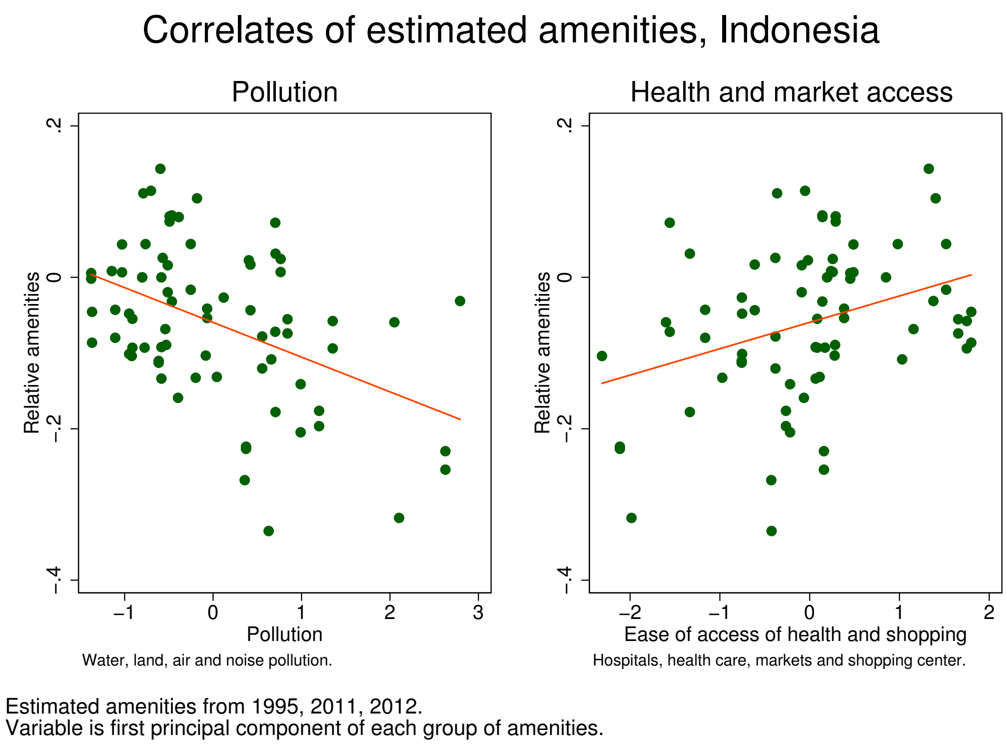
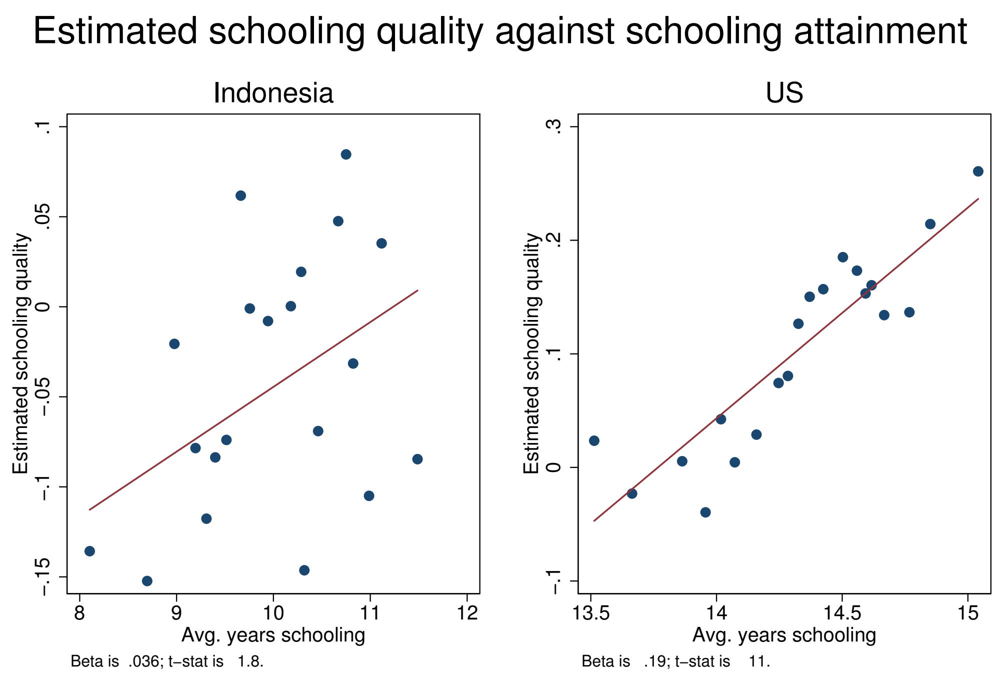
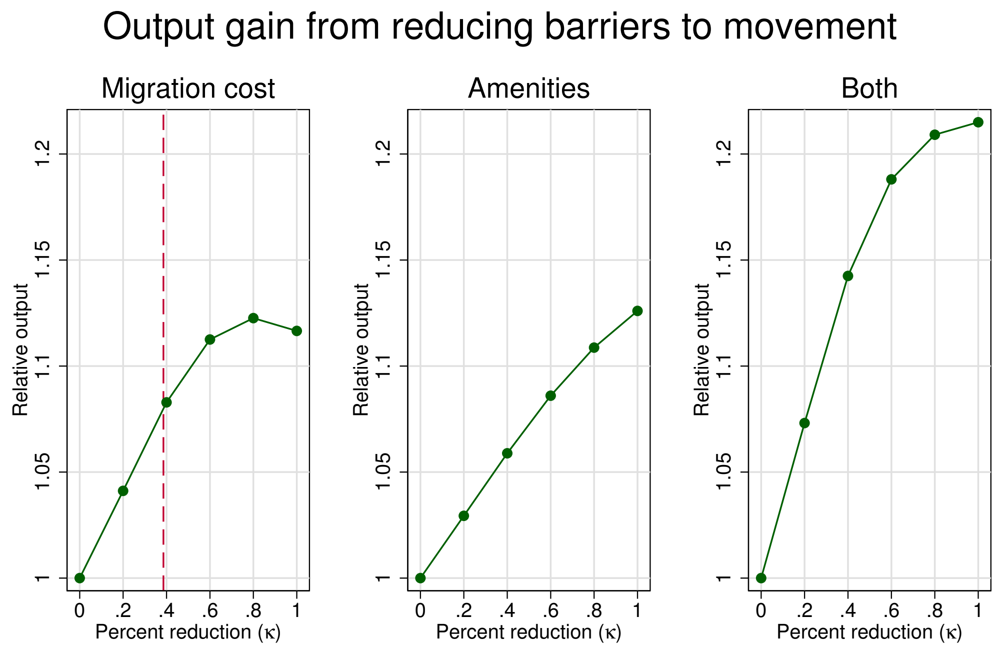



Figure 6: Estimated amenities against measured amenities




Data is: 1990, 2010 for the US. 1995, 2011, 2012 for Indonesia.
Figure shows a binned scatterplot. Year effects are removed from both graphs.

Figure 7: Schooling quality positively correlated with attainment




Data is average across 1995, 2011, 2012 for Indonesia.
The percent reduction, κ, is defined in the text.
The red dashed line shows the US–level of migration costs.

Figure 8: Output gains from reducing barriers to movement

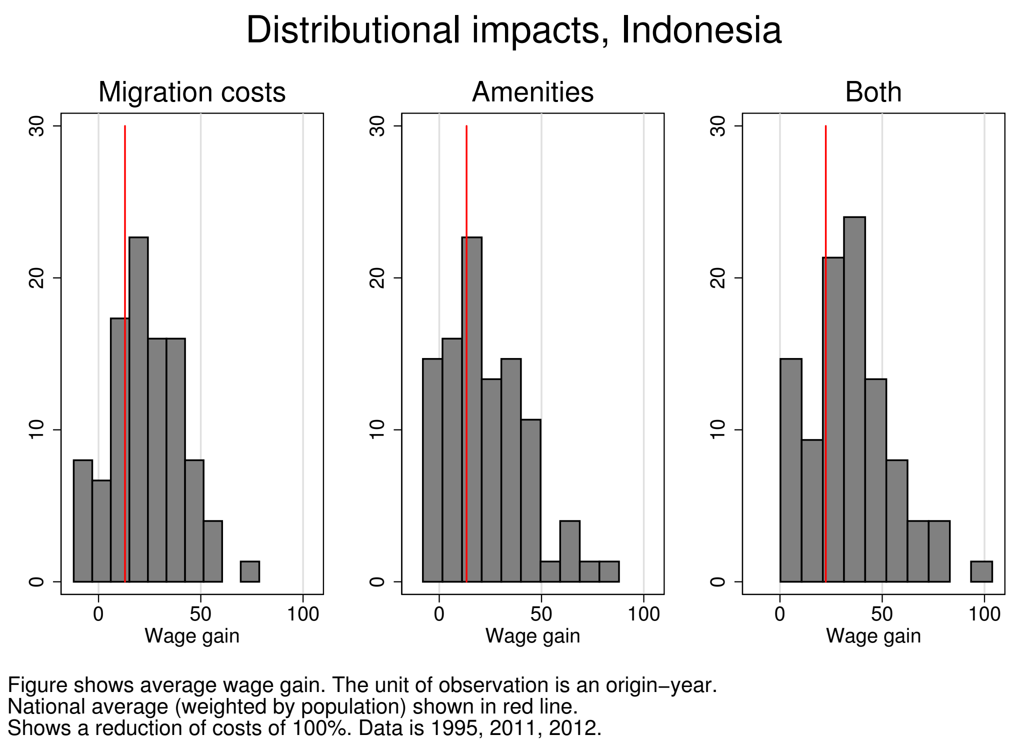



Figure 9: Distributional effects of fully reducing barriers to migration

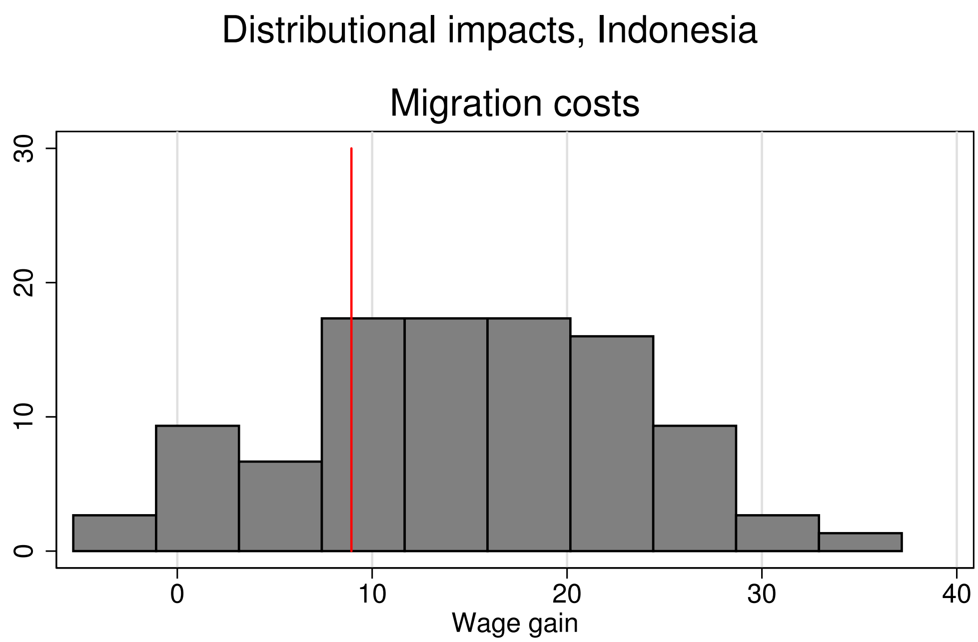



Figure shows average wage gain. The unit of observation is an origin−year.
National average (weighted by population) shown in red line.
Shows a reduction of costs of 40%. Data is 1995, 2011, 2012.

Figure 10: Distributional effects of reducing migration costs to US level

Table 1: Migration rates by origin, Indonesia

| | Rural | Urban | All |
|---|---|---|---|
| *1995* | | | |
| Migration rate | 32.3 | 35.8 | 33.7 |
| Moves within category | 31.1 | 74.6 | 49.4 |
| *2011* | | | |
| Migration rate | 38.7 | 33.7 | 35.7 |
| Moves within category | 24.4 | 84.2 | 58.7 |
| *2012* | | | |
| Migration rate | 38.9 | 34.1 | 35.8 |
| Moves within category | 25.4 | 83.8 | 60.7 |

*Notes:* Data source: 1995 Supas; 2011 Susenas; 2012 Susenas. Migration is measured as living in a regency other than the birth regency. Regencies are classified as rural or urban based on the share of their population that report being rural; we choose the cutoff to classify the regency as rural to match the national urbanization rate for each year.

Table 2: Five facts about migration in Indonesia

| | Movement costs | | Selection | | Compensating Diff. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Dep. variable | $\log \pi_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ |
| Log distance | -0.717 | 0.029 | | 0.007 | |
| | (0.009)*** | (0.001)*** | | (0.002)*** | |
| Log share migrating | | | -0.039 | -0.031 | |
| | | | (0.001)*** | (0.003)*** | |
| Amenities | | | | | -0.023 |
| | | | | | (0.010)** |
| Destination x Year FE | yes | yes | yes | yes | no |
| Origin x Year FE | yes | yes | yes | yes | yes |
| Destination FE | | | | | yes |
| No. of obs. | 25540 | 25244 | 25244 | 25244 | 25050 |

*Notes:* $\log \pi_{odt}$ is the log of the share of population migrating from $o$ to $d$ in year $t$. $\log w_{odt}$ is the log of the average wage of migrants from origin $o$ in destination $d$ in time $t$. An observation is an origin-destination regency pair. Datasource: 1995 SUSENAS, 2011 SUSENAS, 2012 SUSENAS. Number of observations changes between columns because not all pairs with positive migration flows have observed wages. Amenity measure is the standardized value of the first principal component. Two-way clustering of standard errors at the origin-year and destination-year reported in Columns (1)-(4). Clustered standard errors, at the level of the origin-year, reported in Column (5).

Table 3: Estimated Frechet parameters

| | (1) Indonesia | (2) U.S. |
|---|---|---|
| $\rho$ (correlation) | 0.74*** | 0.90*** |
| | (0.031) | (0.017) |
| $\theta$ (dispersion) | 12.5*** | 27.6*** |
| | (1.73) | (3.42) |
| $\tilde{\theta} = \theta(1-\rho)$ | 3.18 | 2.69 |
| Mean migration cost (iceberg) | 0.50 | 0.22 |

*Notes:* Source: estimates from structural estimation of model. Transport costs estimated non-parametrically. Bootstrapped standard errors reported.

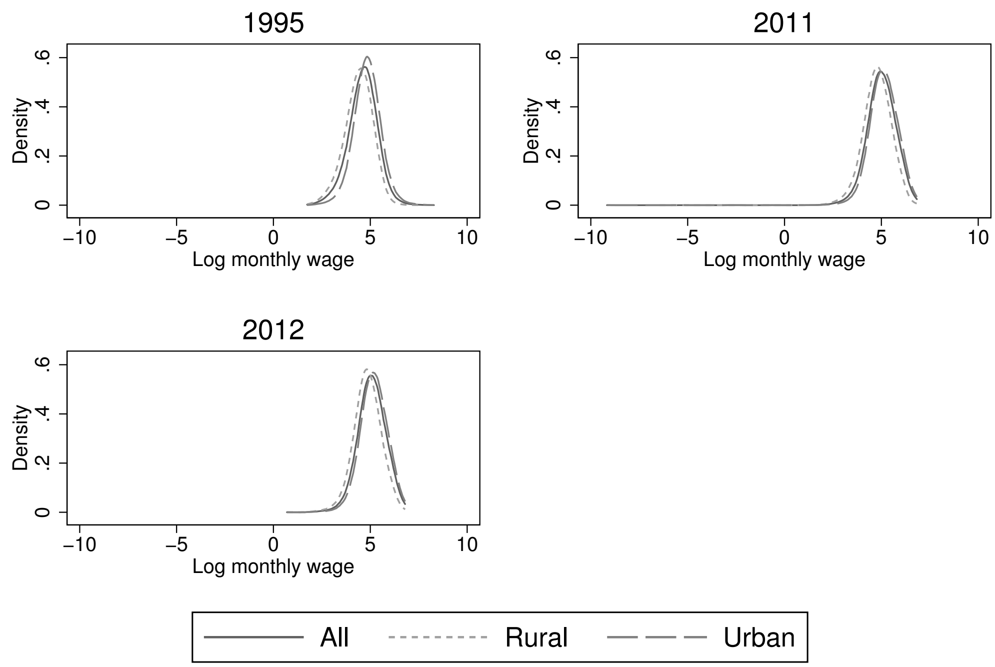
Table 4: Output gain from reducing migration barriers

| | (1) Migration costs | (2) Amenities | (3) Both |
|---|---|---|---|
| *Indonesia* | | | |
| Baseline | 1.117 | 1.126 | 1.215 |
| No selection | 0.970 | 1.125 | 1.132 |

*Notes:* Table shows the output gain from removing the barrier completely. Data is 1995, 2011, 2012 for Indonesia; 1990 and 2010 for the US. No selection recalculates the output gain shutting down the role for comparative advantage.

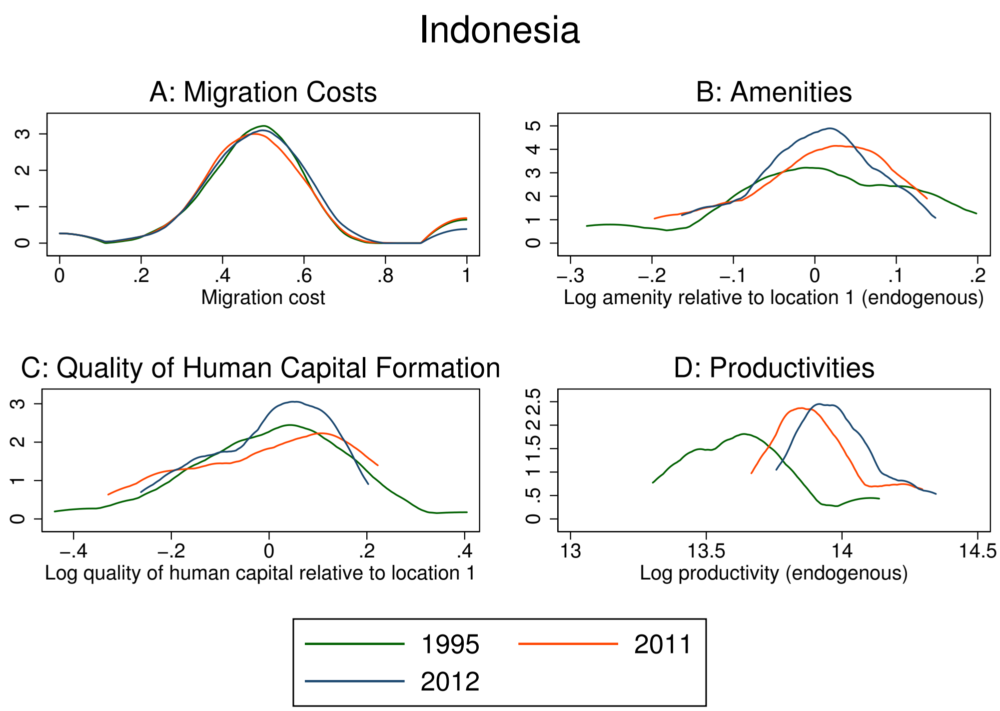
# Appendices

## A Tables and Figures



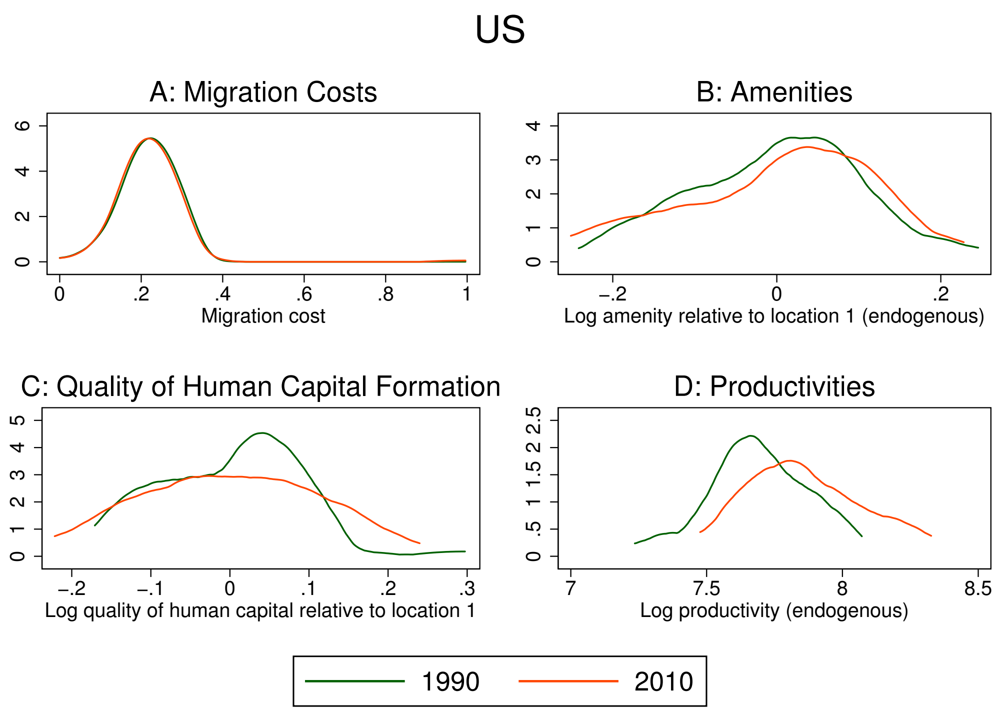

Unit of observation is the individual.
Source: 1995 SUPAS, 2011 SUSENAS; 2012 SUSENAS.

Appendix Figure 1: Distribution of wages: individual level




Amenities and quality of human capital formation have been normalized to have mean zero.

Appendix Figure 2: Distribution of Estimated Structural Parameters, Indonesia




Amenities and quality of human capital formation have been normalized to have mean zero.

Appendix Figure 3: Distribution of Estimated Structural Parameters, US

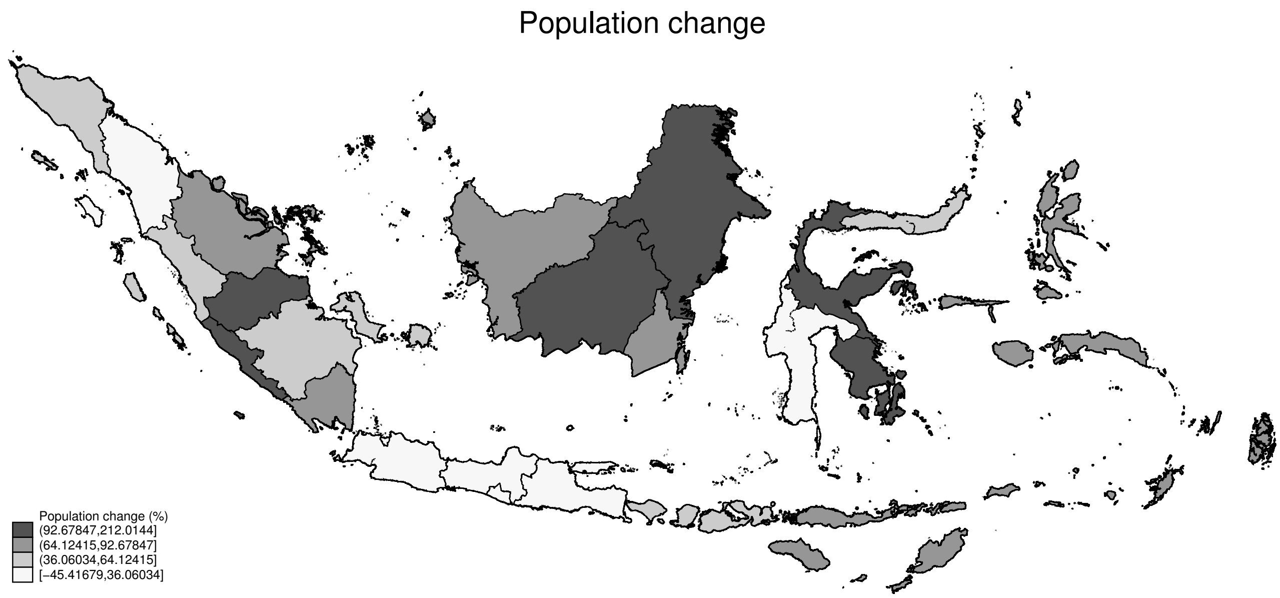



Figure shows population change if migration costs went to zero. Data is 1995


Figure shows wage change experienced by people in an origin if migration costs went to zero. Data is 1995

Appendix Figure 4: Maps of population and wage changes: reduction in migration costs, Ind

Population change

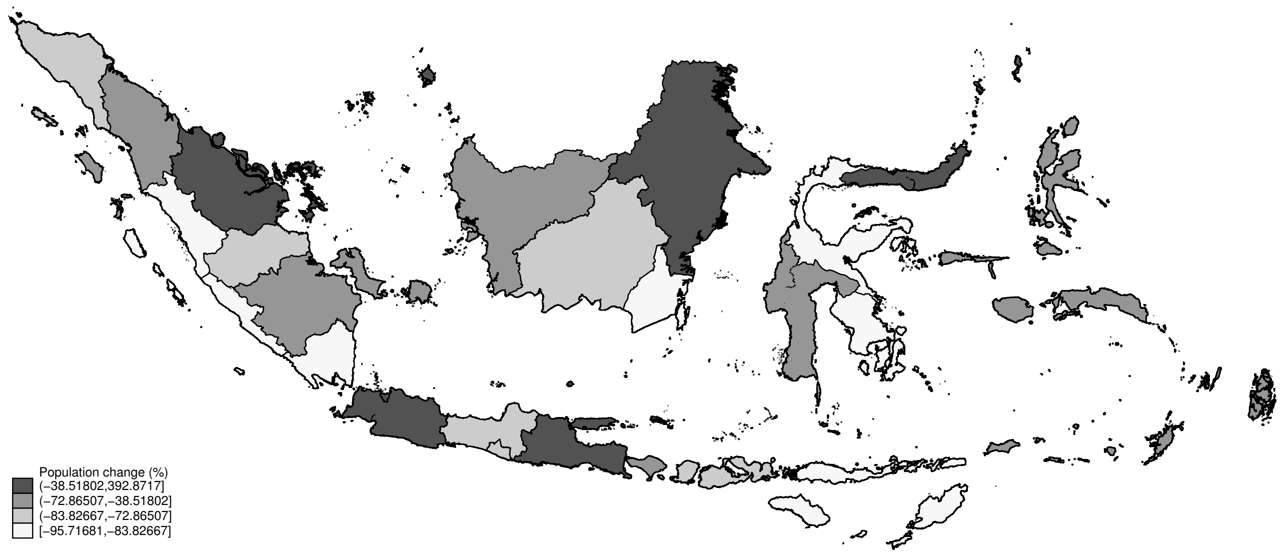
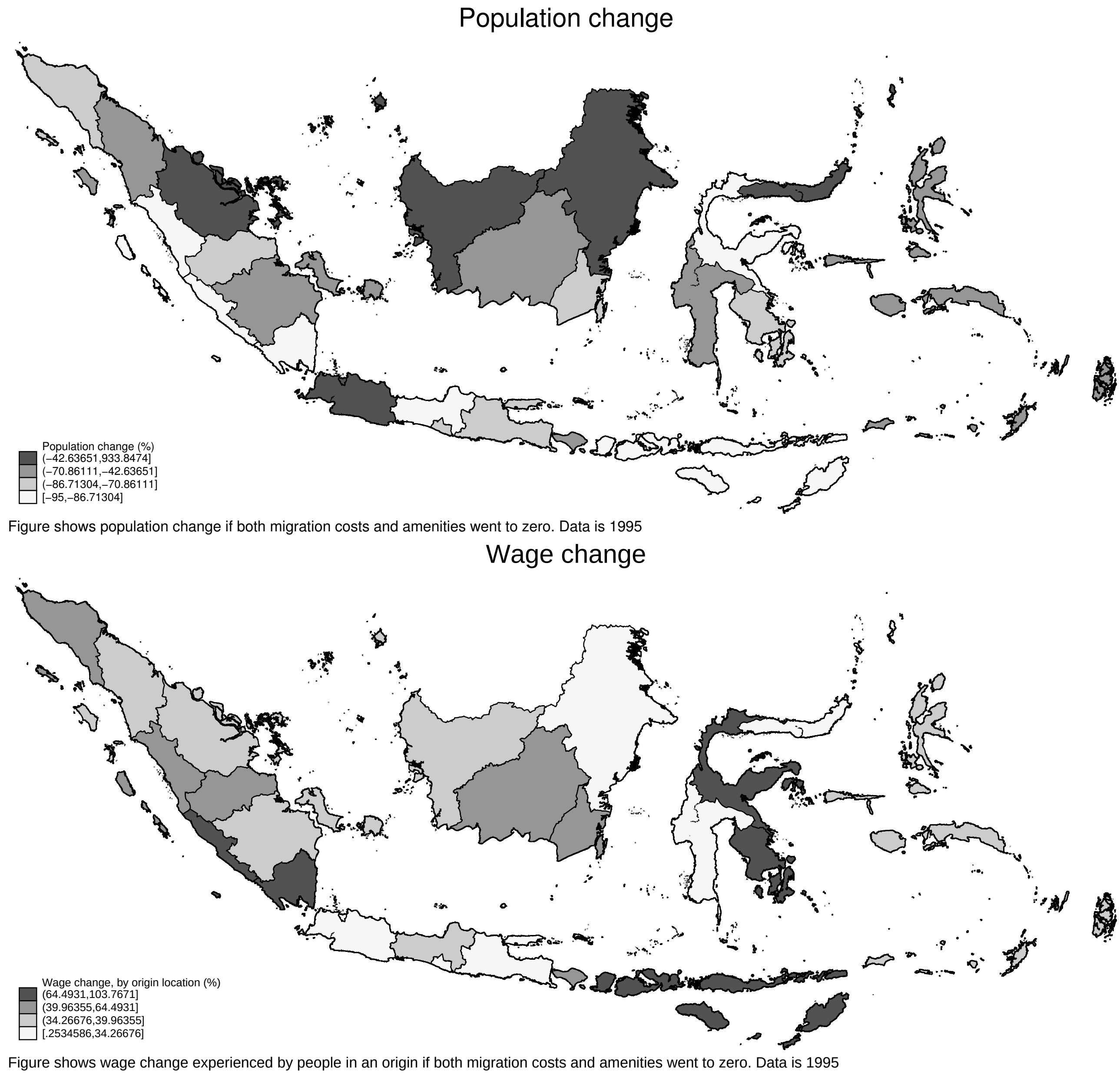



Figure shows population change if amenities went to zero. Data is 1995

Wage change


Figure shows wage change experienced by people in an origin if amenities went to zero. Data is 1995

Appendix Figure 5: Maps of population and wage changes: reduction in amenities, Ind




Appendix Figure 6: Maps of population and wage changes: reduction in both migration costs and amenities, Ind

Appendix Table 1: Summary statistics for Indonesia and US sample

| | Indonesia | | | | United States |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | 1995 | 2011 | 2012 | 1990 | 2010 |
| *Demographic* | | | | | |
| Average age | 38.58 | 39.63 | 39.86 | 40.78 | 44.30 |
| Average age (mig) | 38.43 | 39.31 | 39.71 | 41.26 | 44.95 |
| Years school | 8.01 | 9.85 | 10.12 | 13.48 | 15.24 |
| Years school (mig) | 10.17 | 11.12 | 11.33 | 14.03 | 15.61 |
| High education | 0.82 | 0.92 | 0.93 | 0.85 | 0.93 |
| *Financial* | | | | | |
| Monthly non-zero wage | 129.62 | 188.06 | 198.72 | 4424.07 | 4985.62 |
| Monthly non-zero wage (mig) | 174.82 | 236.25 | 246.36 | 4829.80 | 5667.14 |
| GDP per capita | 1328.48 | 3177.26 | 3337.69 | 39982.11 | 48377.39 |
| *Migration* | | | | | |
| Share migrating | 0.34 | 0.36 | 0.36 | 0.40 | 0.41 |
| Share low educ migrate | 0.15 | 0.19 | 0.19 | 0.31 | 0.30 |
| Share high educ migrate | 0.39 | 0.37 | 0.37 | 0.42 | 0.46 |
| Share migrating (prov) | 0.18 | 0.22 | 0.23 | | |
| Number of obs | 59,006 | 68,510 | 69,330 | 1,958,123 | 336,033 |
| Sum of sample weights | 13,454,136 | 17,324,774 | 17,736,228 | 39,251,152 | 32,404,124 |

*Notes:* Sample is male head of household, between 15 and 65 years old. Migration defined at the regency level in Indonesia, and at the state level in the US. Low education is 3 years of schooling or less in Indonesia; 12 years or less in the US. Data source: Indonesia: 1995 SUPAS, 2011 SUSENAS, 2012 SUSENAS. US: 1990 ACS and 2010 ACS. GDP data from the World Bank Development Indicators Database. All financial values reported in 2010 USD.

Appendix Table 2: Summary statistics for IFLS sample

| | (1) 1993 | (2) 1997 | (3) 2000 | (4) 2007 |
|---|---|---|---|---|
| *Demographic* | | | | |
| Average age | 42.27 | 43.04 | 42.40 | 43.41 |
| Average age (mig) | 41.98 | 42.87 | 42.05 | 43.40 |
| Years school | 5.18 | 5.83 | 7.27 | 7.93 |
| Years school (mig) | 6.42 | 6.92 | 8.69 | 8.74 |
| High education | 0.73 | 0.79 | 0.83 | 0.87 |
| *Financial* | | | | |
| Monthly non-zero wage | 149.58 | 140.18 | 13.86 | 7.32 |
| Monthly non-zero wage (mig) | 203.82 | 182.02 | 13.65 | 6.59 |
| GDP per capita | 1181.78 | 1552.12 | 1662.29 | 2276.16 |
| *Migration* | | | | |
| Share migrating | 0.32 | 0.35 | 0.39 | 0.47 |
| Share low educ migrate | 0.22 | 0.23 | 0.27 | 0.35 |
| Share high educ migrate | 0.36 | 0.38 | 0.41 | 0.49 |
| Share migrating (prov) | 0.14 | 0.16 | 0.24 | 0.23 |
| Number of obs | 5,496 | 5,625 | 7,709 | 9,975 |
| Sum of sample weights | 5,501 | 5,952 | 8,430 | 10,809 |

*Notes:* Sample is male head of household, between 15 and 65 years old. Migration defined at the regency level. Low education is 3 years of schooling or less in Indonesia. Data source: 1993, 1997, 2000 and 2007 IFLS surveys. GDP data from the World Bank Development Indicators Database. All financial values reported in 2010 USD.

Appendix Table 3: Four facts about migration in the U.S.

| | Movement costs | | Selection | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dep. variable | $\log \pi_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ |
| Log distance | -0.553 | 0.023 | | -0.004 |
| | (0.018)*** | (0.002)*** | | (0.004) |
| Log share migrating | | | -0.043 | -0.050 |
| | | | (0.003)*** | (0.006)*** |
| Destination x Year FE | yes | yes | yes | yes |
| Origin x Year FE | yes | yes | yes | yes |
| Destination FE | | | | |
| No. of obs. | 5084 | 5076 | 5076 | 5076 |

*Notes:* $\log \pi_{odt}$ is the log of the share of population migrating from $o$ to $d$ in year $t$. $\log w_{odt}$ is the log of the average wage of migrants from origin $o$ in destination $d$ in time $t$. An observation is an origin-destination state pair. Datasource: 1990 Census, 2010 ACS Number of observations changes between columns because not all pairs with positive migration flows have observed wages. Amenity measure is the standardized value of the first principal component. Two-way clustering of standard errors at the origin-year and destination-year reported in Columns (1)-(4). Clustered standard errors, at the level of the origin-year, reported in Column (5).

Appendix Table 4: Five facts about migration in Indonesia, province-level

| | Movement costs | | Selection | | Compensating Diff. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Dep. variable | $\log \pi_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ |
| Log distance | -0.606 | 0.041 | | 0.019 | |
| | (0.013)*** | (0.003)*** | | (0.009)** | |
| Log share migrating | | | -0.066 | -0.036 | |
| | | | (0.005)*** | (0.012)*** | |
| Amenities | | | | | -0.059 |
| | | | | | (0.074) |
| Destination x Year FE | yes | yes | yes | yes | no |
| Origin x Year FE | yes | yes | yes | yes | yes |
| Destination FE | | | | | yes |
| No. of obs. | 1452 | 1444 | 1444 | 1444 | 1444 |

*Notes:* $\log \pi_{odt}$ is the log of the share of population migrating from $o$ to $d$ in year $t$. $\log w_{odt}$ is the log of the average wage of migrants from origin $o$ in destination $d$ in time $t$. An observation is an origin-destination province pair. Datasource: 1995 SUSENAS, 2011 SUSENAS, 2012 SUSENAS. Number of observations changes between columns because not all pairs with positive migration flows have observed wages. Amenity measure is the standardized value of the first principal component. Two-way clustering of standard errors at the origin-year and destination-year reported in Columns (1)-(4). Clustered standard errors, at the level of the origin-year, reported in Column (5).

Appendix Table 5: Four facts about migration in Indonesia, IFLS data

| | Movement costs | | Selection | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dep. variable | $\log \pi_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ | $\log w_{odt}$ |
| Log distance | -0.571 | 0.023 | | -0.088 |
| | (0.018)*** | (0.013)* | | (0.047)* |
| Log share migrating | | | -0.053 | -0.195 |
| | | | (0.023)** | (0.083)** |
| Destination x Year FE | yes | yes | yes | yes |
| Origin x Year FE | yes | yes | yes | yes |
| Destination FE | | | | |
| No. of obs. | 613 | 228 | 228 | 228 |

*Notes:* $\log \pi_{odt}$ is the log of the share of population migrating from $o$ to $d$ in year $t$. $\log w_{odt}$ is the log of the average wage of migrants from origin $o$ in destination $d$ in time $t$. An observation is an origin-destination province pair. Datasource: 1993, 1997, 2000, 2007 IFLS. Number of observations changes between columns because not all pairs with positive migration flows have observed wages. Amenity measure is the standardized value of the first principal component. Two-way clustering of standard errors at the origin-year and destination-year reported in Columns (1)-(4). Clustered standard errors, at the level of the origin-year, reported in Column (5).

Appendix Table 6: Estimated Frechet parameters (parameterized model)

| | (1) | (2) |
|---|---|---|
| | Indonesia | U.S. |
| $\rho$ (correlation) | 0.74*** | 0.90*** |
| | (0.031) | (0.017) |
| $\theta$ (dispersion) | 12.5*** | 27.6*** |
| | (1.73) | (3.42) |
| $\tilde{\theta} = \theta(1-\rho)$ | 3.18 | 2.69 |
| Mean migration cost (iceberg) | 0.41 | 0.18 |

*Notes:* Source: estimates from structural estimation of model. Transport costs constrained to be a function of log distance. Bootstrapped standard errors reported.

Appendix Table 7: Correlation of estimated amenities with data

| | (1)<br>1995<br>b/se | (2)<br>2011<br>b/se | (3)<br>2012<br>b/se |
|---|---|---|---|
| Water pollution (past year) | -1.33*** | -0.57** | -0.54** |
| | (0.51) | (0.25) | (0.22) |
| Land pollution (past year) | 0.91 | -3.09*** | -2.87*** |
| | (1.63) | (1.03) | (0.89) |
| Air pollution (past year) | -0.45*** | 0.100 | -0.14 |
| | (0.17) | (0.33) | (0.29) |
| Noise pollution (past year) | -1.47 | | |
| | (1.18) | | |
| Main road village lighting | 0.23 | 0.41* | 0.39** |
| | (0.42) | (0.23) | (0.20) |
| Has movie theater | -6.49 | -19.3 | -45.2** |
| | (4.38) | (24.6) | (20.0) |
| Ease of reaching hospital | 0.18** | 0.058 | 0.16*** |
| | (0.080) | (0.078) | (0.062) |
| Ease of reaching puskesmas/other health facility | 0.26* | 0.028 | 0.20** |
| | (0.15) | (0.12) | (0.099) |
| Ease of reaching market with permanent building | 0.18** | | |
| | (0.090) | | |
| Ease of reaching shopping complex | 0.22*** | | |
| | (0.077) | | |
| Flooding | | -0.28 | -0.057 |
| | | (0.29) | (0.27) |
| Earthquake | | -0.025 | -0.047 |
| | | (0.13) | (0.11) |
| Whirlwind/tornado/hurricane | | 0.36 | -0.12 |
| | | (0.24) | (0.22) |
| Drought | | -0.33 | 0.27 |
| | | (0.72) | (0.64) |
| Outbreak (last year): Vomiting/diarrhea | | -0.52 | -0.47 |
| | | (0.43) | (0.38) |
| Outbreak (last year): Malaria | | 0.28 | 0.42** |
| | | (0.25) | (0.21) |
| Outbreak (last year): Bird flu (1 case is considered an outbreak) | | -6.08 | -5.87 |
| | | (5.18) | (4.57) |
| Outbreak (last year): Tuberculosis | | -0.37 | -0.98* |
| | | (0.68) | (0.57) |

*Notes:* Data source: 1996 and 2011 PODES data and estimates from model. Table shows the regression coefficient between the estimated amenity value and the amenity measure given in each row. 1996 PODES data are correlated with the model estimates for 1995; 2011 PODES data are correlated with model estimates for 2011 and 2012.

Appendix Table 8: Robustness: effect of migration costs on growth, Indonesia

| | (1)<br>Sub. elasticity = 4 | (2)<br>Sub. elasticity= 6 | (3)<br>Sub. elasticity = 8 |
|---|---|---|---|
| *Productivity spillover = 0* | | | |
| $\lambda = -0.08$ | 1.115 | 1.115 | 1.115 |
| $\lambda = -0.05$ | 1.115 | 1.115 | 1.115 |
| $\lambda = 0$ | 1.115 | 1.114 | 1.113 |
| *Productivity spillover = 0.05* | | | |
| $\lambda = -0.08$ | 1.122 | 1.122 | 1.122 |
| $\lambda = -0.05$ | 1.122 | 1.121 | 1.121 |
| $\lambda = 0$ | 1.121 | 1.119 | 1.117 |
| *Productivity spillover = 0.08* | | | |
| $\lambda = -0.08$ | 1.125 | 1.125 | 1.125 |
| $\lambda = -0.05$ | 1.125 | 1.125 | 1.125 |
| $\lambda = 0$ | 1.124 | 1.121 | 1.118 |

*Notes:* Table shows the effect of reducing migration costs to zero on labor output. Table shows different combinations of amenity and productivity spillovers, for different values of substitution parameter. Calculated for model with selection.

Appendix Table 9: Robustness: effect of relative amenities on growth, Indonesia

| | (1)<br>Sub. elasticity = 4 | (2)<br>Sub. elasticity= 6 | (3)<br>Sub. elasticity = 8 |
|---|---|---|---|
| *Productivity spillover = 0* | | | |
| $\lambda = -0.08$ | 1.036 | 1.046 | 1.054 |
| $\lambda = -0.05$ | 1.046 | 1.059 | 1.069 |
| $\lambda = 0$ | 1.062 | 1.082 | 1.097 |
| *Productivity spillover = 0.05* | | | |
| $\lambda = -0.08$ | 1.042 | 1.057 | 1.068 |
| $\lambda = -0.05$ | 1.054 | 1.072 | 1.087 |
| $\lambda = 0$ | 1.074 | 1.102 | 1.126 |
| *Productivity spillover = 0.08* | | | |
| $\lambda = -0.08$ | 1.046 | 1.065 | 1.080 |
| $\lambda = -0.05$ | 1.059 | 1.083 | 1.102 |
| $\lambda = 0$ | 1.082 | 1.118 | 1.150 |

*Notes:* Table shows the effect of reducing relative amenities to zero on labor output. Table shows different combinations of amenity and productivity spillovers, for different values of substitution parameter. Calculated for model with selection.

Appendix Table 10: Robustness: effect of both migration costs and amenities on growth, Indonesia

| | (1) Sub. elasticity = 4 | (2) Sub. elasticity= 6 | (3) Sub. elasticity = 8 |
|---|---|---|---|
| *Productivity spillover = 0* | | | |
| $\lambda = -0.08$ | 1.134 | 1.142 | 1.148 |
| $\lambda = -0.05$ | 1.148 | 1.159 | 1.167 |
| $\lambda = 0$ | 1.161 | 1.176 | 1.187 |
| *Productivity spillover = 0.05* | | | |
| $\lambda = -0.08$ | 1.144 | 1.156 | 1.167 |
| $\lambda = -0.05$ | 1.160 | 1.176 | 1.190 |
| $\lambda = 0$ | 1.176 | 1.197 | 1.215 |
| *Productivity spillover = 0.08* | | | |
| $\lambda = -0.08$ | 1.151 | 1.167 | 1.182 |
| $\lambda = -0.05$ | 1.169 | 1.189 | 1.208 |
| $\lambda = 0$ | 1.186 | 1.213 | 1.238 |

*Notes:* Table shows the effect of reducing both migration costs and amenities to zero on labor output. Table shows different combinations of amenity and productivity spillovers, for different values of substitution parameter. Calculated for model with selection.