Spatial Labor Market Power in Sub-Saharan Africa:

The Roles of Self-Employment and Migration*

Samuel Marshall

University of Warwick

October, 2024

Abstract

Labor markets in low-income countries are characterized by large gaps between rural and urban income, between wage and self-employment income, and high rates of self-employment. Standard explanations for these features are frictions that prevent the efficient allocation of resources. I propose an alternative mechanism: firm labor market power. I develop a spatial general equilibrium model of monopsony to disentangle the role of labor market power from migration costs and job search costs. I identify the labor supply curve using Tanzania's 2010 sectoral minimum wage law. I find that rural labor markets are less competitive than their urban counterparts. This finding is driven by the higher share of wage workers employed in large firms. Moving to the competitive equilibrium causes total output to rise by 4.8%. Conversely, reducing migration costs by 10% reduces total output by 4.2%. This counterintuitive finding is explained by the fact that workers choose where to live and work based on the total value of wages and amenities. This creates a wedge between the productively efficient and welfare maximizing labor allocations. The standard result that reducing migration costs causes output to rise is reconciled through either a unidirectional decrease in migration costs in the direction of the city or a symmetric reduction and competitive labor markets.

*I am grateful to Clement Imbert, David Lagakos, and Federico Rossi for supervising this project. For data support, I thank James Mbongo and Saruni Njipay at the Tanzania National Bureau of Statistics. For helpful comments and suggestions, I thank Wiji Arulampalam, Johannes Boehm, Francisco Buera, Stefano Caria, Kevin Donovan, James Fenske, Martin Fiszbein, Ludovica Gazze, Doug Gollin, Junnan He, Kyle Herkenhoff, Allan Hsiao, Illenin Kondo, Angelica Martinez Leyva, Isabelle Mejean, Simon Mongey, Claire Montialoux, Tommaso Porzio, Clara Santamaria, Marta Santamaria, John Sturm, Mike Waugh, Yanos Zylberberg, and seminar participants at the University of Warwick, the BSE summer forum Migration workshop, the University of Sussex, the PSE-CEPR policy forum, the 2023 EEA-ESEM congress, the Sapienza PhD Conference, the 2nd Naples School of Economics PhD and Post-Doctoral Workshop, Boston University, the 2024 STEG annual meeting, the Bristol Applied Economics Meeting, Sciences Po, the XXVII Vigo Workshop on Macroeconomic Dynamics, the Federal Reserve Bank of Minneapolis and the Oxford Development Workshop.

1 Introduction

A fundamental difference between rich and poor countries is their labor markets. Low-income countries are characterized by large gaps between rural and urban income, between wage and self-employment income, and high rates of self-employment.[1] Standard models reconcile these stylized facts through various frictions (*e.g.* Restuccia, Yang, and Zhu, 2008). However, a job is more than a wage; it is a location, an environment, tasks, and a part of the worker's identity. People value these things which can be a source of labor market power for firms, allowing them to pay wages below the marginal product of labor.[2] This can be a source of labor misallocation even when workers confront no frictions. In this paper, I quantify the role of labor market power in explaining these stylized facts. Doing so reveals new insights about the labor supply curve, namely that reducing frictions may not lead to a more productive labor allocation.

There are two sources of labor market power: monopsony, in which all firms can mark down wages because workers value job amenities, and oligopsony, in which relatively large firms set wages strategically. That markets are not perfectly competitive has been well established in high-income countries.[3] Recent cross-country evidence finds that labor markets in poorer countries are less competitive (Armangué-Jubert, Guner, and Ruggieri, 2023), although these effects are mitigated by high rates of self-employment (Amodio, Medina, and Morlacco, 2024). An open question that remains at all levels of development is how labor market power varies across space, and what role it plays in the distribution of labor. I aim to make progress in answering that question here.

The contribution of this paper is to quantify the spatial distribution of labor market competition across space in a low-income country context. I do so by constructing a spatial general equilibrium model of monopsony which accounts for self-employment, job search costs and migration costs. I apply the model to Tanzania in 2010 when the nation enacted its first minimum wage law, which I use to identify the labor supply elasticity. The model replicates the standard result that wages are more

[1] These facts have been remained at the center of economic development for decades (Lewis, 1954; Harris and Todaro, 1970; Kuznets, 1973; Gollin, 2008; Young, 2013; Poschke, 2024)

[2] This may not be the case if there are strong labor unions (Manning, 2004), minimum wages (Ashenfelter, Farber, and Ransom, 2010), wage indexing (Guillouzouic, Henry, and Monras, 2024) or other labor market interventions.

[3] See Yeh, Macaluso, and Hershbein (2022) for the US; Manning (2003) for the UK; Hirsch, Jahn, Manning, and Oberfichtner (2022) for Germany.

competitive in small firms, while making two new predictions. First, wages are more competitive in markets with more self-employment. Wage-employment and self-employment are not disjoint labor markets. Transitions between the two are high (Donovan, Lu, and Schoellman, 2023). When either job search costs or self-employment earnings are high, firms must offer higher wages to pull workers out of self-employment. Second, wage competition has an inverse-u shape relationship with migration patterns. In low emigration areas, high migration costs trap workers in isolated labor markets, making it hard to leave. This allows firms to post lower wages because workers cannot emigrate to more competitive markets. In high immigration areas, workers are arriving for other reasons, *e.g.* amenities, number of job opportunities. This results in a surplus of labor that allows firms to post lower wages. The most competitive wages are those in markets in which migration flows are most tenuous and firms must post higher wages to prevent migrants from redirecting. These features suggest some ambiguity as to whether urban or rural labor markets are more competitive. In practice, rural labor markets are less competitive because their workers are, on average, employed in larger firms.

I then use the model to quantify the relative importance of labor market power, job search costs, and migration costs on labor misallocation both across space and into self-employment. Moving to competitive wages increases firm employment by 4.4 percentage points, which translates into an increase in total output of 4.8%. These gains are overshadowed by the gains from decreasing search costs by 10%, which causes firm employment to rise by more than 34 percentage points and output by 22%. However, these results contrast with a 4.2% decline in total output for an equivalent reduction in migration costs. The latter result is reconciled through the fact that when people choose where to work, they also choose where to live, which may not be the most productive location.[4] Average wages in Dar es Salaam, Tanzania's megacity, are just 11% higher than rural wages, despite being by far the most productive area. However, if I reduce migration costs and move to competitive wages, total output rises. I develop these results via the following steps.

To quantify the spatial distribution of labor market power, I develop two-sector spatial general equilibrium framework of monopsonistic competition that is able to match several stylized facts in the

[4]The idea that location amenities compensate for wages was first proposed by Rosen (1974) and Roback (1982). More recently Albouy (2008) showed that the rank-ordering of locations in the US is nonsensical when not accounting for variation in the cost-of-living. Indeed, when I do not deflate prices the output loss is even larger.

data. The model builds closely on Berger, Herkenhoff, and Mongey (2022)'s model of monopsonistic competition, which I combine with a spatial equilibrium framework.[5] The main facts that the model needs to replicate are that more than 85% of workers are engaged in self-employment while their average labor income is lower than that in nearby wage jobs. I match these observations by having workers choose between two sectors of production, firms and self-employment, as in Lewis (1954). Workers face two types of frictions–search costs, which make it costly to access jobs in firms, and migration costs which may prevent them from optimally locating across space.[6] To the best of my knowledge, this paper is the first to construct a unified spatial general equilibrium framework of monopsonistic competition with self-employment.

In the model, all self-employed workers earn the average product of labor.[7] Wage workers, on the other hand, are paid a markdown on the marginal product of labor. Wage markdowns are generated through an upward sloping labor supply curve and firms setting wages via Bertrand competition. The labor supply curve is generated through amenities which are of three types: non-rival job and location amenities, and idiosyncratic amenities, which are the features of a job for which the value is specific to the worker.[8] In the absence of frictions, amenities create variation in wages across firms and space, but strategic wage setting is needed to generate wage markdowns.[9]

Wage markdowns are calculated as $\varepsilon/(1+\varepsilon)$, where ε is the labor supply elasticity. In the perfect

[5]The spatial component is based on the framework developed by Monte, Redding, and Rossi-Hansberg (2018) and most closely follows Bryan and Morten (2019). See Redding and Rossi-Hansberg (2017) for a recent review.

[6]search costs aim to capture the present discounted value of searching for a job as well as all of the unobservable frictions that prevent workers from engaging in wage work, *e.g.* information asymmetries between firms and workers (Abebe, Caria, Fafchamps, Falco, Franklin, and Quinn, 2020; Alfonsi, Bandiera, Bassi, Burgess, Rasul, Sulaiman, and Vitali, 2020), local commuting costs (Monte et al., 2018; Abebe et al., 2020), search frictions (Abebe, Caria, and Ortiz-ospina, 2021), and demands for time in the household. That migration costs exist and are large has been well established in the literature. See Lagakos (2020) for a recent review. However, as these costs are unobservable, I abstain from claiming that this feature of the model matches a stylized fact in the data. However, the returns to migration are large. Lagakos, Marshall, Mobarak, Vernot, and Waugh (2020) find that migrants in Tanzania earn on average 11% more after migrating and those that migrate from rural to urban districts earn 21% more.

[7]The self-employment sector as a whole exhibits decreasing returns to scale in each market. This feature of the model aims to capture the fact that in rural districts there is a fixed amount of land on which to do agriculture, while in urban districts increasing the number of *e.g.* street vendors reduces the earnings of each one. This assumption is not quantitatively important for the findings.

[8]Examples are flexible hours for job amenities and sunny weather for location amenities. Examples of idiosyncratic amenities are working with a friend or being close to home. In an extension, I allow location amenities to be endogenous. This reflects the idea that congestion may reduce the value of using a park, for example.

[9]In Bertrand competition, firms choose a wage observing the labor supply curve and the wages in all other firms. The functional form of the markdown will depend upon the extent to which firms internalize the effect of their wage on the wage in other firms.

competition model, labor supply is perfectly elastic and this term reduces to unity–workers are paid their marginal product. What varies across models of monopsonistic competition is the functional form of the labor supply elasticity. In this model, each firm faces its own labor supply curve which is a function of three elasticities: the between-firm elasticity (η), the sector elasticity, between wage-work and self-employment (γ), and the migration elasticity (θ).

In the next step, I turn to the data to empirically estimate these elasticities using Tanzania's 2010 minimum wage law. The legislation specified a specific level for 20 industries as well as a national floor for all others. The law had a heterogeneous effect across firms in two ways. First, firms in the same industry had different levels of exposure to the minimum wage due to differences in their pre-existing wages; firms with lower wages needed to raise their wages more to comply with the minimum wage law. Second, firms in different industries with the same wage bill needed to raise their wages by different amounts to be compliant.[10] I identify the between-firm elasticity (η), using changes in wages and employment following the introduction of the minimum wage law. Following Card and Krueger (1994), I instrument for wages using the gap between the pre-policy-period wage-bill and the minimum-wage compliant wage-bill. Because firms are setting wages strategically, these results will be biased.[11] The exception is very small firms with no oligopsony power who face an isoelastic labor supply elasticity which is precisely η. Thus, I interact the gap instrument with the firm's share of local firm employment. I estimate a between-firm elasticity of 2.5 which is considerably lower than Berger et al. (2022)'s estimate for the US (10.85) but is comparable to Franklin, Imbert, Abebe, and Mejia-Mantilla (2024)'s estimate for Ethiopia (3.36).

I estimate the sector elasticity (γ) via the model-generated moment that relates the ratio of firm-

[10] A potential concern for identification is the well-documented lack of enforcement of minimum wage laws in developing countries (Basu, Chau, and Kanbur, 2010). Indeed, minimum wage compliance is not perfect. 19% of wage workers were paid below the minimum wage in 2007, the last observed pre-policy year. The share fell to 10% in 2010, representing a 47% rate of compliance, which is consistent with the average rate of compliance in Sub-Saharan African countries (Bhorat, Kanbur, and Stanwix, 2017). By 2013, the rate of non-compliance fell below 5%, likely reflecting both an adjustment period (Clemens and Strain, 2022) and changes in nominal wages due to inflation (Kaur, 2019). For tractability, in the model I assume full compliance and do not consider how the level of the minimum wage is likely to affect compliance.

[11] The reduced form estimate is confounded by the Nash-equilibrium response of firms to changes in the wage in other firms, except for very small firms. Put differently, a change in the wage offered by any firm will affect the wage offered by all other firms. However, as the firm's market share declines, its effect on the wages of other firms decreases. The within-market elasticity is identified by the limiting case in which the firm is atomistic and its choice of wages does not affect those of other firms.

to self-employment to relative wages in each labor market. The market-level firm-wage is a CES aggregator of each firms wage where the elasticity is the between-firm elasticity, η. I instrument for the wage ratio with the one I would calculate if all firms were paying exactly the minimum wage. I estimate a sector elasticity of 1.5. Since the market-wage index is a function of η, I re-estimate the sector elasticity with different calibrated values of η. The ratio of the sector- to firm-elasticity is roughly constant at 0.6 across the alternative calibrations. This indicates that the way in which firms compete with one-another is different from the way in which they compete for workers from self-employment.

I then use the spatial equilibrium framework to estimate the migration elasticity (θ). The flow of migrants between two locations depends upon the relative wages, relative amenities, and migration costs. To disentangle the role of relative amenities on migration flows, I again instrument for each market-wage index with the equivalent index as-if all firms were paying the minimum wage. My ability to use this instrument relies critically upon the fact that there is spatial variation in industrial composition, and hence applicable minimum wages, across space. I estimate a migration elasticity of 1.4. This value is higher than Berger et al. (2022)'s estimate for the US (0.42), but is comparable to Tombe and Zhu (2019)'s estimate for China (1.5), Bryan and Morten (2019)'s estimate for Indonesia (3.2), and is consistent with their finding that migration elasticities are larger in developing than in developed countries.

With all of the parameters of the model in hand, I am able to estimate the markdown in each firm. This allows me to map the spatial distribution of wage markdowns. There is substantial heterogeneity in markdowns across space with most workers being paid between 66-71% of marginal product. This implies that the earnings and productivity gap of wage workers will not be equal. Indeed, because rural labor markets are less competitive, the rural-urban income gap for wage workers overstates the productivity gap. I then use the model to simulate the counterfactuals discussed earlier. The surprising finding that reducing migration costs causes total output to fall is robust to different assumptions on labor market competition, congestion and agglomeration forces, and excluding self-employment from the analysis. Finally, I conclude with a discussion of the observed patterns of urbanization without growth in sub-Saharan Africa (Henderson, Roberts, and Storeygard, 2013; Gollin, Jedwab, and Vollrath, 2016; Henderson and Kriticos, 2018). I find suggestive evidence that this model is better able to replicate these patterns than a standard model in which workers face frictions but choose where to live and work based solely on wages, *i.e.* there is no wedge between the output and welfare maximizing labor supply distribution.

This paper contributes to several strands of the literature. I provide the first direct estimates of wage markdowns in a low-income country context. In a recent cross-country analysis, Armangué-Jubert et al. (2023) estimate wage markdowns as low as 55% in low-income countries. I find that workers are paid between 64-71% of their marginal product in Tanzania. The difference in our estimates can be explained as follows. First, markets appear less competitive in their analysis because they do not account for self-employment, which reduces the market share of each firm and hence their ability to dictate local wages. Second, estimation of the wage markdowns depends critically upon identification of the labor supply elasticity.[12] Several recent papers have estimated wage markdowns in Latin America with widely varying findings. In atomistic firms, in which markdowns are driven entirely by the labor supply elasticity, markdown estimates range from 98% of marginal product in Peru (Amodio et al., 2024), 70% in Columbia (Amodio and de Roux, 2021), and 50% in Brazil (Felix, 2022).

Second, this paper makes a theoretical contribution by constructing a simple spatial general equilibrium framework of monopsony that accounts for self-employment. Typically, estimation of the labor supply elasticity requires data on firm hires or exits (Manning, 2003). The general equilibrium framework developed by Berger et al. (2022) alleviated this constraint. To account for spatial frictions, I combine their model with the spatial general equilibrium framework developed by Monte et al. (2018). These models have been used in the development context to study both migration (Bryan and Morten, 2019) and commuting frictions (Franklin et al., 2024). I add to this a sectoral choice between

[12] As noted above, wage markdowns are calculated as $\varepsilon/(1+\varepsilon)$, where ε is the labor supply elasticity, which I identify using Tanzania's 2010 minimum wage law. A closely related literature studies the interaction between firms with labor market power and minimum wages (Manning, 2006, 2019; Berger, Herkenhoff, and Mongey, ming). The direct effect of the minimum wage is less clear in this context due to the well documented lack of enforcement in developing countries (Basu et al., 2010; Rani, Belser, and Oelz, 2013; Bhorat et al., 2017; Mansoor and O'Neill, 2021). In the development context, the employment effects of the minimum wage vary from no disemployment effects in Brazil (Almeida and Carneiro, 2012; Derenoncourt, Noelke, Weil, and Taska, 2021) to positive effects on formal employment in Indonesia (Magruder, 2013). To focus ideas, in this analysis, I abstract from the compliance decision of firms and use the minimum wage only as a means of identifying the labor supply elasticity.

wage-work and self-employment in the spirit of Lewis (1954).

Third, this paper contributes to the literature studying income gaps in developing countries. Kuznets (1973) advocated for development through structural transformation–the reallocation of productive inputs from agriculture to manufacturing and eventually to services. Today this transition has yet to occur in many countries and a large share of workers remain in low-productivity agriculture (Gollin, Lagakos, and Waugh, 2014). Most agriculture occurs in rural areas and, more recently, the literature has turned to studying the related rural-urban income gap as a proximate cause for sectoral income gaps.[13] While some authors argue that workers are optimally located across space (Young, 2013; Hicks, Kleemans, Li, and Miguel, 2021), the rural-urban income gap is often reconciled through high costs of migration (Lagakos et al., 2020).[14] This paper focuses on two types of frictions that propagate income gaps: job search costs and migration costs. I find that reducing search costs causes output to rise.[15] While this result may seem obvious, the reason is more nuanced. Job search costs prevent workers from accessing the full range of jobs. A directed reduction to a particular firm may reduce output if the firm is unproductive. Conversely, I find that reducing migration costs causes output to fall. This result is a continuation of the downward revisions to estimates of the gains from reducing migration costs (Bryan and Morten, 2019). While Bryan and Morten (2019) find modest gains from total output, their results follow from workers moving to places where they are more productive rather than that which maximizes welfare. The gap between the output and welfare maximizing labor allocations is inline with the rethinking of whether labor is misallocated in agriculture in the US context (Herrendorf and Schoellman, 2015). The decline in total output that I find is linked to a net decline in urban employment. This finding is not without precedent. Faber

[13]Henderson and Kriticos (2018) documents that there is more agriculture happening in African cities than in other developing countries. However, this urban agriculture appears to primarily be an artifact of the data: smaller cities have rural areas within their administrative boundaries. Indeed there is very little agriculture in primary cities. These findings do highlight two important features of cities in Africa though: they are small and have not induced a transition out of agriculture in the surrounding areas.

[14]The literature has proposed several mechanisms through which monetary costs may preclude profitable migration including risk in finding employment and near subsistence levels of consumption at origin (Bryan, Chowdhury, and Mobarak, 2014) or village risk-sharing networks (Morten, 2019). However the main cost of migration is non-monetary (Lagakos, Mobarak, and Waugh, 2023; Imbert and Papp, 2020; Bryan, Chowdhury, Mobarak, Morten, and Smits, 2021). Indeed, reviewing the literature, Lagakos (2020) concludes that while early evidence pointed toward monetary costs as the primary constraint against migration, more recent evidence has determined that migration costs are largely non-monetary.

[15]See Caria and Orkin (2024) for a recent review of urban labor market frictions.

(2014) finds that Chinese highways led to a decrease in local rural output. Imbert, Seror, Zhang, and Zylberberg (2022) find that an influx of immigrants in China led to a decline in labor productivity in manufacturing firms. They attribute the decline in productivity to sticky capital allocation within firms. Applying the methodology of Au and Henderson (2006), Dar es Salaam is not necessarily too large, but its economy is overly concentrated in services (Gollin et al., 2016; Henderson and Kriticos, 2018). From an amenity perspective, Lagakos et al. (2023) find substantial disutility of poor urban housing when interpreting the experimental results of Bryan et al. (2014) in a general equilibrium framework; increasing the quality of urban housing is equivalent to raising migrant wages by 21%.

2 Data

The introduction of Tanzania's first minimum wage law in 2010 makes it an ideal point in time to estimate the labor supply curve. Thus, I focus my efforts on collecting data around this point in time. No single data source contains all of the information that I need, so I combine data from several sources including: the Tanzanian Employment and Earnings Survey (EES), the 2010 Tanzanian National Panel Survey (NPS), which is a part of the World Bank's Living Standards Measurement Study (LSMS) surveys, the 2002 and 2012 Censuses, Tanzania's 2014 Integrated Labor Force Survey (ILFS), and industrial minimum wage levels from the Tanzanian gazette, a monthly bulletin that reports new laws. I describe the main source of data used for each piece of the model below and brief discussion when there is more than one alternative.

Population Distribution I estimate the initial population distribution using the 2002 and 2012 census. I define the population in each district as the total prime-aged employment (inclusive of the self-employed). I do not directly observe the population in 2009. To estimate it, I use the value in 2012 and the growth rate (g_o) as implied from the 2002 and 2012 censuses

$$L_o^{2009} = L_o^{2012} \times (1 + g_o)^{-3}$$

To account for re-districting that occurred during this period, I use the time-consistent district boundaries from Integrated Public Use Microdata Series (IPUMS) International to define the set of districts used in this analysis.

Firm Wages and Employment Data on firm employment and earnings is taken from the Tanzanian EES, an annual survey of firms. Marshall (2023) showed that the average wages and employment in this data is very similar to that in the Census, LSMS, and ILFS. I report the average wages and employment in Table 2. The dataset covers all firms with at least fifty employees and a sample of firms with fewer. In total, around 10,000 firms are surveyed each year. The survey covers all sectors of the economy, and is much larger than other manufacturing surveys. This coverage comes at the cost of information in the survey. For each firm, I know employment by group, male, female, prime-aged, *etc.* and total payments to each of those groups. However, I do not observe any single wage. For the main analysis, I calculate the wage for each firm as the total payments to prime-aged non-foreign-born workers over total employment in this group.

I am partially able to observe the distribution of wages in the firm through a series of questions that ask, "how many workers are paid between X and Y." In Section 6, I use this information to estimate the share of workers in each firm who are paid below the minimum wage. The survey was not run in 2008 or 2009, so I use the period 2005-2007 to estimate the firm's exposure to the minimum wage in the pre-policy period.

Self-Employment Earnings Self-employment income is notoriously difficult to calculate (Gollin et al., 2014). To this end, I estimate earnings relative to average firm wages in each market using household consumption per adult equivalent from the 2010 LSMS as

$$w_{sd} = w_{fd}^{\text{EES}} \times \frac{\text{consumption-equivalent}_d^{\text{self-employed}}}{\text{consumption-equivalent}_d^{\text{employed}}}$$

Where w_{fd}^{EES} is the average firm wage in district d in the EES, and the consumption equivalent values are the average among prime-aged individuals who report that their main occupation in the past year was either self-employment or wage-work.

Alternatively, I could directly use information on self-employment earnings from both the 2010 LSMS and 2014 ILFS. Both of these suffer from the aforementioned reporting issues. Those engaged in self-employment agriculture consume some of their output and small household businesses may either under-report or not know their total earnings. Additionally, the 2014 ILFS, while quite large at 40,000 observations, occurred four years after the minimum-wage was implemented. Thus the deflated earnings from that year may not reflect the average income in 2010 if their was selection out of self-employment over that period. Self-employment income in the 2010 LSMS is not ideal because the the sample is much smaller at 12,000 observations, and the sampling scheme was not designed to be representative at the district level.

To calculate the total number of workers in self-employment, I first estimate the share of prime-aged workers in wage-work in each district in 2010 using the growth rate of employment between 2002 and 2012 from the census. I then calculate self-employment to be consistent with the observed number of workers in the EES

$$n_{sd} = (1 - s_d^{\text{Census}}) \times n_{fd}^{\text{EES}}$$

Where s_d is the share of employment in district d in wage-work, as defined in the model and n_{fd}^{EES} is the total employment in the EES. It is important to note that the total firm employment in each market is the weighted sum of each firm's employment, but the employment in any given firm is not weighted. This keeps me from over-weighting any single firm observation.

Migration For the main analysis, I use the one-year migration flows from the 2012 Census. I limit my attention to the set of prime-aged individuals (15-65) to abstract from family moves (under 15) and return migration (over 65). Although these flows correspond to the year 2011, the coverage in the census is much larger than that in any of the surveys. However, given that the question is limited to moves during the past year, the matrix of migration between district pairs is still quite sparse. Hence, for the estimation of the migration elasticity in Section 7, I aggregate to the regional level (23 distinct units). When I estimate the structural migration costs, I use the implied ten-year flows between districts following the procedure outlined in Appendix B.2.

As a robustness exercise, I use data from the 2012, and 2014 LSMS and the 2014 ILFS to estimate migration flows between district pairs. These surveys ask about the year of migration, so I have a better idea of when the migration happened. With the 2014 surveys, the reported migration year may be inaccurate due to recall bias (Kirchberger, 2021). I address this concern by focusing on migration flows during the five-year period 2010-2014. As noted above, the ILFS was designed to be representative at the district level while the LSMS surveys were not. To account for differences in the sampling, I first weight the migration flows from the 2012 and 2014 LSMS surveys by the number of years since 2010. I then equally weight these shares with those from the ILFS.

Minimum Wages I get industry level minimum wages for the 2010 law from the Tanzania National Gazette, a monthly bulletin that includes all new national regulations. I assign to each two-digit ISIC code a corresponding minimum wage. I report the minimum wage-ISIC crosswalk in Appendix Table 23. For industries with sub-sector levels (*e.g.* mining has three levels that are distinguished by the firm's license class) that cannot be discerned from standard ISIC codes, I assume that all firms in that sector are subject to the lowest minimum wage level.

3 Context & Motivating Facts

Labor markets in developing countries differ from those in developed countries along several dimensions that are relevant for labor market power. I first document several stylized facts about self-employment to motivate the theoretical framework before discussing the firm size distribution and migration frictions.

The Self-Employment Rate is High The high rate of self-employment may ameliorates the labor market power of firms. In Table 3, I report the rate of employment by type in the 2012 Census. 70% of prime-aged individuals are engaged in some type of employment. Of those, 85% are self-employed. To put this number in context, in the US, just 9% of workers are self-employed. Self-employment here should not be conflated with informal employment. In Brazil where informal employment is a primary concern, the rate of self-employment is 25%.[16] Self-employment is a much larger share of the economy in Tanzania than in India (49%) or the USA in 1910 (28%). The lack of jobs is likely to deliver a great deal of market power to firms. However, the high rate of transition between self-employment and wage work means that the potential earnings in self-employment are likely to limit the extent to which firms can mark down wages (Donovan et al., 2023).

Tanzania's 70% employment rate is an outlier. This number is driven by the rate of self-employment in rural areas. Indeed, when restriction to urban areas only, the employment rate is more similar to that in the USA or Brazil. However, the high rate of self-employment in Tanzania is not an outlier for sub-Saharan Africa. In Table 4, I report the employment share and self-employment rate by rural-urban designation for several sub-Saharan African countries. For many countries, the self-employment rate is above 75%.

Average Wage > Average Self-Employment Earnings One may expect that the high rate of self-employment in this context reflects higher earnings than in wage-work. However, measuring relative earnings may be difficult if the self-employment are under-reporting (Herrendorf and Schoellman, 2015). I ameliorate this concern by plotting the average log-consumption per adult equivalent in self-employment against that in wage work in Figure 1. Above the black indicates are districts in which self-employment consumption is below that in wage-work. While consumption is typically higher among wage-workers, that is not uniformly true.

Competition between Firms is not the same as competition between firms and self-employment In Figure 2, I plot the log wage against the employment market-share for firms and self-employment. Average wages are generally rising with employment share for firms. This is consistent with more productive firms being larger and paying higher wages. However, for self-employment that relationship does not hold. This suggests that firms compete with one another differently from how they compete for labor from self-employment.

[16] See (Ulyssea, 2018, 2020; Derenoncourt et al., 2021). In this analysis, I do not distinguish between formal and informal employment within the firm.

Firm Size Distribution I report the distribution of firm size and employment in Table 5. The missing missing middle firm size distribution in middle income countries described by (Hsieh and Olken, 2014) is not present here. The employment share is rising with firm size–a larger share of the labor force works in firms with 10-49 employees than in firms with fewer than ten employees. This difference in the employment distribution is likely due to the higher rate of self-employment here than in India, Indonesia or Mexico. Comparing rural and urban districts in Tanzania, the share of workers in large firms (50+ employees) is higher in rural (66%) than urban districts (53%). However, the distribution of firm size exhibits the same pattern in both rural and urban districts. Approximately two-thirds of firms have fewer than ten employees.

Migration Costs Finally, migration is likely to affect the labor market power of firms. In many developing countries, 20-25% of individuals migrate out of rural areas as young adults (Young, 2013). While the literature has largely focused on rural to urban moves, there is a non-negligible rate of migration between urban and between rural areas. In Table 6, I report the five-year migration rates across a number of surveys. Indeed, rural to urban moves account for less than half of the total migration out of rural areas, while the majority of migration is between urban areas.

To assess whether migration is lower in more isolated areas, I plot the one-year emigration rate in the 2012 census in Figure 3. Emigration rates are highest near the four main cities. The area of higher migration rates extends geographically further around Dar es Salaam, Tanzania's largest city. Emigration rates in the hinterlands are as low as 3%. This spatial depiction of emigration rates is consistent with a gravity model of migration. Individuals who are further away from the city face higher costs to migrate.

3.1 Tanzania's Minimum Wage Law

The Labour Institutions Order of 2010 created Tanzania's first minimum wage law. The law set forth specific levels for 20 sectors and a national minimum wage for all others.[17] The law had

[17] See Appendix C for details of the Wage Order. The interested reader is directed to Marshall (2023) for a more detailed discussion of the effects of the law.

a heterogeneous effect across firms for two reasons. First, firms in the same sector were exposed differently to the law based on their pre-existing wages. Second, firms with equivalent wages in sectors with different minimum wages needed to raise their wages by different amounts to be compliant with the law. This together with differences in sectoral composition across districts in Tanzania lead to a heterogeneous impact of the law across space.

A major concern for this analysis is that firms did not comply with the minimum wage law. It has been documented that the rate of compliance and enforcement of minimum wage laws in sub-Saharan Africa is low (Bhorat et al., 2017), and that compliance is lower in countries with complex minimum wage laws (Rani et al., 2013). However, in Section 6, I show that the rate of compliance was high. Reconciling these differences is outside the scope of this paper and is an area for future research.

4 Economic Framework

The aim of this section is to develop a general equilibrium framework that disentangles the roles of labor market power, search costs, and migration costs on labor misallocation. The model combines Berger et al. (2022)'s monopsony framework with the spatial general equilibrium model employed by Bryan and Morten (2019).[18] New here is the addition of self-employment as a job choice, and search costs that keep workers out of wage-work.

The economy is composed of a unit measure of workers indexed by $\omega \in [0, 1]$ and a discrete set of locations indexed by o or $d \in D$. In each location there are $(M_d - 1)$ firms indexed by (f) and a self-employment sector indexed by (s). I use the word job and index (i) to refer to the set of firms and the self-employment option. Locations are characterized by a non-rival amenity value (B_d), a labor market friction (δ_d), and an initial population $L_o \sim F(L)$.

[18]Here I deviate from Berger et al. (2022) by explicitly assuming that a labor market is a location rather than a location-industry pair. To keep things simple, I abstract from any sector-specific constraints arising from the minimum wage law.

The spatial general equilibrium framework is based largely on the commuting model developed by Monte et al. (2018), later used in a development context by Franklin et al. (2024). The commuting costs in these models are more akin to migration costs in that they explain the spatial misallocation of labor, albeit at a smaller scale. search costs, as I define them here may encompass commuting costs, but also include information asymmetries, familial constraints, *etc.*.

4.1 Technology

Both firms and the self-employed produce a homogeneous final good which is traded in a perfectly competitive national market at a price P which is normalized to one.

Firms Each firm is endowed with productivity $A_{f(d)} \sim F(A^d)$ and produces output (y_f), measured in value-added, using labor (n_f) as its sole input according to the production function:

$$y_f = A_f n_f^{\alpha} \tag{1}$$

Where $\alpha \in (0, 1)$, implying that production in each firm exhibits decreasing returns to scale. I also implicitly assume that expected firm productivity is independent across locations.

Self-Employment In each location there is a self-employment productivity $A_{sd} \sim F(A^s)$ that is common to all workers. Output is produced using the same technology as firms as if the self-employment sector were operating as a single firm:

$$y_{sd} = A_{sd} n_{sd}^{\alpha}$$

Each worker in self-employment earns the average product of labor $w_{sd} = A_{sd} n_{sd}^{\alpha-1}$.

Total output in each location and aggregate output are defined as

$$y_d = y_{sd} + \sum_{f \in d} y_f \quad ; \quad Y = \sum_d y_d$$

4.2 Preferences and Choices

In this section, I develop a discrete choice migration model and derive the labor supply curve to each firm from each origin. In what follows, I use the term job to refer to the set of firms in each location and the option of self-employment.

Individuals are heterogeneous in two dimensions: their place of birth (o), which determines the migration cost to all other locations, and their idiosyncratic tastes over jobs. These tastes are described by a vector of preferences (idiosyncratic amenities) for each firm and self-employment in each location $\zeta(\omega)$ which is distributed according to the multivariate Gumbel distribution

$$F(\zeta_{11},\dots,\zeta_{id},\dots,\zeta_{(M_D+1)D}) = -\exp\left[-\sum_d\left(\left(\sum_{f\in d} e^{-\eta\,\zeta_{f(d)}}\right)^{\frac{\gamma}{\eta}} + e^{-\gamma\,\zeta_{sd}}\right)^{\frac{\theta}{\gamma}}\right]$$

It is important to note here that the idiosyncratic amenities include a value for self-employment in each location. To be consistent with the data, two relationships between the elasticities must hold. First $\eta > \gamma$. This implies that if a firm were to raise its wage, it would pull more workers from other firms in the local labor market than from self-employment. Second, $\gamma > \theta$. This implies that if the wage in any firm or in self-employment were to go up, more labor would reallocate to that job from within the market than from other markets. In practice, the distributional assumption captures the fact that people have personal preferences over both locations and firms. One's preferences over firms within a market are correlated because they share a location. For example, an individual with a family network in a certain location will have a higher amenity value for all firms in that location.

Workers choose a job (i) from the set of firms (f) or self-employment (s), in a market (d) to maximize their indirect utility

$$v(\omega|o(\omega)) = \max_d\left\{B_d\tau_{od}\,\max\left\{\delta_d\max_{f\in d}\left\{e^{\zeta_{fd}(\omega)}b_f w_f\right\},\ e^{\zeta_{sd}(\omega)}w_{sd}\right\}\right\} + \Pi\partial\omega$$

The job choice is made in three steps. First, in each market (d) the worker chooses a firm based on the wage (w_f), the firm's non-rival amenity value (b_f), and their idiosyncratic amenity draw (ζ_f).[19] In the second step, the worker chooses between working in self-employment, where they earn (w_{sd}) and have idiosyncratic value (ζ_{sd}) or working at the firm they chose in the first step. However, there are search costs which keep workers out of wage-work which are captured by (δ_d). This value is common to all workers but may vary across markets.[20] I assume that $\delta_d \in (0,1]$. The lower bound ensures that

[19] While I use the term amenity here, b_f captures both positive and negative aspects of the firm, *e.g.* flexible hours, and hazardous work. Equivalently, b_f can be modeled as the location parameter for ζ_f in the amenity vector.

[20] To build intuition for these choices, first consider commuting costs, which are captured by this friction. If most

in each market a non-zero measure of workers will supply labor to firms. When $\delta_d = 1$, there are no labor market frictions but some workers will still be self-employed because of the distributional assumptions on the idiosyncratic preferences.

In the final step, the individual chooses among all markets (d) based on their indirect utility of working in their job choice, the non-rival amenity in d, (B_d) and the cost of migrating from their birth location o to d, (τ_{od}).[21] It is assumed that $\tau_{od} \in [0, 1]$ and $\tau_{oo} = 1$. The latter assumption states that there is no cost for not migrating while the former implies that a percentage of utility is lost for migrating to any other district. The special case where $\tau_{od} = 0 \, \forall \, d$ corresponds to infinite costs of migration. I make the additional assumption that migration costs are symmetric; that is $\tau_{od} = \tau_{do}$.

Firm profits (Π) are redistributed lump-sum across all workers (including the self-employed). This has no effect on choices, but does affect welfare.

5 Equilibrium

In this section, I present the equilibrium of the model and show how wage markdowns are affected by self-employment.

Labor Supply Under the distributional assumptions on the amenities, total labor supply from o to firm f or self-employment s in market d, can be expressed as

$$n_{fdo} = \left(\frac{b_f w_f}{W_{fd}}\right)^{\eta} \left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma} \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta} L_o \tag{2a}$$

$$n_{sdo} = \left(\frac{w_{sd}}{W_d}\right)^{\gamma} \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta} L_o \tag{2b}$$

jobs are at mining firms which are located far from villages where workers live, then the commuting costs will be high. Conversely, if most jobs are in local manufacturing in the village, then the costs will be low. Just as no two firms are located on top of each other (usually), the firm specific commuting time will depend upon where the firm is physically located and where the individual lives. This variation is captured by the individual's firm amenity.

[21] As with firm amenities, B_d captures both positive aspects of the location *e.g.* clean air, markets, natural beauty, as well as negative aspects, *e.g.* pollution, crime. Under the assumption that the average job amenity in each location is one, B_d can equivalently be written as the location parameter in the amenity distribution. The net location parameter for each firm is then $B_d b_f$. I avoid this simpler formulation for clarity here.

Where the wage indices are given by

$$W_{fd} := \left[\sum_{f\in d}(b_f w_f)^{\eta}\right]^{\frac{1}{\eta}} \quad , \quad W_d := [w_{sd}^{\gamma} + (\delta_d W_{fd})^{\gamma}]^{\frac{1}{\gamma}} \quad , \quad \mathbf{W_o} := \left[\sum_{d}(\tau_{od} B_d W_d)^{\theta}\right]^{\frac{1}{\theta}}$$

The first term in (2a) is the probability of choosing firm f conditional upon choosing employment in market d. Among firms, the labor share to firm f is increasing in the own wage and amenity at the rate η. When $\eta \to \infty$, markets are perfectly competitive–the wage-amenity value in each firm will equate. This implies that even under perfectly competitive markets, wage dispersion can be explained by firm amenities.

When $\eta < \infty$, workers have tastes over firms within the market and do not supply all of their labor to the firm with the highest value. The elasticity between firms is one source of market power within a market. This term is independent of the worker's origin and implies that workers from all origins will supply labor in equal proportion across firms.[22]

The second term in (2a) is the probability of choosing wage-work in market d. When $\gamma \to \infty$, some labor will still be misallocated into self-employment because of the search costs (δ_d). For $\gamma < \infty$, workers are elastic between wage-work and self-employment. This means that in the absence of labor market frictions, if all firms in the market lower their wages uniformly, not all labor will reallocate to self-employment.

The third term in (2a) is the probability of choosing market d. This expression says that first, locations with higher amenities (B_d) have more migrants from all origins. Second, there will be more migrants from o in places with lower costs of migration (τ_{od} closer to one).[23] Third, migration into any location is increasing in the number of firms; what matters for migration is not only the value of the wages, but the number of wage offers. When $\theta \to \infty$, the wage-amenity value will not necessarily equate across markets due to variation in the supply of labor across locations and the costs

[22] In practice, this means that a change in total labor supply in market d will only affect the distribution of workers across firms through the distortionary effect on relative wages of self-employment. When labor supply rises, wages in self-employment decline more slowly than those in firms, increasing the relative share of labor engaged in self-employment.

[23] However, for any two origins o and o' with $\tau_{od} > \tau_{o'd}$ does not imply that there will be more migration form o. This is because the cost of migration relative to other locations matters. For example, two rural villages with high costs of migration to Dar es Salaam may differ in their migration there because one is closer to a smaller city.

of migration (τ_{od}).

Welfare Following Franklin et al. (2024), the welfare value for workers born in o can be expressed by the expected utility

$$E[v_o] = \sum_d (B_d \tau_{od})^{\theta} \left(w_{sd}^{\gamma} + \left(\delta_d^{\eta} \sum_{f \in d} (b_f w_f)^{\eta} \right)^{\frac{\gamma}{\eta}} \right)^{\frac{\theta}{\gamma}} \psi + \Pi L_o \tag{3}$$

Where ψ is the Euler-Mascheroni constant $(\sim\ 0.577)$.

Aggregate Labor Supply Curve Because the measure of workers in the total economy is unitary, the labor supply to each firm can be be expressed as the share of workers in market d that are employed in firm f, multiplied by the share of workers in market d from o. Summing over origins, the aggregate labor supply curve to firm f can be expressed as

$$n_f = \left(\frac{b_f w_f}{W_{fd}} \right)^{\eta} \left(\frac{\delta_d W_{fd}}{W_d} \right)^{\gamma} \left[\sum_o \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}} \right)^{\theta} L_o \right] \tag{4}$$

Firm Problem Labor market competition is Bertrand, so firm f in market d takes as given local competitor's wages $\{w_{-f(d)}\}$, the local self-employment wage w_{sd}, the wage level in all other markets $\{W_{-d}\}$, and the aggregate distribution of labor $\{L_o\}$, and chooses its wage (w_f) to maximize profits

$$\max_{w_f} \quad A_f n_f^{\alpha} - w_f n_f \tag{5}$$

Subject to

$$n_f = \sum_o n_{fdo}(w) \quad , \quad n_{fdo}(w) = \left(\frac{b_f w_f}{W_{fd}} \right)^{\eta} \left(\frac{\delta_d W_{fd}}{W_d} \right)^{\gamma} \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}} \right)^{\theta} L_o \quad , \quad o = 1, \ldots, D$$

Under the assumption of Bertrand competition, the firm understands that $\partial n(w_f, W_{fd}, W_d, \mathbf{W_o})/\partial w_f \neq 0$, $\partial W_{fd}/\partial w_f \neq 0$ and $\partial W_d/\partial w_f \neq 0$. That is, firm wages affect hiring both directly and indirectly through the market level wage index W_d.

Self-Employment operates like a firm in this context. However, it differs from firms in that it does not compete in Bertrand competition. It pays each worker the average revenue product of labor. This implies that all profits from self-employment are paid to the worker.

Equilibrium Definition Given an initial population distribution $\{L_o\}_{o=1}^D$, market amenities $\{B_d\}_{d=1}^D$, firm amenities $\{b_f\}$, search costs $\{\delta_d\}$, and origin-destination migration costs $\{\tau_{od}\}$, a *spatial oligopsonistic Nash-Bertrand* equilibrium is: (1) a household labor supply curve for each origin $n(w_f, W_{fd}, W_d, \mathbf{W_o})$, (2) firm wages $\{w_f\}$, (3) self-employment wages $\{w_{sd}\}$, (4) quantities of labor $\{n_f\}$, $\{n_{sd}\}$ and $\{n_{fdo}\}$, (5) profits (Π), and (6) aggregate wage indices for each origin $\{\mathbf{W_o}\}$ and market level wage indices $\{W_{fd}\}$ and $\{W_d\}$ for each destination that satisfy the following conditions:

1. Given wages $\{w_f\}$ and $\{w_{sd}\}$, search costs $\{\delta_d\}$, migration costs $\{\tau_{od}\}$, amenities $\{B_d\}$ and $\{b_f\}$ and profits (Π), household optimization implies the labor supply curve $n(w_f, W_{fd}, W_d, \mathbf{W_o})$ for each origin o.

2. For every firm f in location d: given competitor wages $\{w_{-f(d)}\}$, the self-employment wage w_{sd}, the aggregate wage indices $\{\mathbf{W_o}\}$ from each origin and the labor supply curve from each origin $n(w_f, W_{fd}, W_d, \mathbf{W_o})$, firm f's optimization yields wage w_f and employment n_f.

3. Firm wage decisions are consistent with the market $\{W_{fd}\}$, $\{W_d\}$ and aggregate $\{\mathbf{W_o}\}$ wage indices for each origin, as well as profits (Π)

4. Markets clear $n_f = \sum_o n_{fdo} \, \forall f, d$, $n_{sd} = \sum_o n_{sdo} \, \forall d$, and $\sum_d \left(n_{sd} + \sum_f n_{f(d)} \right) = \sum_o L_o$.

5.1 Markdowns in the Model

The labor supply elasticity to firm f in market d is given by

$$\varepsilon_f = \eta \left(1 - s_f\right) + \gamma s_f \left(1 - s_d\right) + \theta s_f s_d \left(1 - \frac{\sum_o s_{od}^2 L_o}{s_m(d)}\right) \tag{6}$$

Where s_f is firm f's share of firm-employment its local market, d, s_d is the share of workers in d who are employed in a firm, s_{od} is the share of workers from o who migrate to d and $s_m(d)$ is market d's share of total employment. These shares can be expressed as

$$s_f = \left(\frac{b_f w_f}{W_{fd}}\right)^{\eta} \quad ; \quad s_d = \left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma} \quad ; \quad s_{od} = \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta} \quad ; \quad s_m(d) = \sum_o s_{od} L_o$$

The labor supply elasticity in (6) affects wages through the markdown (μ_f) that firms pay on workers' marginal product

$$w_f = \mu_f \, mrpl_f \qquad ; \qquad \mu_f = \frac{\varepsilon_f}{1+\varepsilon_f} \tag{7}$$

Where $mrpl_f = \alpha A_f n_f^{\alpha-1}$ is the marginal revenue product of labor. In the case of perfect competition, $\varepsilon = \infty$, the markdown is one and workers are paid their marginal product. When firms have market power, $\varepsilon < \infty$ and workers are paid a fraction of their marginal product. Thus, lower values of ε_f imply that a firm will pay its workers a lower share of their marginal product.

For an atomistic firm $(s_f \to 0)$, equation (6) reduces to η–the monopsony limit. These firms have no local wage-setting power, and the elasticity to these firms is governed entirely by the preference parameter η. Under the assumption that $\eta > \gamma > \theta$, wages will be closest to the competitive level in small firms.

Ignoring the final term in (6) momentarily, the labor supply elasticity in a market with a single firm would simplify to $\gamma(1 - s_d)$–the oligopoly limit. This means that labor is less elastic in markets with less self-employment and wages are further from the competitive level. When there is more than one firm in the market, the firm's labor supply elasticity is a weighted average of η and $\gamma(1 - s_d)$. Larger firms have more weight on the latter term and markdown their wages more than small firms.

In a world with homogeneous migration costs $(\tau_{od} = \tau \, \forall \, o, d)$, s_{od} is invariant across markets and the final term in (6) reduces to $\theta s_f s_d(1 - s_m)$. This implies that firms in larger markets will have lower labor supply elasticities than those in smaller markets. Intuitively, this is because there is less labor to pull from other labor markets.[24] When τ_{od} varies across markets, the final term in (6) is like a Herfindahl-Hirschman index of immigration concentration. Labor is less elastic in locations with more immigration.

[24]It follows that for a single labor market $s_m = 1$ or when migration costs are infinite $\tau_{od} = 0 \, \forall \, o, d$, the labor supply elasticity reduces to $\varepsilon_f = \eta\,(1 - s_f) + \gamma s_f\,(1 - s_d)$. When there is a continuum of labor markets (6) reduces to $\varepsilon_f = \eta\,(1 - s_f) + \gamma s_f\,(1 - s_d) + \theta s_f s_d$.

Comparative Statics Under the assumption that $\eta > \gamma > \theta$, the implications of equation (6) can be summarized as follows

1. For any two firms f' and f'' in the same market d such that $s'_f > s''_f$, then $\mu_{f''} > \mu_{f'}$. That is wages are less competitive in the larger firm.

2. For any market d, holding fixed the firm shares $\{s_f\}$ and immigration shares $\{s_{od}\}$ and comparing two equilibria with $s'_d > s''_d$, then $\mu_{f(d')} \leq \mu_{f(d'')}$, with strict inequality whenever $s_f > 0$. That is, wages are less competitive in markets with less self-employment.

3. For any two firm in different markets $f(d')$ and $f(d'')$ such that $s_{f(d')} = s_{f(d'')}$ and $s_{d'} = s_{d''}$ and $\sum_o s^2_{od'} L_o/s_m(d') > \sum_o s^2_{od''} L_o/s_m(d'')$, $\mu_{f(d')} \leq \mu_{f(d'')}$, with strict inequality whenever $s_{f(d')} > 0$ and $s_{d'} > 0$. That is, wages are less competitive in markets with a higher concentration of immigration.

The term 'immigration concentration,' in result (3) is somewhat misleading in that immigration includes non-migration. That is, a location can have a high immigration concentration if either it receives a large share of the migrants from other locations or it has very little emigration.

Comparison with Berger et al. (2022) This model differs from that in Berger et al. (2022) on two key dimensions: there is self-employment and there is a discrete number of markets. These modelling choices reflect differences in context. As shown in Table 3, 85% of prime-aged workers in Tanzania are engaged in self-employment versus just 10% in the US. Here, a labor market is a district whereas in their context it is a commuting zone crossed with an industry. I define a labor market as a district for two reasons. The first is due to data limitations: districts are the most disaggregated level of geography that is available. The second is due to context. It is not obvious that firms in different industries are not competing for the same workers. Moreover, self-employment competes with all industries. If I were to ignore self-employment and assume a continuum of markets, equation (6) would reduce to $\varepsilon_f = \eta\,(1 - s_f) + \theta s_f$. This still differs from their formulation, $\varepsilon_f^{BHM} = (\eta^{-1}(1 - s_f^{WB}) + \theta^{-1} s_f^{WB})^{-1}$, slightly because I have assumed Bertrand rather than Cournot competition. Where s_f^{WB} is the firm's wage-bill rather than employment market share. However, these two formulations yield quantitatively similar predictions for wage markdowns.

5.2 Misallocation

I now turn to each of the three sources of labor misallocation in the model–wage markdowns, search costs, and migration costs–to quantify their role in both welfare and output.

Wage Markdowns The estimated wage markdown depends upon the type of competition that the researcher assumes. These are summarized as follows

$$\varepsilon_f = \begin{cases} \infty & \text{Perfect Competition} \\ \eta & \text{Monopsonistic Competition} \\ \eta + s_f\left(\gamma - \eta\right) + s_f s_d\left(\theta - \gamma\right) & \text{Local Oligopoly} \\ \eta + s_f\left(\gamma - \eta\right) + s_f s_d\left(\theta\left(1 - \frac{\sum_o s_{od}^2 L_o}{s_m(d)}\right) - \gamma\right) & \text{Spatial Oligopoly} \end{cases}$$

Under perfect competition, the firm is a price taker and markdowns are zero. Non-uniform wages across firms will arise through variation in the labor supply curve to each firm. Under monopsonistic competition, the firm internalizes how a change in their wage affects their own labor supply, but they do not internalize how a change in their wage affects other wages in the market, *i.e.* $\partial W_{fd}/\partial w_f = 0$. This implies a uniform markdown across firms.

Under local oligopoly, the firm additionally internalizes how a change in their own wage affects the labor supply in the rest of the local market, but does not internalize how it affects migration decisions, *i.e.* $\partial \mathbf{W_o}/\partial w_f = 0$. As discussed in Section 5.1, markdowns will be larger in bigger firms and in markets with less self-employment. Under spatial oligopoly, the firm internalizes how a change in their wages affects migration decisions. This leads to less competitive wages in markets with a higher density of immigration.

Search Costs Individuals are misallocated into self-employment whenever

$$1 > \frac{e^{\zeta_{sd}(\omega)} w_{sd}}{\max_{f \in d} \left\{ e^{\zeta_{fd}(\omega)} b_f w_f \right\}} > \delta_d$$

This inequality states that the share of self-employment that is misallocated is higher in markets with larger labor market frictions. As the friction decreases ($\delta_d \to 1$), the share of workers in self-employment will decline. This causes the average productivity of those workers who remain in self-employment to increase. Thus, removing search costs will raise output by increasing the share of labor in firms and by increasing the productivity of the self-employed. It follows from equation (3) that removing this type of friction will have an unambiguous positive effect on welfare.

Migration Costs This type of friction also acts like a tax on welfare. If workers are misallocated across space, than any reduction in migration costs that causes a worker to migrate will cause welfare to increase. However, its effect on output is ambiguous. With no migration costs, labor supply to each firm can be expressed as

$$n_f = \left(\frac{b_f w_f}{W_{fd}} \right)^{\eta} \left(\frac{\delta_d W_{fd}}{W_d} \right)^{\gamma} \left(\frac{B_d W_d}{\mathbf{W}} \right)^{\theta}$$

Where $\mathbf{W} = \left[\sum_d \left(B_d W_d \right)^{\theta} \right]^{\frac{1}{\theta}}$. To fix ideas, ignore the self-employment margin momentarily. With no migration costs, the labor supply to each firm is independent of the origin of the worker. Labor supply is productively inefficient whenever $b_f \neq 1$ or $B_d \neq 1$. If the migration frictions were such that they kept workers out of high amenity jobs or locations, then removing them will cause output to fall. To see why this can be true, consider the labor supply to each market with migration costs

$$n_d = \left[\sum_o \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}} \right)^{\theta} L_o \right]$$

To fix ideas, suppose that there are two locations: a low-amenity high-productivity city and a high-amenity low-productivity village. First consider an initial population distribution such that all workers were born in the village. Comparing the equilibria with and without migration costs, the latter would have a higher share of the population in the city and output would be higher.

Now consider an initial population distribution such that all workers were born in the city. With migration costs, some people will migrate to the village because it has a high amenity and the marginal product of labor there will be very high. Without migration costs even more people will leave the city and total output will decline. Hence, removing migration costs may cause output to fall if productivity and amenities are negatively correlated.[25]

6 Minimum Wage Compliance

The aim of this section is to show that the minimum wage law had an effect on wages. I consider two measures of non-compliance: the share of workers paid below the minimum wage and the gap between the minimum wage compliant wage-bill and the firm's wage bill. Following Marshall (2023), I define the rate of Employment Non-Compliance (ENC) for firm f in industry i in district d in year t to the minimum wage law as

$$ENC_{fidt} = \frac{\sum_{r \in R_t} n_{rfidt} \, \mathbf{1}\left[\underline{w}_i - w_r > 0\right]}{\sum_{r \in R_t} n_{rfidt}}$$

Where $r \in R_t$ are the set of wage ranges which vary across years with the set $R_t = [o, \infty) \, \forall t$ and n_{rfidt} is the number of workers with wages in range r. $\underline{w}_i$ is the industry minimum-wage that applies to firm f and w_r is the average wage of workers in wage range r.[26] The term ENC_{fidt} can be interpreted as the share of workers paid below the minimum wage level. An alternative measure is the

[25] In each of these hypothetical examples, it is guaranteed that the labor share will change with a measure of workers. When the set of workers is discreet, the number of workers in the initial location will be weakly lower. This counter-intuitive result is not necessarily a shortcoming of the model. Consider that 20-25% of young people emigrate from rural areas to cities as young adults (Young, 2013). In doing so, they may not have complete information about the job opportunities or living conditions in the city to which they are migrating and may lose their local land rights. Thus they may by trapped in the city despite a preference to return home.

[26] Assuming a uniform distribution between each range r's lower-bound wage ($\underline{w}_r$) and upper-bound wage ($\bar{w}_r$). Under this distributional assumption, $w_r = (\underline{w}_r + \bar{w}_r)/2$. When the minimum wage falls between these two numbers, the number of workers in range r who are paid below the minimum wage is unknown. The interested reader is directed to Marshall (2023) who considers several methods for counting these workers.

gap between the firm's current wage-bill and the minimum wage compliant wage-bill defined as[27]

$$GAP_{fidt} = \frac{\sum_{r \in R_t} n_{rfidt} \min\{0, \underline{w}_i - w_r\}}{\sum_{r \in R_t} n_{rfidt} w_r} \tag{8}$$

When multiplied by 100, GAP_{fidt} can be interpreted as the percent by which a firm would need to raise its wages to be fully compliant with the minimum wage law. These measures are typically used to assess the bite of the minimum wage in the pre-policy period. Here, I use them to assess the degree of compliance once the minimum wage law was enacted. I estimate the rate of non-compliance as an event study, limiting my attention to the period 2005-2013 to avoid changes caused by the 2014 reform.

$$Y_{fidt} = \sum_t \delta_t + \boldsymbol{\mu}_i + \boldsymbol{\lambda}_d + \varepsilon_{fidt} \tag{9}$$

Where $\boldsymbol{\mu}_i$ are industry fixed effects and $\boldsymbol{\lambda}_d$ are district fixed effects. I exclude an intercept term so that the δ_t coefficients can be interpreted as the share of employment below the minimum wage level in each year. I plot the ENC and GAP coefficient estimates in Figure 4. In 2007, roughly 19% of workers earned a wage below the proposed minimum wage levels. When the law was enacted in 2010, the share fell to 10%, representing a 47% rate of compliance, consistent with the findings of Bhorat et al. (2017) for sub-Saharan African countries. I find a similar pattern when using the GAP measure. By 2013, the share of workers paid below the minimum wage falls below 5%. The declining rate of non-compliance is likely attributable to two factors. First, inflation reduces the real cost of employing workers at the minimum wage (Kaur, 2019). Second, even in developed countries, where enforcement is strongest, firms are not fully compliant with unforeseen increases in the minimum wage (Clemens and Strain, 2022). Taken together, this evidence suggests that firms adjusted wages in response to the legislation.

[27] Card and Krueger (1994) first used this measure to estimate the positive employment effects of New Jersey's minimum wage law on fast food workers. It remains a useful measure for analysis at the firm level and has been taken up more recently in minimum wages studies in the UK (Draca, Machin, and Van Reenen, 2011) and Germany (Dustmann, Lindner, Schönberg, Umkehrer, and vom Berge, 2020).

7 Estimation of the Structural Parameters

In this section I estimate the full set of parameters that govern firm and worker behavior in the model: the three elasticities $(\eta,\ \gamma,\ \theta)$, firm and market amenities $\{b_f\}$, $\{B_d\}$, search costs $\{\delta_d\}$ and migration costs $\{\tau_{od}\}$, and firm and self-employment productivity $\{A_f\}$ and $\{A_{sd}\}$. I begin with the interactions between firms and build outwards.

Between-Firm Elasticity (η) I estimate η via the reduced form labor supply elasticity. As shown by Berger et al. (2022), these estimates do not account for the strategic interaction between firms, making markets appear less competitive than they are. However, the strategic response to a marginal change in the wage offer of a firm is decreasing in the size of the firm making the change. In particular $\lim_{s_f \to 0} \varepsilon_f = \eta$. Put differently, for very small firms, only the monopsony channel is relevant–they have no local wage setting power. Hence, the structural and reduced form estimates will align for very small firms.

Since I cannot match firms across years, I predict the exposure to the minimum wage, $\widehat{\text{GAP}}_{fidt}$, in the pre-policy period, 2005-2007, for firm (f) in industry (i) in district (d) in year (t) via a random forest regression algorithm. GAP_{fidt} is the GAP measure (8) defined in Section 6, and represents the fraction by which a firm would need to raise its wages to be fully compliant. A value of zero indicates that the firm was already paying each of its workers a wage greater than the minimum wage. The set of independent variables include s_f, $s_f s_m$, $s_f s_d s_m$, an indicator for each sector, district, and region, the log Herfindahl-Hirschman Index, the log number of firms, the log employment share in the largest industry, log total employment in the district, log firm employment in the district, an indicator for whether the firm is privately owned, in a tradeable industry, in a non-agricultural industry, and the year. The inclusion of the year allows me to predict the exposure of each firm in 2009. I then follow the same procedure to predict the employment weighted exposure to the minimum wage, the ENC.

I then estimate the within-market elasticity via Instrumental Variables (IV), using information from both the GAP and ENC.

$$\log n_{fd} = \beta_0 + \beta_1 \log w_{fd} + \beta_2 \log w_{fd} \times s_f + \beta_3\, s_f s_d \times \log w_{fd} + \Gamma X_{fd} + \epsilon_{fd} \tag{10}$$

Where the endogenous variables are $\log w_{fd}$, $s_f\ \log w_{fd}$, and $s_f s_d\ \log w_{fd}$. X_{fd} includes controls for the firm's employment share, s_f, the firm's share of market employment, $s_f s_d$, log total employment, inclusive of self-employment, in district d, the log Herfindahl-Hirschman Index, and dummy variables for each region. I report the estimation results of (10) in Table 7. η can be interpreted as the coefficient on log wage. Column 1 reports the OLS estimates. In columns 2-4, I vary the set of included instruments. Column 2 restricts the set of instruments to the interactions of $\widehat{GAP}$ with s_f and $s_f s_d$. The estimated elasticity is 3.9, but the F-stat is low. Column 3 restricts the set of instruments to the interactions of $\widehat{ENC}$ with s_f and $s_f s_d$. The estimated elasticity is lower at 2.5 and the F-stat is sufficiently high at 22. In column 4, I include both sets of instruments and the estimated elasticity falls to 2.1. The first-stage F-stat is lower with the inclusion of the additional instruments. I calibrate $\eta = 2.5$ to estimate the rest of the parameters and use $\eta = 3.9$ as a robustness.

Firm Amenities (b_f) To estimate the employment elasticity, I first need to calculate W_{fd}, which is a function of $\{b_f\}$. I calibrate $\eta = 2.5$ using an intermediate value from the estimates in Table 7. I then estimate b_f for each firm via the following procedure. I begin with the initial guess of $b_f^0 = 1\,\forall\, f$. I then iterate on the following loop until $b_f^0 - b_f^1 < \varepsilon$, where ε is the tolerance threshold.

1. $W_{fd} = \left[\sum_{f\in d}\left(b_f^0 w_f\right)^{\eta}\right]^{\frac{1}{\eta}}$
2. $b_f^1 = (s_f^{\text{data}})^{1/\eta}\frac{W_{fd}}{w_f}$
3. $b_f^1 = \frac{b_f^1}{\sum_{f\in d} b_f^1}$
4. $b_f^0 = b_f^1$

Step 3 ensures that the value of W_{fd} is normalized across markets and that the b_f do not capture anything related to the aggregate market amenities.

Sector Elasticity (γ) Dividing the share of firm-employment by the share of self-employment and taking logs yields the expression

$$\log\left(\frac{s_d}{1-s_d}\right) = \gamma \log\left(\frac{W_{fd}}{w_{sd}}\right) + \gamma \log \delta_d$$

which I use to estimate γ instrumenting for log W_{fd}/w_{sd} using the firm-wage index evaluated at the minimum wage $W_{fd}(\underline{w})$.

$$\log\left(\frac{s_d}{1-s_d}\right) = \beta_0^d + \beta_1^d \log\left(\frac{W_{fd}}{w_{sd}}\right) + \Gamma^d X_d + \epsilon_d \tag{11}$$

Where W_{fd} is calculated using the firm amenities estimated above with η =2.5. Γ^d includes controls for the district's log share of total employment, s_m, an indicator for whether the district is urban. I report the estimation results of (11) in Table 8. γ can be interpreted as the coefficient on log W_{fd}/w_{sd}. Column 1 reports the OLS estimates. In columns 2-7 I report the IV estimates under various values of η. For each iteration, I re-estimate the implied firm amenities for the calibrated between-firm elasticity. Column 2 reports the estimated value of γ (1.5) under the preferred η calibration. Across each of the various calibrations of η the ratio of γ/η is relatively constant between 0.46 - 0.6. This implies that changing the value of η will move the point estimate of markdowns in the smallest firms, but will not affect the range of markdowns.

The F-stat in the preferred specification is 14, but it ranges from 118 for $\eta = 1$ to 4 for $\eta = 5$. The reason for this is that when $\eta = 1$, workers are very inelastic, so to match the observed labor shares, the b_f do not need to vary as much, and they will account for less of the variation in W_{fd}. On the other hand, when $\eta = 5$, workers are very elastic with respect to firm wages and the b_f need to be much larger to match the labor shares. This implies that more of the value of W_{fd} will be determined by the firm amenities and hence the minimum wage instrument will be less correlated with the wage index. For the remaining estimation, I calibrate $\gamma = 1.5$.

search costs (δ_d) To calculate each market wage index, W_d, I need to calculate δ_d. Having estimated η and γ, this can be solved for explicitly by inverting the market's firm-employment share, s_d.

$$\delta_d = \left(\frac{s_d}{1-s_d}\right)^{\frac{1}{\gamma}} \frac{w_{sd}}{W_{fd}}$$

Between-Market Elasticity (θ) I estimate θ via Poisson using the migration flows between regions.[28] The model-based relationship between migrants and non-migrants for any o can be expressed as

$$\log\left(\frac{n_{od}}{n_{oo}}\right) = \theta \log\left(\frac{W_d}{W_o}\right) + \theta \log\left(\frac{B_d}{B_o}\right) + \theta \log \tau_{od}$$

The main estimating equation is then the empirical equivalent:

$$n_{od} = n_{oo} \exp\left[\theta \log\left(\frac{W_d}{W_o}\right) + \alpha_1 \log \text{dist}_{od} + \alpha_2 \log \text{stock}_{od} + \mu_o\right] + \varepsilon_{od} \tag{12}$$

Where n_{od} and n_{oo} are the annualized number of migrants from o to d and non-migrants, respectively in the 2012 census. W_d and W_o are the market wage indices for the destination and origin, respectively. I calculate these as an employment-weighted average of the values for each district in the region. dist_{od} is the distance in kilometers between o and d, stock_{od} is the stock of migrants in d from o, and μ_o are origin fixed effects. I instrument for W_d in (12) using the equivalent CES aggregator of the applicable minimum wages for each firm defined as

$$\log\left(\frac{W_d(\underline{w})}{W_o(\underline{w})}\right) = \log\left(\frac{\delta_d W_{fd}(\underline{w})}{\delta_o W_{fo}(\underline{w})}\right)$$

Where the expression simplifies because there is no minimum wage in self-employment.

To control for variation in prices across locations, I spatially deflate wages using the LSMS price index. I report the estimation results for (12) Table 9. θ can be interpreted as the coefficient on W_d/W_o.

[28] This is a higher level of aggregation than that used in the rest of the analysis. I use regional rather than district migration flows because the latter is more sparse and susceptible to over-fitting (Dingel and Tintelnot, 2021). An observation of zero migrants between a district pair may be caused by sample size limitations. That is, there may be some migrants between the pair, but none that were in the sampling frame. As a robustness exercise, I estimate θ using district migration flows in Section 7.1.

Columns 1-3 report the estimates using nominal wages, while columns 4-6 use spatially deflated wages. The reduced form estimate for the minimum wage in columns 1 and 4 is 0.3, suggesting the migrants are aware of the minimum wage policy and respond to it when making migration decisions. The preferred estimate for the migration elasticity in column 6 is 1.4. This value is larger than the Poisson estimate in column 5, 1.0. The larger coefficient in column 6 implies that wages are negatively correlated with amenities: locations with higher amenities can offer lower wages. This finding is consistent with Rosen (1974) but conflicts with recent findings in Africa (Gollin, Kirchberger, and Lagakos, 2021).

Migration Costs (τ_{od}) Following a similar procedure to Bryan and Morten (2019), I use the assumption of symmetric migration costs to express τ_{od} in terms of the migration shares.[29]

$$\tau_{od} = \left(\frac{s_{od}}{s_{oo}} \times \frac{s_{do}}{s_{dd}} \right)^{\frac{1}{2\theta}}$$

This implies the two principle assumptions on the migration costs. First $\tau_{oo} = 1$ is true for all θ. Second, no migration between any pair of markets implies that the migration cost is infinite.

I estimate s_{od} using migration flows between districts in the 2012 census. I observe two types of migration flows in the census, those in the past year and those since birth. Neither reflect the lifetime decision to migrate which is what I aim to capture with the migration costs. To close this gap, I calculate the implied migration shares that I would have observed over a ten-year period following the procedure outlined in Appendix B.2.

Market Amenities (B_d) The amenity in each market can be solved for as an implicit function by rearranging $s_m(d)$:

$$B_d = \frac{s_m^{\text{data}}(d)^{\frac{1}{\theta}}}{W_d} \left[\sum_o L_o \left(\frac{\tau_{od}}{\mathbf{W_o}} \right)^{\theta} \right]^{\frac{-1}{\theta}}$$

[29] I derive this expression explicitly in Appendix B.2. Bryan and Morten (2019) express the migration costs as $\log \pi_{od} - \log \pi_{oo} + \log \pi_{do} - \log \pi_{dd} = 2\theta\tau_{od}$. This expression is equivalent to that but is well-defined for the case when $\pi_{od} = 0$.

I estimate B_d following an iterative procedure similar to that used to estimate b_f. I begin with an initial guess of $B_d^0 = 1\,\forall\, d$. I then iterate on the following loop until $abs(B_d^0 - B_d^1) < \varepsilon^B$

1. $\mathbf{W_o} = \left[\sum_d \left(\tau_{od} B_d^0 W_d\right)^\theta\right]^{\frac{1}{\theta}}$
2. $s_m^{\text{model}} = \left(\frac{\tau_{od} B_d^0 W_d}{\mathbf{W_o}}\right)^\theta$
3. $B_d^1 = \frac{s_m^{\text{data}}(d)^{\frac{1}{\theta}}}{W_d}\left[\sum_o L_o\left(\frac{\tau_{od}}{\mathbf{W_o}}\right)^\theta\right]^{\frac{-1}{\theta}}$
4. $B_d^0 = B_d^1\left(\frac{s_m^{\text{data}}(d)}{s_m^{\text{model}}(d)}\right)^\rho$

Where $\rho \in (0,1)$ in the final step dampens the updating procedure to improve the rate of convergence.

Firm and Self-Employment Productivity $\{A_f, A_{sd}\}$ The productivity in each firm can be solved for by rearranging the equilibrium expression for wages in each firm

$$A_f = \left(\frac{w_f}{\alpha \mu_f}\right) n_f^{1-\alpha}$$

I calibrate $\alpha = 0.65$ from Gollin (2002). However, the value of A_f will vary based on the implied value of μ_f. Hence, when I assume other forms of competition between firms, I change my estimate of A_f.

The value of self-employment productivity is invariant to the assumption on firm competition and is given by

$$A_{sd} = w_{sd} n_{sd}^{1-\alpha}$$

Model Fit Table 1 summarizes the estimation results from the above procedure. I report the key moments in the data and the simulated model in Table 10. With firm and location amenities, the model is able to match the labor share in each firm and in self-employment in each location precisely.

7.1 Robustness

Alternative η Estimation To assess whether my estimate of the between-firm elasticity is dependent upon the estimation method, I follow the two-step procedure in Berger et al. (2022) to estimate η. In

Table 1: Model Calibration and Estimation

Parameter	Value	Parameter	Mean	Range
α	0.65	b_f	1.00	[0.046, 19.462]
η	2.50	B_d	1.42	[0.08, 13.373]
γ	1.50	A_f	0.02	[0.002, 0.769]
θ	1.40	A_{sd}	0.17	[0.028, 0.565]
		δ_d	0.03	[0.004, 0.116]
		τ_{od}	0.03	[0, 1]

the first step, I use the predicted instruments $\widehat{GAP}$ and $\widehat{ENC}$, denoted by $\hat{Z}$ to estimate log-wages and log-employment in 2010 as a function of the firms' market share interacted with the instrument

$$y_{fd} = \beta_0^y + \beta_1^y s_f + \beta_2^y \hat{Z}_f + \beta_3^y s_f \times \hat{Z}_f + \beta_4^y s_f s_d \times \hat{Z}_f + \beta_4^y s_f s_d + \Gamma^y X_{fd} + e_{fd}^y \tag{13}$$

Where the dependent variable y_{fd} takes log-average-wage in the firm and log-employment. To be consistent with the estimation in section 7, in X_{fd} I include region fixed effects, the log district employment and the log Herfindahl-Hirschman Index. The reduced form labor supply elasticity is calculated by differentiating (13) with respect to the instrument $\hat{Z}_f$

$$\hat{\varepsilon}(s_f) = \frac{\partial \log n_{f(d)} / \partial \hat{Z}_f}{\partial \log w_{f(d)} / \partial \hat{Z}_f} = \frac{\hat{\beta}_2^n + \beta_3^n s_f + \beta_4^n s_f s_d}{\hat{\beta}_2^w + \beta_3^w s_f + \beta_4^w s_f s_d} \tag{14}$$

The between-firm labor supply elasticity is then estimated by taking the limit of (14).

$$\hat{\eta} = \lim_{s_f \to 0} \hat{\varepsilon}(s_f) = \frac{\hat{\beta}_2^n}{\hat{\beta}_2^w}$$

Table 17 presents the estimation results for (13). Columns 1 and 2 use $\widehat{ENC}$ as the instrument. The estimated elasticity is 2.7, slightly higher than the value estimated via instrumental variables, but not statistically different. In columns 3 and 4, I use $\widehat{GAP}$ as the instrument. Again, the estimated elasticity is similar (4.8) to that estimated earlier. The standard errors are tighter when using the ENC versus the GAP instrument. This finding is consistent with the significantly higher first-stage F-statistic found when using the ENC in the instrumental variables estimation. This confirms that the within-market elasticity estimate is robust to alternative empirical specifications.

Alternative θ Estimation Estimation of the migration elasticity relies upon two key assumptions: that the census migration flows accurately reflect the true migration flows and that the migration cost is symmetric. I test these assumptions in Table 18. Column 1 reports the IV-Poisson estimation results from Table 9. Column 2 limits the sample to only pairs with non-zero migration flows. Almost all pairs have non-zero observed flows in the census and the results are unchanged. In column 3, I estimate migration costs with log distance between districts, the log stock of migrants in d from o and the log employment ratio. The estimated migration elasticity is slightly lower at 0.9. In columns 4-6, I repeat the same set of exercises using migration flows from the LSMS and ILFS. The baseline elasticity estimate is higher (2.1). More than 200 of the 552 region pairs have non-zero migration flows, and when I limit the sample to migrating pairs the elasticity estimate falls to 1.9. Finally, when I use the asymmetric migration costs, the elasticity is 1.3, nearly identical to the baseline estimation.

To assess the sensitivity of the estimation to the level of aggregation, in Table 19 I report the estimation results at the district level. The estimated migration elasticity under the assumption of symmetric migration costs is much lower at 0.7. When I use asymmetric migration costs, the estimate falls to 0.3. However, these results are likely to understate the true migration elasticity. Migrants may not have perfect information about the wages in a specific location, but may have a general idea. Hence their exact choice of location may not accurately reflect their sensitivity to the differences in wages between the origin and destination.

7.2 Discussion

Markdowns Across Space I plot the spatial distribution of wage markdowns in Figure 5. Average wage markdowns range from 0.66 in some rural districts to 0.71 in Dar es Salaam, implying that urban labor markets are more competitive than rural labor markets. This finding is made more stark when I plot the average worker markdown against population density in Figure 6. This finding is intuitive– firm employment is more concentrated in rural labor markets. However, this finding is slightly more precarious than that. As noted in Section 3, urban labor markets have less self-employment and higher immigration rates which make markets less competitive. To illustrate this point, in Figure 7, I plot the markdown curve for the average rural and urban district. The curve relates the firm's share of firm employment to its equilibrium wage markdown. The urban curve is below the rural curve, implying that for two firms of equal employment share, the firm in the urban market will pay workers a lower share of marginal product. In the figure, I show where the mean worker is in a rural and urban market. The average worker in a rural labor market is employed in a larger (by employment share) firm the average urban worker.

Rural-Urban Productivity Gap The persistently high share of workers engaged in low-productivity agriculture is a puzzle (Gollin et al., 2014). The gap between agricultural and non-agricultural earnings has been conflated with the rural-urban income gap since Kuznets (1973). When firms markdown wages, the income and productivity gaps are not the same. I report both the rural-urban and agricultural productivity gaps by worker type in Table 11. At the top of the table, I compare the earnings for the rural self-employed with those in several definitions of urban markets. For the self-employed, there is no markdown, so the earnings and productivity gap are the same. Self-employed workers in Dar es Salaam are 47% more productive than those in rural areas. Looking further down the table, the gap between rural wage-workers and those in Dar es Salaam is tighter. Workers in Dar es Salaam at 8.7% more productive but earn 10.8% more. Because wage markdowns are lower in urban areas, the income gap is narrower than the productivity gap. The often cited rural-urban income gap arises once we look at all workers together. The average urban worker earns 53% more than a rural worker, but is more than twice as productive. This implies that the rural-urban income gap actually understates the productivity gap. The reason for this is that there are more workers in self-employment in rural areas where they are much less productive than those in cities.

In Panel B, I instead consider the agricultural productivity gap. Non-agricultural wage workers earn 24% more than those in agriculture and the productivity gap is nearly identical. This suggests that the rural-urban income gap understates the agricultural productivity gap. However, when I look at all workers, the earnings and productivity gaps are nearly identical to those in Panel A.

Together, this evidence highlights two important findings. First, because rural labor markets are less competitive, the rural-urban income gap among the employed overstates the productivity gap. Second, the well-documented rural-urban and agricultural productivity gaps arise because of compositional differences. As shown in Table 3, rural labor markets have higher rates of self-employment and a larger share of that employment is in agriculture.

Migration Costs I plot the average migration cost in each district in Figure 8 using the estimated value of θ. While average migration costs are high (see table 1), the map is visually consistent with what one might expect. Migration costs are lowest in urban districts. The exception is the Zanzibar archipelago, but there is little migration between the islands and the mainland. Migration costs are also lower in the Northwest around Geita where there is a large gold mine.

Search Costs A firm can post a vacancy for two reasons–it is expanding and is creating new positions or an employee has left and they have not been able to fill the position. To abstract from the first type of vacancy, I plot the number of unfilled positions per employee against the estimated search costs in the top panel of Figure 9. The relationship is negative. Recall that higher values of δ_d imply lower frictions. The figure can then be read as locations with larger labor market frictions have more unfilled job vacancies. This makes intuitive sense. If the cost of finding a job is higher, there will be more unfilled positions.

In the bottom panel of Figure 9, I plot the number of hires per vacancy. Here the relationship is positive and can be interpreted as districts with smaller labor market frictions hire more workers per posted vacancy. This relationship also makes sense. Locations with search costs are able to fill more of their vacancies. It is worth noting here that the hiring and vacancy patterns are not used in any way to estimate any of the parameters in the model.

8 Decomposing Labor Misallocation

Two separate literatures have evolved in development that study labor market frictions. The first has aimed to show that profitable migration is not undertaken on account of high costs of migration (*e.g.*

Lagakos, 2020). The second aims to identify the main sources of job search costs (Abebe et al., 2020; Alfonsi et al., 2020). Yet there is little evidence on the relative importance of these two types of frictions. The model allows me to quantify the relative importance of these two types of frictions in a unified framework while also addressing a third source of misallocation: labor market power. I quantify the role of each friction on labor misallocation in Table 12. Column 1 reports statistics under the simulated baseline equilibrium. In column 2, I turn off labor market power. This has no effect on the labor supply curve, but all firms become price takers and are in effect now choosing quantities of labor. Moving to a competitive equilibrium increases total output by 4.8% and welfare by 1.3%. Both the rural-urban income gap and the wage to self-employment income gap rise. This result is driven by an increase in both urban and firm employment as a result of higher wages.

In columns 3 and 4, I reduce job search costs by 10%. Under the labor market power equilibrium, total output increases by 21.7% and welfare by 83%. Firm employment expands to 48% resulting in the wage to self-employment income gap flipping. Now the self-employed earn more than the average firm worker. When I move to the competitive equilibrium the income pack falls back closer to unity. This suggests that even small reductions in job search costs are likely to have large effects on output.

In columns 5 and 6, I reduce migration costs by 10%. Under the labor market power equilibrium, total output falls by 4.2%. This result is driven by a 3.8% decline in urban employment. However, welfare increases by 122%. These results are consistent with workers choosing their place of work and residence taking into account the total value of amenities and wages. When I additionally move to the competitive equilibrium, total output rises by 0.5%. This reversal is caused by an increase in firm employment due to higher wages and a lower rate of emigration from urban areas. In both counterfactuals, the rural-urban income gap rises. This result is driven by an increase in rural labor supply which further drives down wages relative to those in cities. In the next section, I dig further into this counterintuitive result to explain precisely the mechanics behind it.

8.1 Reducing Migration Costs

In this section, I examine the effect of reducing migration costs when viewed through the lens of an inelastic labor supply curve. The idea of reallocating labor into cities as a means to development was first put forth by Kuznets (1973). Recently, the idea that labor is spatially misallocated due to migration frictions has been studied experimentally (Bryan et al., 2014; Brooks and Donovan, 2020). These studies focus on the idea of moving labor from rural areas to cities.

In Table 13, I take this idea literally by considering asymmetric reductions in migration costs. In column 3, I reduce the cost of migrating into urban districts only. The urban share of employment rises from 14.2% at baseline to 25.7% and total output rises by 3.8%. This counterfactual is most similar to the experimental design of giving bus tickets to rural workers in Bangladesh to seasonally migrate to Dhaka (Bryan et al., 2014). In column 4, I consider the opposite design in which I only reduce the cost of migrating out of urban districts. Output falls, but only by 1.4%. This suggests that most of the output decline is coming from rural to rural moves rather than from emigration out of urban markets. I investigate this hypothesis in column 5 by reducing migration costs between rural districts only. Total output falls by 5%–more than under a symmetric decline in migration costs. However, the urban employment share falls by more than under the symmetric reduction. This can be attributed to the general equilibrium effects that make some rural locations more attractive to urban workers because of emigration out of that location while there is no counteracting effect to pull workers into urban districts because the migration costs have not declined.

Congestion and Agglomeration An alternative explanation for the migration cost result is that the model does not account for congestion or agglomeration forces. If amenities or productivity are endogenous, this may affect the counterfactual predictions. Following Bryan and Morten (2019), I make amenities and productivity endogenous as

$$B_d = \bar{B}_d L_d^{\lambda} \quad ; \quad A_i = \bar{A}_i L_d^{\phi}$$

Congestion forces are present whenever $\lambda < 0$. This implies that more people reduce each person's ability to enjoy public goods. Similarly, agglomeration forces are present whenever $\phi > 0$. This implies that firms are more productive when there are more people in a location. This captures, for example, knowledge spillovers from change encounters of workers in different firms. This is a slight deviation from Bryan and Morten (2019), who use total human capital to measure agglomeration, but is consistent with Au and Henderson (2006).

I report the counterfactual equilibrium for a 10% decline in migration costs under various calibrations of λ and ϕ in Table 14. Column 1 restates the equilibrium outcomes with no congestion or agglomeration as shown in column 5 of Table 21. In column 2, I introduce congestion forces only by making amenities endogenous. I follow Bryan and Morten (2019) and calibrate $\lambda = -0.04$. Both total output and urban employment fall by less than in the baseline model. This result may seem surprising considering that congestion externalities are typically associated with urban areas. Here, congestion reduces total migration by reducing the utility gain of the marginal migrant. This causes fewer people to emigrate from urban areas. In column 3, I introduce agglomeration forces only by making productivity endogenous. I calibrate $\phi = 0.05$, the high end of values considered by Bryan and Morten (2019). Output and the urban employment share fall slightly more than in baseline. This result follows the same logic as that for congestion. Agglomeration increases the value of immigrating for the marginal migrant, increasing total migration. This also makes it even worse for people left behind because productivity is even lower. In column 4, I turn on both congestion and agglomeration forces. The agglomeration effect slightly wins out and the overall effect on output is an addition 0.2% decline relative to baseline. In column 5, I ask whether an impossibly high value of congestion could possibly recover an overall gain in output. To do so, I calibrate $\phi = 0$ and $\lambda = -0.5$. The urban share of employment is 1.2% higher than in the baseline case, and the fall in output is lower at 2.6%.

Compensating Differentials One might expect that if output is going down because the compensating differentials (amenities) in low productivity areas are too high, then equating them would recover the expected output gains. However, that is not quite the correct comparison. If there are no compensating differentials, then the initial population distribution will change as well, and that is the outcome from which to compare. In Table 20, I compare these outcomes turning off each type of amenity. Equating location amenities causes total output to rise by 1%. Equating job amenities actually causes output to fall by 4.9%. This finding is driven by a lower share of wage employment. When job amenities are equated, the high amenity firms that were attracting a disproportionate amount of labor before lose that edge. When I equate both types of amenities, output rises by just 0.5%.

Alternative Assumptions on Firm Competition To assess whether the migration cost result is driven by by how I model competition, in Appendix Table 21, I simulate a 10% reduction in migration cost under various assumptions on labor market competition. In columns 1 and 2, I report the baseline results assuming spatial labor market power. In columns 3 and 4, I re-estimate firm productivity assuming that firms are competing for workers in monopsonistic competition. That implies that firms internalize how a change in their own wage affects their own labor supply, but not how it affects the wages of other firms in their local market. This implies that all firms pay a fixed proportional markdown. When I reduce migration costs in column 4, the results are essentially unchanged. This is because most firms are quite small and effectively have no oligopoly power. In columns 5 and 6, I assume that firms are instead price takers and are choosing employment, implying that markdowns are zero. Output falls by 3.8%, less than in the labor market power framework.

Ignoring the Self-Employment Margin What would the counterfactual predictions about lowering migration cost look like if I ignored the self-employment sector? To answer this question, I run the same set of exercises in Table 21 as if I did not observe self-employment. Specifically, I remove self-employment from the choice set of workers. I report the estimation results in Appendix Table 22. A 10% reduction in migration costs causes total output to fall by 1.4%. This finding is independent of the form of labor competition that I assume.

To see why the form of competition doesn't matter, recall from equation (4) that when migration costs change, the firm's share of local employment will not change–s_f is constant. What will change is s_d because the marginal marginal-product differs between firm and self-employment. When there is no self-employment, $s_d = 1$, so when labor reallocates across markets, wages will change at the same rate in firms as within a market, regardless of the form of competition.

Is Dar es Salaam too big? Au and Henderson (2006) proposed that the optimal city employment follows an inverted u-shape, with the peak being determined by the ratio of value-added in manufacturing to services. While the authors were interested in whether Chinese cities were too small, the hypothesized that cities in other regions may be too big. In Table 15, I follow their logic to assess whether Dar es Salaam is too large. In 2012, Dar es Salaam had 1.7 million employed persons and a value-added ratio of 0.16. To put this in context, in Au and Henderson (2006) the typical Chinese city had a ratio of 1.4, with the most services-intensive cities having a value of 0.6. Their model would suggest that employment is approximately half of the optimal level as reported in column (5).

The value-added ratio may be confounded if self-employment is more heavily engaged in manufacturing. In column (7), I report the optimal city size under the assumption that all self-employment is in services and the optimal city size is slightly higher. When I assume that all self-employment is manufacturing in column (9), Dar es Salaam is too big. However, Tanzania's three other cities are still far too small. In column (3), I report the ratio of employment in services and manufacturing among the self-employed. The value is comparable to the value-added ratio among firms, suggesting that the estimates based on firm-employment are reasonable.

Taken together, these results suggest that it is not that Dar es Salaam is too large *per se*, but rather that the city's economy is overly concentrated in services (Gollin et al., 2016; Henderson and Kriticos, 2018). Au and Henderson (2006) do not directly address the reverse question of what is the optimal ratio of value-added between services and manufacturing given a city's size, however we can infer that a number closer to 0.6, that of China's most services intensive cities would be optimal.

8.2 Urbanization without Growth

The finding that further reducing migration costs will cause a fall in output is surprising. In the standard model, there is no tension between what is welfare improving and what is output improving. So what evidence is there that this is the right way to think about labor markets in sub-Saharan Africa? Over the past 60 years, the urban population in sub-Saharan Africa has grown at a faster rate than any other region in the world. At the same time, the region has seen the lowest growth in per-capita income. I plot these trends in figure 10. Urbanization without economic growth is atypical. East Asia has seen the fastest rate of growth over the period and has urbanized almost as much as sub-Saharan Africa.

If we are to take this model seriously, then it should do better at predicting the observed patterns of urbanization without growth in sub-Saharan Africa than the standard model in which migration responds to wage gaps alone.[30] I do this by beginning with the population distribution as it was in 2000 and simulating the model with infinite migration costs to find the baseline level of output in that year. I then slowly reduce migration costs towards their level in 2010 and beyond. I compare the change in urbanization and output under this model and the standard model in Figure 11. Urbanization increases from 9.5% in 2000 in both models. However, it flattens around 14% in the spatial labor market power model while it continues to rise in the standard model. Output on the other hand, initially rises in the spatial labor market power model, reaching a peak of 6%, before falling. In the standard model, output continues to rise, reaching a 17% gain when migration costs reach their 2010 level.

These results should be interpreted as suggestive preliminary evidence that this framework is better able to match the observed patterns of growth and urbanization in the data. Of course, the spatial distribution of both firms and productivity has changed over this period and that is not accounted for here. Moreover, it is likely that urbanization will continue to rise in the data as the number of firms in mega cities such as Dar es Salaam increase.

9 Conclusion

Economic development requires the shift of productive inputs out of agriculture and into manufacturing and services. That these transitions have been happening slowly in low-income countries is typically explained by economic frictions. In particular, labor markets in low-income countries differ from those in high income countries in that there are large gaps between rural and urban income, between

[30] Specifically, I compare the predictions of this model from those in a competitive equilibrium framework with no amenities. This implies that there is no gap between the welfare and output maximizing labor allocations.

wage and self-employment income, and there are high rates of self-employment. In this paper, I consider the role of labor market power in explaining these features. To do so, I construct a spatial general equilibrium model of monopsony that allows me to disentangle labor market power from migration and job search costs. I find that most workers are paid between 66-71% of their marginal product. Rural labor markets are less competitive than their urban counterparts. This implies that the earnings gap between rural and urban workers employed in firms overstates their productivity gap. Moving to a competitive equilibrium causes output to rise by 4.8%. Reducing migration costs, on the other hand, causes total output to fall. This finding is reconciled through the fact that workers choose where to live and work based on the total value of wages and amenities. The gap between rural and urban wages is not high enough to induce additional workers to move into those more productive areas. However, when I reduce migration costs and move to competitive wages, the gap closes enough to cause total output to rise.

References

Abebe, G., A. S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn (2020). Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City. *The Review of Economic Studies* (September), 1–32.

Abebe, G., S. Caria, and E. Ortiz-ospina (2021). The Selection of Talent Experimental and Structural Evidence from Ethiopia. *American Economic Review 111*(6), 1757–1806.

Albouy, D. (2008). Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas. Technical Report No. 14472, NBER.

Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2020). Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda. *Econometrica 88*(6), 2369–2414.

Almeida, R. and P. Carneiro (2012). Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics 4*(3), 64–89.

Amodio, F. and N. de Roux (2021). Labor Market Power in Developing Countries: Evidence from Colombian Plants.

Amodio, F., P. Medina, and M. Morlacco (2024). Labor Market Power, Self-Employment, and Development. Mimeo.

Armangué-Jubert, T., N. Guner, and A. Ruggieri (2023). Labor Market Power and Development. Mimeo.

Ashenfelter, O. C., H. Farber, and M. R. Ransom (2010). Labor Market Monopsony. *Journal of Labor Economics 28*(2), 203–210.

Au, C. C. and J. V. Henderson (2006). Are Chinese cities too small? *Review of Economic Studies 73*(3), 549–576.

Basu, A. K., N. H. Chau, and R. Kanbur (2010). Turning a blind eye: Costly enforcement, credible commitment and minimum wage laws. *Economic Journal 120*(543), 244–269.

Berger, D. W., K. Herkenhoff, and S. Mongey (Forthcoming). Minimum Wages, Efficiency and Welfare. *Econometrica*.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2022). Labor Market Power. *American Economic Review 112*(4), 1147–1193.

Bhorat, H., R. Kanbur, and B. Stanwix (2017). Minimum wages in sub-saharan Africa: A primer. *World Bank Research Observer 32*(1), 21–74.

Brooks, W. and K. Donovan (2020). Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua. *Econometrica 88*(5), 1965–1997.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Under-investment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica 82*(5), 1671–1748.

Bryan, G., S. Chowdhury, A. M. Mobarak, M. Morten, and J. Smits (2021). Encouragement and distortionary effects of conditional cash transfers. Discussion Papers 1085, Economic Growth Center at EliScholar – A Digital Platform for Scholarly Publishing at Yale.

Bryan, G. and M. Morten (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Card, D. and A. Krueger (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review 84*, 772–784.

Caria, S. and K. Orkin (2024). Barriers to Search and Hiring in Urban Labour Markets. *VoxDevLit 10*(1), 1–46.

Clemens, J. and M. R. Strain (2022). Understanding "Wage Theft": Evasion and avoidance responses to minimum wage increases. *Labour Economics 79*(March), 102285.

Derenoncourt, E., C. Noelke, D. Weil, and B. Taska (2021). Spillover Effects from Voluntary Employer Minimum Wages. Technical Report No. 29425, NBER.

Dingel, J. I. and F. Tintelnot (2021). Spatial Economics for Granular Settings. Technical Report No. 27287, NBER.

Donovan, K., W. J. Lu, and T. Schoellman (2023, 05). Labor Market Dynamics and Development. *The Quarterly Journal of Economics 138*(4), 2287–2325.

Draca, M., S. Machin, and J. Van Reenen (2011). Minimum Wages and Firm Profitability. *American Economic Journal: Applied Economics 3*(1), 129–151.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. vom Berge (2020). Reallocation Effects of the Minimum Wage. Discussion Paper CDP 07/20, Centre for Research and Analysis of Migration.

Employment and Labour Relations Act (2004). United Republic of Tanzania.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *Review of Economic Studies 81*(3), 1046–1070.

Felix, M. (2022). Trade, Labor Market Concentration, and Wages.

Franklin, S., C. Imbert, G. Abebe, and C. Mejia-Mantilla (2024, May). Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia. *American Economic Review 114*(5), 1382–1414.

Gollin, D. (2002). Getting Income Shares Right. *Journal of Political Economy 110*(2), 458–474.

Gollin, D. (2008). Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics 55*(2), 219–233.

Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth 21*(1), 35–70.

Gollin, D., M. Kirchberger, and D. Lagakos (2021). Do urban wage premia reflect lower amenities? Evidence from Africa. *Journal of Urban Economics 121*.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics 129*(2), 939–993.

Guillouzouic, A., E. Henry, and J. Monras (2024). Public Sector Pay in Spatial Equilibrium. Discussion Paper No. 19014, CEPR.

Harris, J. R. and M. P. Todaro (1970). Migration, Unemployment and Developmnent: A Two-Sector Analysis. *The American Economic Review 60*(1), 126–142.

Henderson, J. V. and S. Kriticos (2018). The Development of the African System of Cities. *Annual Review of Economics 10*, 287–314.

Henderson, J. V., M. Roberts, and A. Storeygard (2013). Is Urbanization in Sub-Saharan Africa Different? *The World Bank Research Observer* (June), 48.

Herrendorf, B. and T. Schoellman (2015). Why is measured productivity so low in agriculture? *Review of Economic Dynamics 18*(4), 1003–1022.

Hicks, J. H., M. Kleemans, N. Y. Li, and E. Miguel (2021). Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata. *Journal of the European Economic Association 19*(3), 1522–1555.

Hirsch, B., E. J. Jahn, A. Manning, and M. Oberfichtner (2022). The Urban Wage Premium in Imperfect Labor Markets. *Journal of Human Resources 57*(Special Issue Monopsony in the Labor Market), S111–S136.

Hsieh, C.-t. and B. A. Olken (2014). The Missing "Missing Middle". *Journal of Economic Perspectives 28*(3), 89–108.

Imbert, C. and J. Papp (2020). Costs and Benefits of Rural-Urban Migration: Evidence from India. *Journal of Development Economics 146*, 1–17.

Imbert, C., M. Seror, Y. Zhang, and Y. Zylberberg (2022). Migrants and Firms: Evidence from China. *American Economic Review 112*(6), 1885–1914.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review 109*(10), 3585–3616.

Kirchberger, M. (2021). Measuring internal migration. *Regional Science and Urban Economics*.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97–126.

Kuznets, S. (1973). Modern Economic Growth: Findings and Reflections. *The American Economic Review 63*(3), 247–258.

Lagakos, D. (2020). Urban-Rural Gaps in the Developing World: Does Internal Migration Offer Opportunities? *Journal of Economic Perspectives 34*(3), 174–192.

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

Lagakos, D., A. M. Mobarak, and M. Waugh (2023, May). The Welfare Effects of Encouraging Rural-Urban Migration. *Econometrica 91*(3), 803–837.

Lewis, W. (1954). Economic Development with Unlimited Supplies of Labour.

Magruder, J. R. (2013). Can minimum wages cause a big push? Evidence from Indonesia. *Journal of Development Economics 100*(1), 48–62.

Manning, A. (2003). The real thin theory: Monopsony in modern labour markets. *Labour Economics 10*(2), 105–131.

Manning, A. (2004). Monopsony and the efficiency of labour market interventions. *Labour Economics 11*(2), 145–163.

Manning, A. (2006). A Generalised Model of Monopsony. *The Economic Journal 116*(508), 84–100.

Manning, A. (2019). The Minimum Wage and Trade Unions. In *Monopsony in Motion: Imperfect Competition in Labor Markets*, pp. 325–359. Princeton University Press.

Mansoor, K. and D. O'Neill (2021). Minimum wage compliance and household welfare: An analysis of over 1500 minimum wages in India. *World Development 147*, 1–19.

Marshall, S. (2023, April). Minimum Wage Compliance and Migration: The Labor Market Effects of Tanzania's Sectoral Minimum Wage Bill. Mimeo.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–3890.

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

Poschke, M. (2024). Wage employment, unemployment and self-employment across countries. *Journal of Monetary Economics* (September), 103684.

Rani, U., P. Belser, and M. Oelz (2013). Minimum wage coverage and compliance in developing countries. *International Labour Review 152*(3), 381–410.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative Spatial Economics. *Annual Review of Economics 9*, 21–58.

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

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

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy 82*(1), 34–55.

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

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review 108*(8), 2015–2047.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics 12*, 525–546.

Yeh, C., C. Macaluso, and B. Hershbein (2022). Monopsony in the US Labor Market. *American Economic Review 112*(7), 2099–2138.

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

Figures & Tables

Table 2: Firm Monthly Wage and Employment Summary Statistics

Year	Firms	Minimum Wage (2)	Wage (3)	Wage + Inkind (4)	Employees (5)
2005	9035	0	258.1	302.6	19.8
		[0]	[244.8]	[311.3]	[155.2]
2006	11948	0	288.1	339.3	16.3
		[0]	[259.8]	[338]	[76.3]
2007	11056	0	259.8	310.9	19.9
		[0]	[241.2]	[315.2]	[205.5]
2010	8112	86.4	343.4	388.7	24.6
		[28.5]	[248]	[308.6]	[128.6]
2011	8648	75.9	332.9	377.4	24.9
		[24.3]	[250.5]	[301.1]	[132.5]
2012	9417	66.1	349.3	394.6	25.1
		[22.6]	[271.6]	[326.2]	[115.8]
2013	8802	61.3	357.2	391.9	26.4
		[20.3]	[253.5]	[296.9]	[129.1]

Notes: Reporting real average monthly wages in thousands of Tanzanian Shillings (TSH) in the EES. Standard deviation in brackets. Wages are deflated using the Tanzanian CPI but are not adjusted for spatial variation in the price level. Wages are weighted by firm weight and the number of employees at the firm and employment is weighted by the firm sampling weight.

Table 3: Employment by Type

	Employed (%)	*Employment Share (%)*		
		Wage Worker	Self-Employed	
			Total	Agriculture
	(1)	(2)	(3)	(4)
Panel A: *Tanzania*				
Rural	76.5	5.2	94.8	77.9
Urban	61.2	30.0	70.0	28.2
Panel B: *Countries*				
Tanzania	70.2	14.1	85.9	60.1
USA	65.6	90.6	9.4	0.4
Brazil	60.9	74.8	25.2	5.1
India	53.7	51.5	48.5	29.0
USA (1910)	60.8	72.2	27.8	16.2

Notes: Source: IPUMS International. Reporting the shares of employment by geographic designation in the 2012 Tanzania census (Panel A) and national aggregates in Panel B. The sample includes all individuals aged (15-65). Column (1) reports the share of individuals who are employed. Columns (2) and (3) report the employment share in wage-work and self-employment, respectively. Column (4) reports the share of all workers who are engaged in self-employment agriculture. The samples in Panel B are the 2012 Tanzania, 2010 USA, 2010 Brazil, 1910 USA censuses and a 2010 India employment survey.

Table 4: Self-Employment in Sub-Saharan Africa

Census	Employed (%)			*Self-Employment Share (%)*		
	All (1)	Urban (2)	Rural (3)	All (4)	Urban (5)	Rural (6)
Tanzania (2012)	70.2	61.2	76.5	85.9	70.0	94.8
Benin (2013)	55.8	54.9	56.7	84.1	77.2	90.1
Cameroon (2005)	43.9	34.7	54.6	78.0	57.3	93.2
Ghana (2010)	68.7	63.6	74.9	76.1	64.8	87.7
Guinea (2014)	58.8	49.8	64.6	87.4	71.7	95.1
Lesotho (2006)	42.4	51.2	39.4	33.9	18.9	40.7
Liberia (2008)	45.5	36.9	53.9	83.5	71.4	91.6
Malawi (2008)	58.1	54.5	58.8	77.1	47.5	82.5
Mali (2009)	58.2	50.5	60.9	71.4	54.0	76.3
Rwanda (2012)	60.2	56.2	61.1	80.4	46.8	86.9
Senegal (2013)	38.7	42.4	34.8	71.9	59.5	87.3
South Sudan (2008)	65.5	62.8	66.4	86.3	70.4	91.6
Sudan (2008)	41.8	42.8	41.2	60.9	41.3	72.6
Togo (2010)	69.6	64.1	74.0	83.9	71.2	92.5

Notes: Reporting the shares of employment by geographic designation in the Sub-Saharan African countries where available. Columns (4)-(6) report the share of employed persons who are engaged in self-employment. The sample includes all individuals aged (15-65). *Source:* IPUMS International.

Table 5: The Firm Size Distribution in Rural and Urban Markets

	Employment Share			Firm Share		
	Rural	Urban	p value	Rural	Urban	p value
Firm Size	(1)	(2)	(3)	(4)	(5)	(6)
1-9	11.6	15.4	0.136	61.8	62.0	0.962
10-49	24.0	34.2	0.120	30.1	32.3	0.617
50+	65.5	50.4	0.034	8.3	5.7	0.030

Notes: The table reports the average share of employment and firms by firm size in 2010 across districts by type. Reporting the p-value for the two-sided t-test that $\mu_{urban} = \mu_{rural}$. *Source:* EES.

Table 6: Migration Rates by Type

Sample	Year	Last 5 Years				2010 Only			
		Rural to		Urban to		Rural to		Urban to	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LSMS	2010	5.15	2.90	3.86	6.15	5.23	2.42	1.30	3.10
LSMS	2012	6.78	2.99	5.34	7.10	3.17	1.17	0.74	2.55
LSMS	2014	6.74	4.00	6.37	11.01	2.25	0.72	0.46	1.50
ILFS	2014	4.06	3.25	2.24	5.45	3.72	3.25	0.61	2.51

Notes: Reporting five-year migration rates in percent for each sample in columns (1)-(4) and the reported 2010 migration rate in columns (5)-(8). Rural-rural and urban-urban migration episodes exclude migration within the district. All values are weighted by the respective sample's survey weights. Urban districts are defined as those for which at least half of the population was living in an urban area.

Table 7: Between-Firm Elasticity IV Estimation Results

	Dependent Variable: log employment			
	(1)	(2)	(3)	(4)
log wage	0.405***	3.919***	2.474***	2.099***
	(0.082)	(0.683)	(0.381)	(0.406)
log wage $\times$ s_f	-0.058	-39.142***	-20.016**	-11.042*
	(2.539)	(13.337)	(8.428)	(5.974)
log wage $\times$ $s_f\, s_d$	12.598**	13.297***	11.479***	12.887***
	(5.419)	(4.390)	(4.426)	(4.427)
F-statistic		7.958	22.135	15.450
Firms	3880	3880	3880	3880
Controls	Y	Y	Y	Y
Instruments		$\widehat{\text{GAP}}$	$\widehat{\text{ENC}}$	$\widehat{\text{GAP}}$ & $\widehat{\text{ENC}}$
Estimation	OLS	IV	IV	IV

Notes: The table presents OLS and IV estimation results for the reduced form estimation of the between-firm elasticity η. s_f is the firm's share of firm-employment in the local market, s_d is the share of employment in the local market that is engaged in wage-work. Columns (1) reports the OLS estimates. Columns (2)-(4) vary the set of instruments. The $\widehat{GAP}$ instruments are $\{\widehat{GAP}, \widehat{GAP} \times s_f, \widehat{GAP} \times s_f s_d\}$ and the $\widehat{ENC}$ instruments are $\{\widehat{ENC}, \widehat{ENC} \times s_f, \widehat{ENC} \times s_f s_d\}$. Controls include s_f, $s_f s_d$, log total employment in district d (inclusive of self-employment), the log HHI, and region fixed effects. Robust standard errors clustered by district in parenthesis. Reporting the Kleibergen and Paap (2006) cluster robust F-statistic. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 8: Sector Elasticity IV Estimation Results

	Dependent Variable: $s_d/(1-s_d)$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\log W_{fd}/\log w_{sd}$	0.225**	1.537***	0.455***	1.157***	1.897***	2.521**	3.026*
	(0.087)	(0.494)	(0.099)	(0.337)	(0.679)	(1.130)	(1.688)
log s_m	-0.349**	-0.767***	-0.486***	-0.690***	-0.831***	-0.925**	-0.990**
	(0.142)	(0.221)	(0.101)	(0.176)	(0.269)	(0.368)	(0.468)
urban	2.312***	2.128***	1.393***	1.847***	2.431***	3.030***	3.571***
	(0.273)	(0.359)	(0.304)	(0.340)	(0.398)	(0.557)	(0.826)
F-statistic		14.560	117.856	22.399	10.434	6.157	3.855
Districts	103	103	103	103	103	103	103
η	2.5	2.5	1	2	3	4	5
Estimation	OLS	IV	IV	IV	IV	IV	IV

Notes: The table presents OLS and IV estimation results for the sector elasticity. γ can be interpreted as the coefficient on $\log W_{fd}/\log w_{sd}$. s_d and $s_m(d)$ are as defined in the main text. Columns (2)-(7) change the calibrated value of η. Firm amenities are re-estimated for each iteration. Robust standard errors clustered by region in parenthesis. Results are weighted by total district employment. Reporting the Kleibergen and Paap (2006) weak identification F-statistic. $^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Table 9: Migration Elasticity GMM Estimation Results

	Dependent Variables: n_{od}					
	Nominal			Spatially Deflated		
	(1)	(2)	(3)	(4)	(5)	(6)
$W_d(\underline{w})/W_o(\underline{w})$	0.337***			0.360***		
	(0.026)			(0.028)		
W_d/W_o		0.997***	1.082***		1.139***	1.398***
		(0.073)	(0.082)		(0.099)	(0.116)
$\log \tau_{od}$	0.424***	0.426***	0.422***	0.424***	0.433***	0.422***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
F-statistic			1879.844			1345.804
Region Pairs	552	552	552	552	552	552
Origin FE	Y	Y	Y	Y	Y	Y
η	2.5	2.5	2.5	2.5	2.5	2.5
γ	1.5	1.5	1.5	1.5	1.5	1.5
Estimation	Poisson	Poisson	IV-Poisson	Poisson	Poisson	IV-Poisson

Notes: The table presents Poisson and IV-Poisson estimation results for the migration elasticity θ at the regional level. τ_{od} is calculated using the Census migrants since birth. The exposure variable is the number of non-migrants. Results are weighted by destination population. Migrants are counted as the number of prime-aged individuals who moved in the past year in the 2012 census. $W_d(\underline{w})$ and $W_o(\underline{w})$ are the market wage indices calculated using the minimum wage and is the instrument for the wage ratio, W_d/W_o. Columns (1)-(3) use nominal wages, while columns (3)-(6) spatially deflate wages in each region. Robust standard errors in parenthesis. Reporting the Kleibergen and Paap (2006) weak identification F-statistic. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 10: Baseline Calibrated Model Fit

	Rural		Urban	
	Data	Model	Data	Model
	(1)	(2)	(3)	(4)
Wages				
Firms	1.458	1.458	1.325	1.325
Self-Employment	0.916	0.916	1.112	1.112
All Workers	0.967	0.967	1.199	1.199
Markdowns				
Firms	0.705	0.704	0.714	0.714
Self-Employment	1.000	1.000	1.000	1.000
All Workers	0.972	0.972	0.883	0.883
Employment Share				
firm	0.095	0.095	0.407	0.407
self-emp	0.905	0.905	0.593	0.593
All	0.858	0.858	0.142	0.142
Output				
Firms	0.260	0.260	0.165	0.165
Self-Employment	0.711	0.711	0.094	0.094
All Workers	0.971	0.971	0.259	0.259
Output per Worker				
Firms	3.187	3.192	2.857	2.857
Self-Employment	0.916	0.916	1.112	1.112
All Workers	1.131	1.132	1.823	1.823

Notes: Reporting the simulated and data moment averages weighted by number of workers.

Table 11: The Earnings and Output Gap

	Earnings		Output per Worker	
	Value	Gap	Value	Gap
	(1)	(2)	(3)	(4)
Panel A: *Rural-Urban Productivity Gap*				
Self-Employment				
Rural	1.000		1.000	
All Urban	1.314	1.314	1.314	1.314
Tier 2 Cities	1.252	1.252	1.252	1.252
Dar es Salaam	1.470	1.470	1.470	1.470
Wage-Employment				
Rural	1.596		3.508	
All Urban	1.715	1.075	3.702	1.055
Tier 2 Cities	1.431	0.897	3.088	0.880
Dar es Salaam	1.769	1.108	3.812	1.087
All Employment				
Rural	1.043		1.181	
All Urban	1.456	1.396	2.158	1.828
Tier 2 Cities	1.311	1.257	1.861	1.576
Dar es Salaam	1.600	1.534	2.487	2.106
Panel B: *Agricultural Productivity Gap*				
Self-Employment				
Agriculture	1.000		1.000	
Non-Ag.	1.580	1.580	1.580	1.580
Wage-Employment				
Agriculture	1.247		2.723	
Non-Ag.	1.552	1.244	3.371	1.238
All Employment				
Agriculture	1.003		1.021	
Non-Ag.	1.570	1.565	2.220	2.175

Notes: Panel A reports the earnings by rural-urban status. Gaps are calculated relative to the value for rural by employment type. Panel B reports the gaps between agricultural and non-agricultural employment. Earnings and output per worker in rural self-employment are normalized to one.

Table 12: Quantifying Labor Misallocation

	Baseline		10% Reduction in Search Costs		10% Reduction in Migration Costs	
	Labor Market Power	Competitive Equilibrium	Labor Market Power	Competitive Equilibrium	Labor Market Power	Competitive Equilibrium
	(1)	(2)	(3)	(4)	(5)	(6)
Total Output	1.000	1.048	1.217	1.249	0.958	1.005
Output per Worker						
Rural	0.894	0.892	0.909	0.911	0.894	0.890
Urban	1.643	1.601	1.482	1.448	1.916	1.854
Welfare	1.000	1.013	1.831	2.044	2.231	2.308
Average Wage	0.978	1.088	0.981	1.159	0.945	1.050
Urban-Rural Gap	1.414	1.532	1.346	1.438	1.646	1.775
Self-Emp Income Gap	1.579	1.992	0.787	0.990	1.677	2.106
Markdown						
Rural	0.705	1.000	0.701	1.000	0.704	1.000
Urban	0.714	1.000	0.714	1.000	0.714	1.000
Employment Share						
Firm	0.139	0.183	0.484	0.568	0.125	0.167
Urban	0.142	0.152	0.158	0.165	0.104	0.114

Notes: Odd numbered columns report the results under spatial labor market power. Even numbered columns report the results when moving to competitive equilibrium. Columns (3) and (4) reduce job search costs by 10%, *i.e.* $\delta_d = 0.9 * \delta_d^{data} + 0.1 * 1$ for all d. Columns (5) and (6) reduce migration costs by 10%, *i.e.* $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ for all o, d. Welfare and total output are normalized to one at baseline.

Table 13: Asymmetric Reductions in Migration Costs

	Baseline	Symmetric	To Urban Only	From Urban Only	Rural to Rural
	(1)	(2)	(3)	(4)	(5)
Total Output	1.000	0.958	1.038	0.986	0.950
Output per Worker					
Rural	0.894	0.894	0.902	0.897	0.897
Urban	1.643	1.916	1.284	1.772	2.040
Welfare	1.000	2.231	1.118	1.149	2.000
Average Wage	0.978	0.945	0.991	0.970	0.940
Urban-Rural Gap	1.414	1.646	1.096	1.520	1.743
Self-Emp Income Gap	1.579	1.677	1.412	1.621	1.697
Markdown					
Rural	0.705	0.704	0.705	0.704	0.704
Urban	0.714	0.714	0.714	0.714	0.714
Employment Share					
Firm	0.139	0.125	0.174	0.131	0.122
Urban	0.142	0.104	0.257	0.118	0.090

Notes: Columns (1) and (2) restates the results in columns (1) and (5) in Table 12. Column (3) reduces migration costs by 10% only when the destination is an urban district, *i.e.* $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ if d urban. Column (4) reduces migration costs by 10% only when the origin is an urban district, *i.e.* $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ if o urban. Column (5) reduces migration costs by 10% only when both the origin and destination are rural districts, *i.e.* $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ if o and d are rural.

Table 14: Simulated Reduction in Migration Costs with Endogenous Amenities and Productivity

	Baseline	Congestion Only	Agglomeration Only	Congestion & Agglomeration	High Congestion
	(1)	(2)	(3)	(4)	(5)
Total Output	0.958	0.960	0.954	0.956	0.974
Output per Worker					
Rural	0.894	0.894	0.897	0.897	0.893
Urban	1.916	1.904	1.905	1.892	1.811
Welfare	2.231	2.226	2.238	2.232	2.184
Average Wage	0.945	0.947	0.943	0.944	0.959
Urban-Rural Gap	1.646	1.635	1.629	1.619	1.557
Self-Emp Income Gap	1.677	1.673	1.672	1.668	1.643
Markdown					
Rural	0.704	0.704	0.704	0.704	0.705
Urban	0.714	0.714	0.714	0.714	0.714
Employment Share					
Firm	0.125	0.126	0.124	0.125	0.130
Urban	0.104	0.105	0.102	0.104	0.116
λ	0.000	-0.040	0.000	-0.040	-0.500
ϕ	0.000	0.000	0.050	0.050	0.000

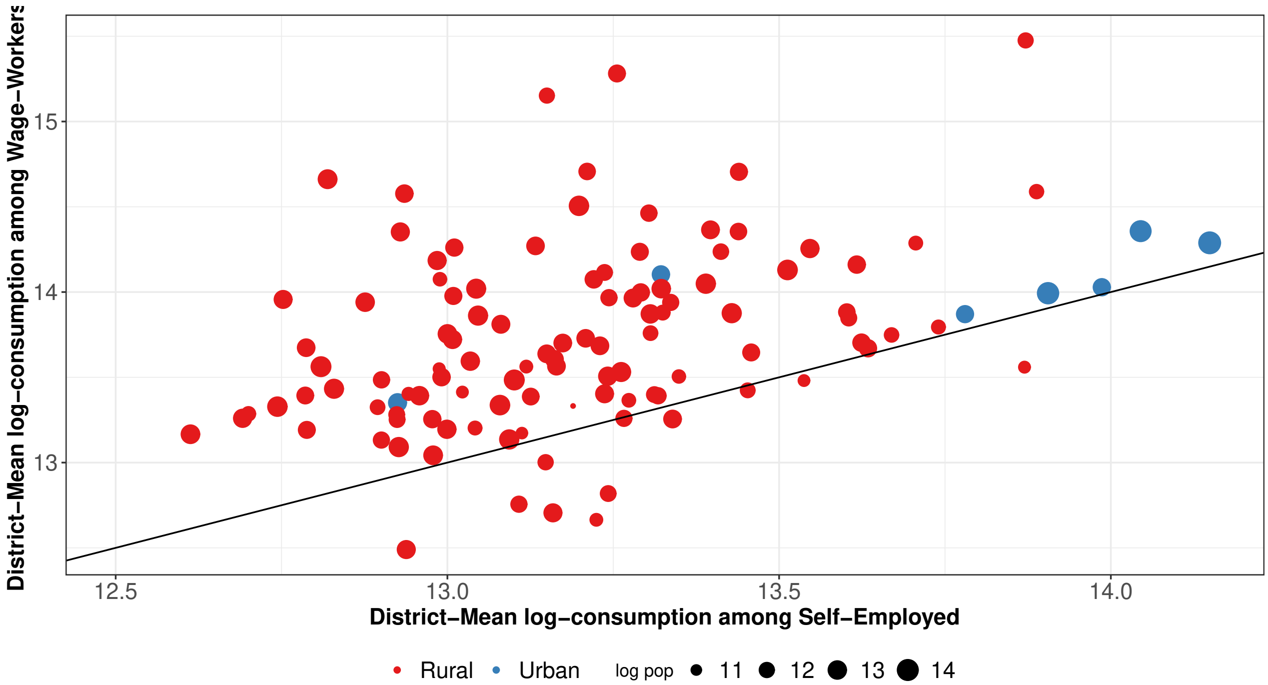
Notes: The table reports the simulated counterfactual equilibrium for a 50% decline in migration costs under various assumptions on the engoeneity of amenities and productivity. All counterfactuals use the spatial oligopoly firm competition. Output and welfare can be interpreted relative to a baseline value of one. Column 1 restates the baseline counterfactual reported in Table 21. Column 2 endogenizes amenities only; column 3 endogenizes productivity only; column 4 endogenizes both productivity and amenities. Column 5 has very high congestion forces and no agglomeration forces.

Table 15: Optimal City Size

			Self-Emp	Firms		Lower Bound		Upper Bound	
City	Pop.	Empl.	Ratio	MS	N^*	MS	N^*	MS	N^*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dodoma	394	160	0.176	0.013	3517	0.006	3543	1.252	980
Arusha	291	114	0.184	0.213	2862	0.130	3116	0.847	1487
Dar es Salaam	4288	1731	0.202	0.164	3009	0.103	3204	0.751	1642
Mwanza	612	196	0.176	0.145	3068	0.071	3313	1.197	1036

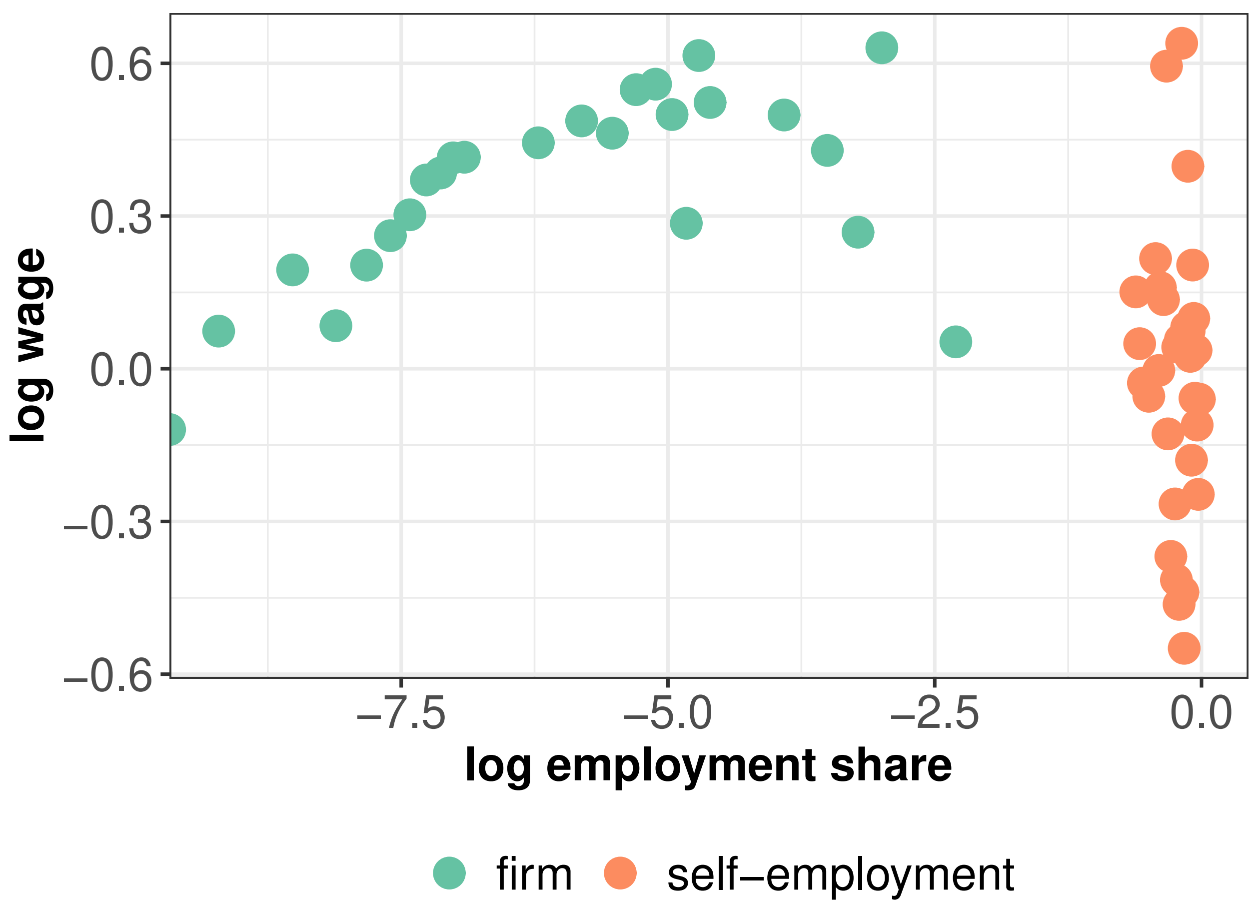
Notes: Population, employment ant N^* are in thousands. The self-employment ratio is the ratio of manufacturing to services in the 2012 census. Population and employment are totals from the 2012 census. MS is the manufacturing to services value-added ratio. N^* is the optimal city employment for that ratio using the values from Au and Henderson (2006).

Figure 1: Average Consumption by Main Occupation



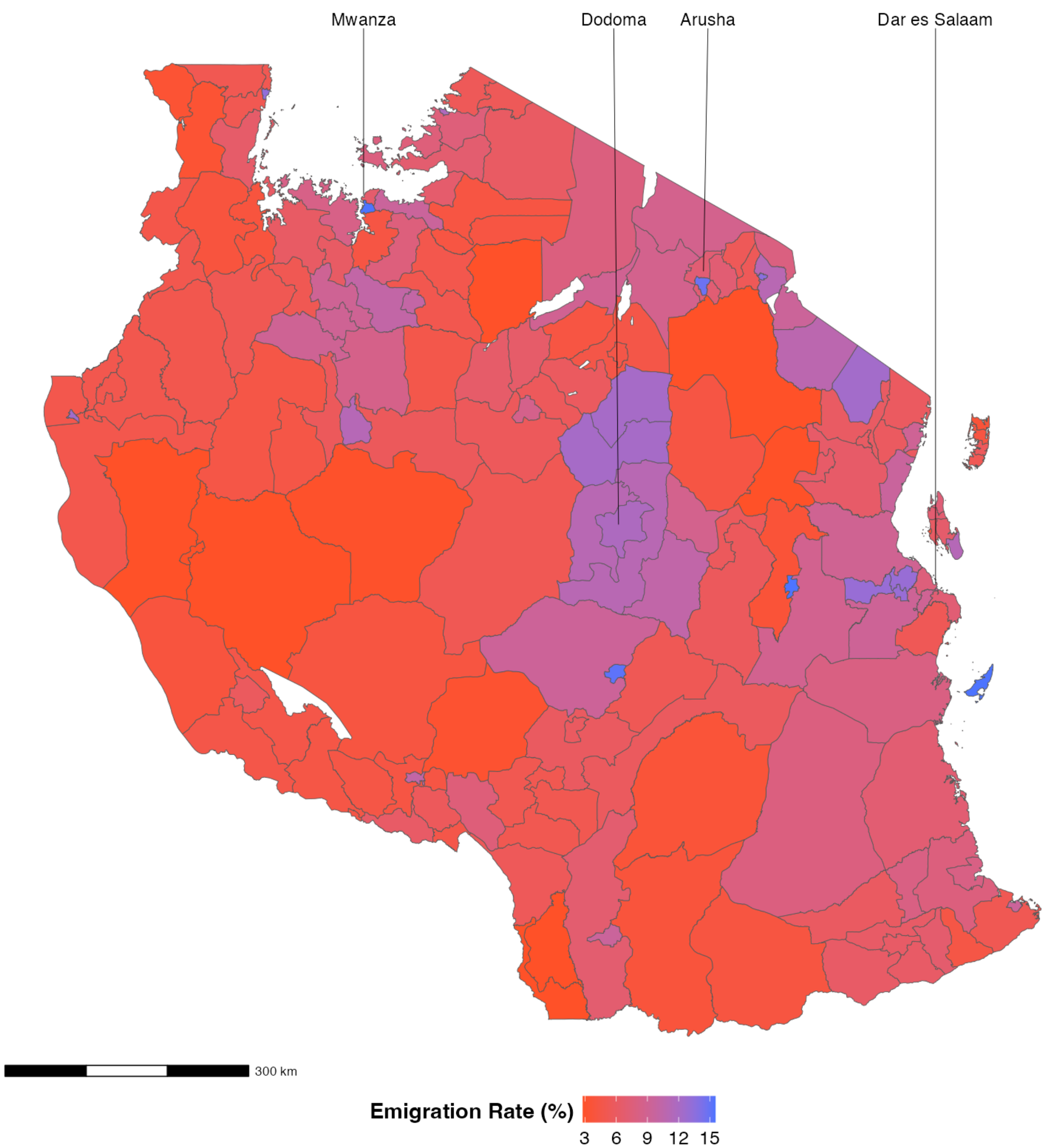
Notes: The figure plots the average log-consumption-per-adult-equivalent by district in the 2010 LSMS. The sample is limited to prime-aged individuals whose main occupation in the last twelve months was either self-employment or work for a wage.

Figure 2: Observed Labor Supply Curve by Employment Type



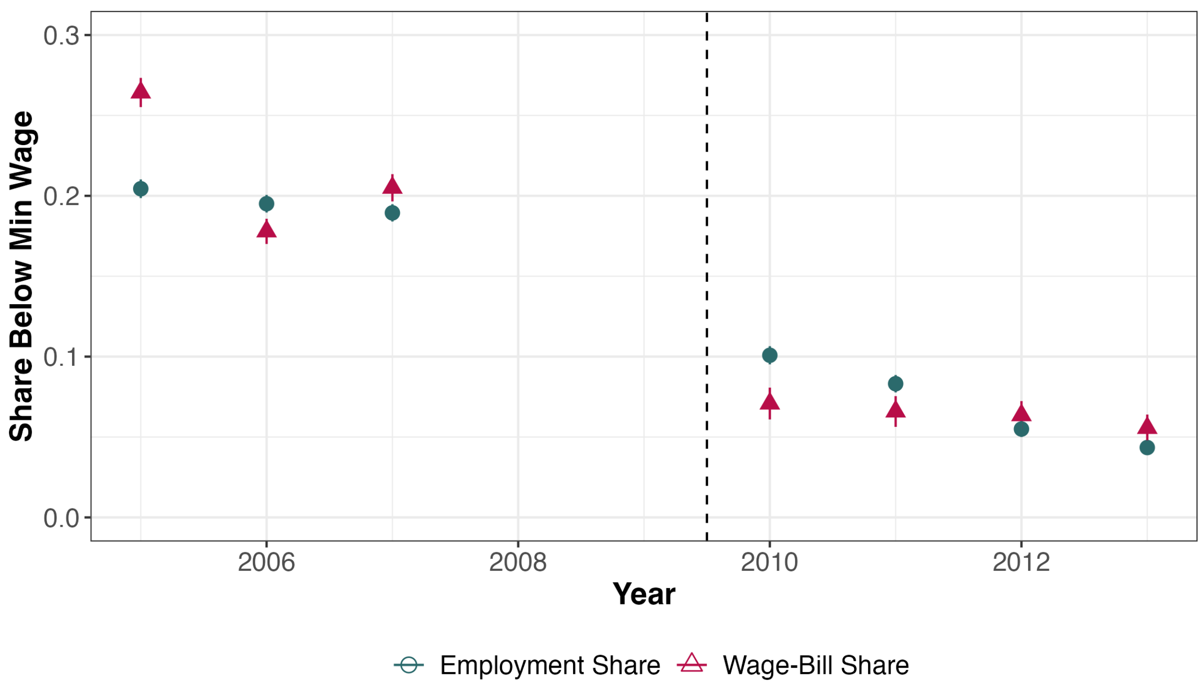
Notes: The figure plots the average wage against the share of market employment by employment type. Each firm point represents the average wage among all firms whose market share is within $1x10^{-}4$.

Figure 3: Prime-Aged Emigration



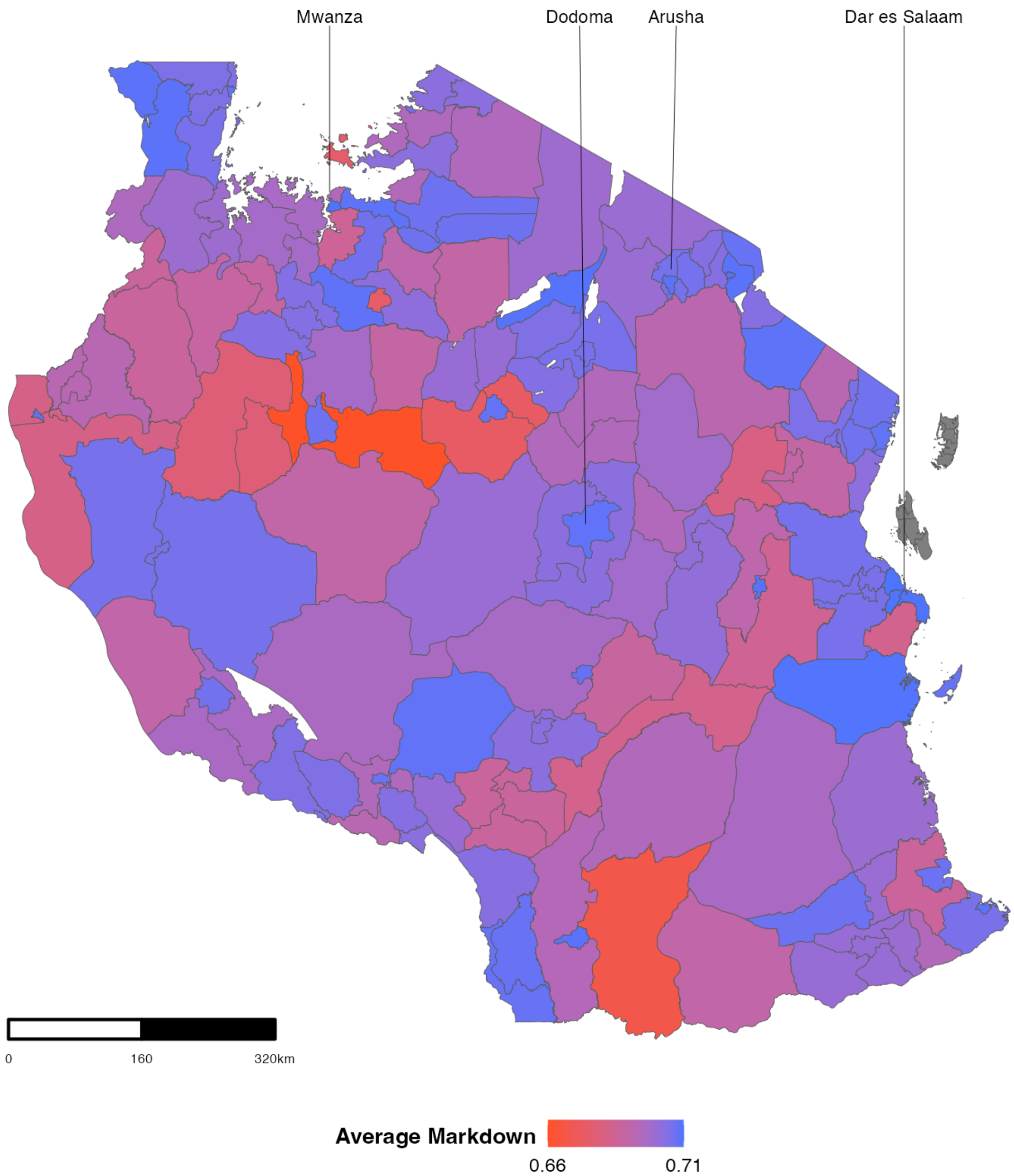
Notes: Displaying one-year emigration rate among individuals aged 15-65 in the 2012 census. Emigration Rates are winsorized at the 2nd and 98th percentiles.

Figure 4: Non-Compliance Rate Event Study



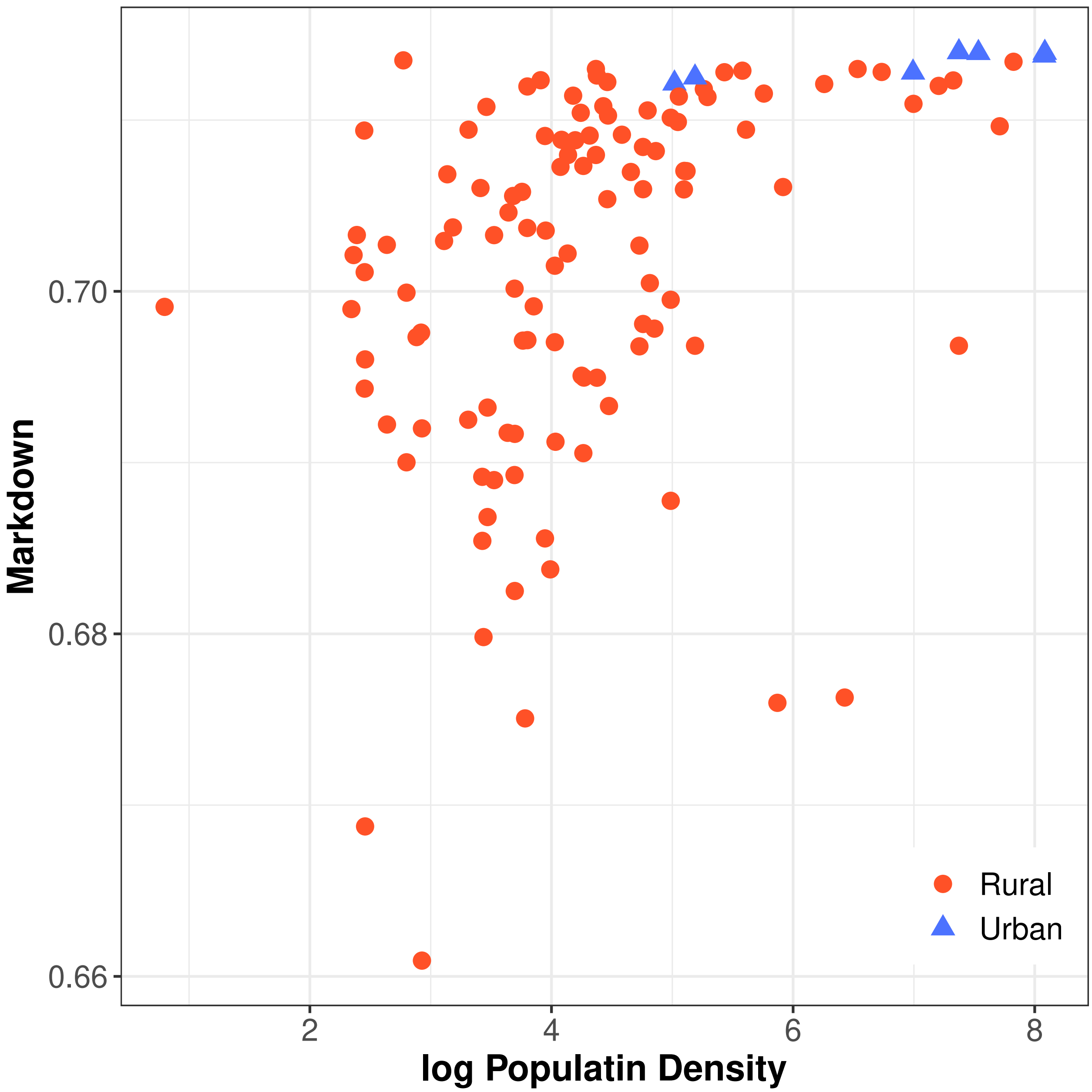
Notes: Plotting the coefficient estimates and 95% confidence interval from equation (9) for the employment non-compliance rate (ENC) and GAP measure. The dashed line indicates the date when the minimum wage law was implemented.

Figure 5: Spatial Distribution of Markdowns



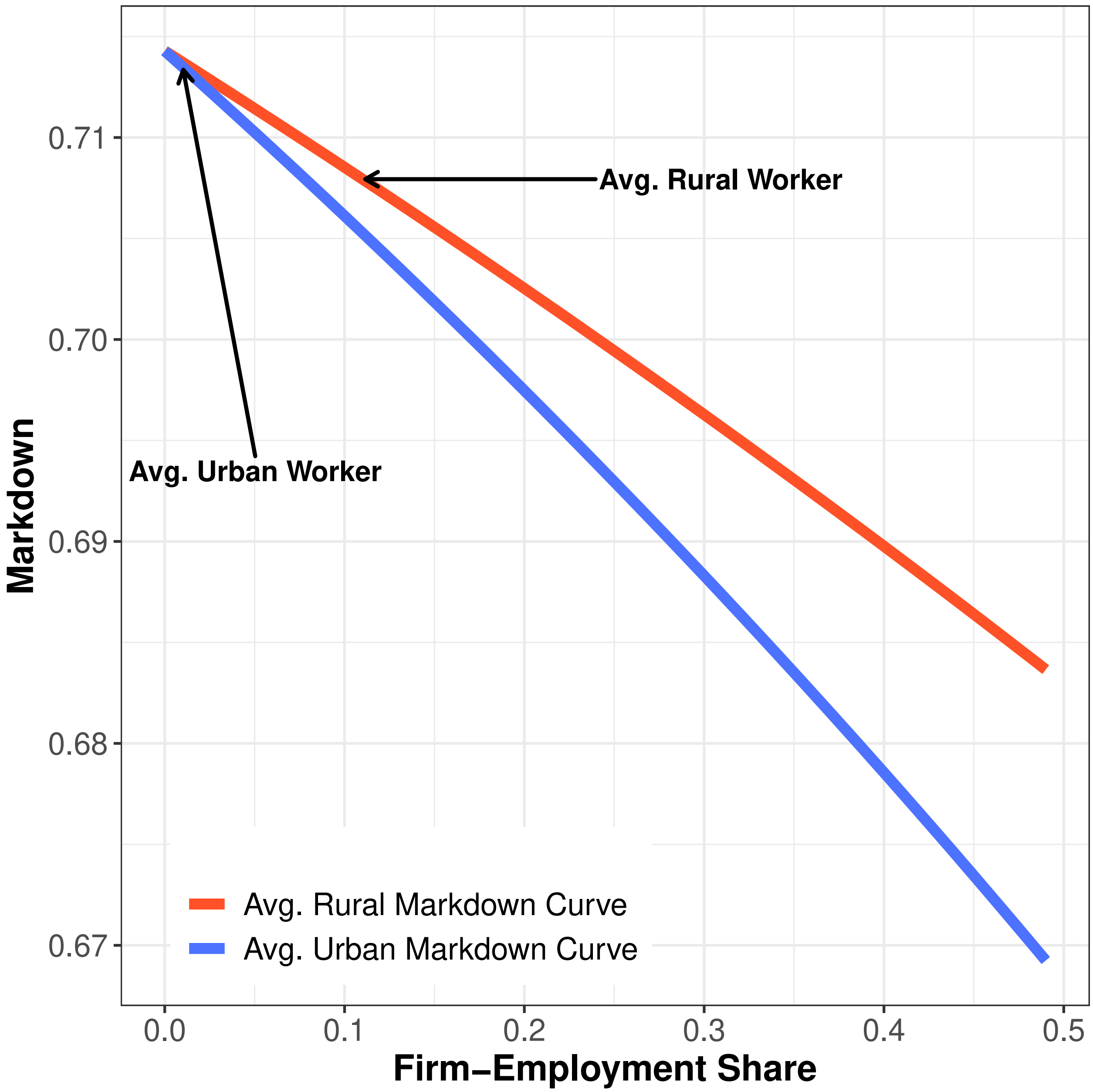
Notes: The figure plots the estimated average markdown among wage-workers in each district.

Figure 6: Wage Markdowns by Population Density



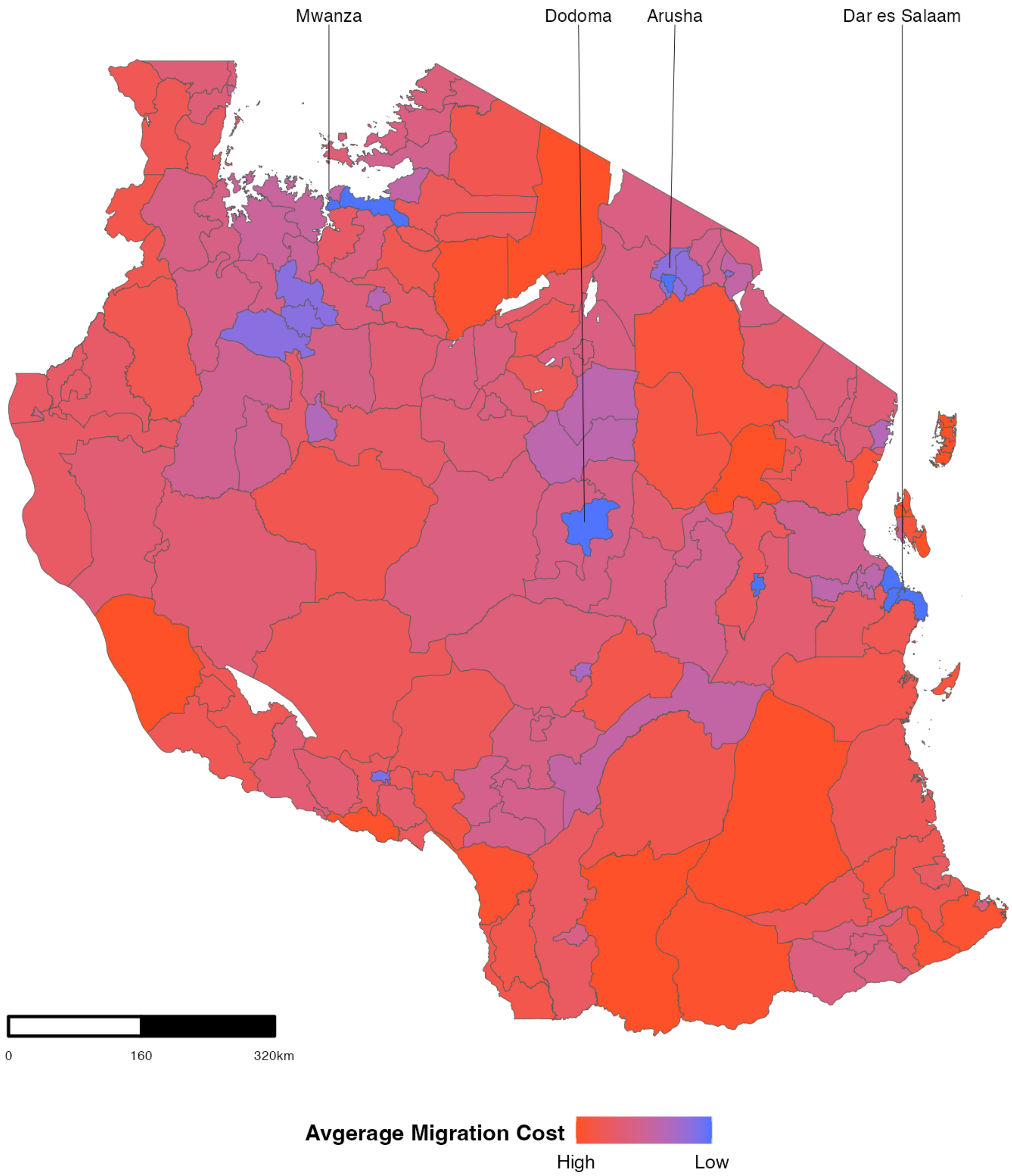
Notes: The figure plots the estimated average wage markdown among firm employees by district against the reported population density in the Census.

Figure 7: Equilibrium Markdown Curve in Urban and Rural Districts



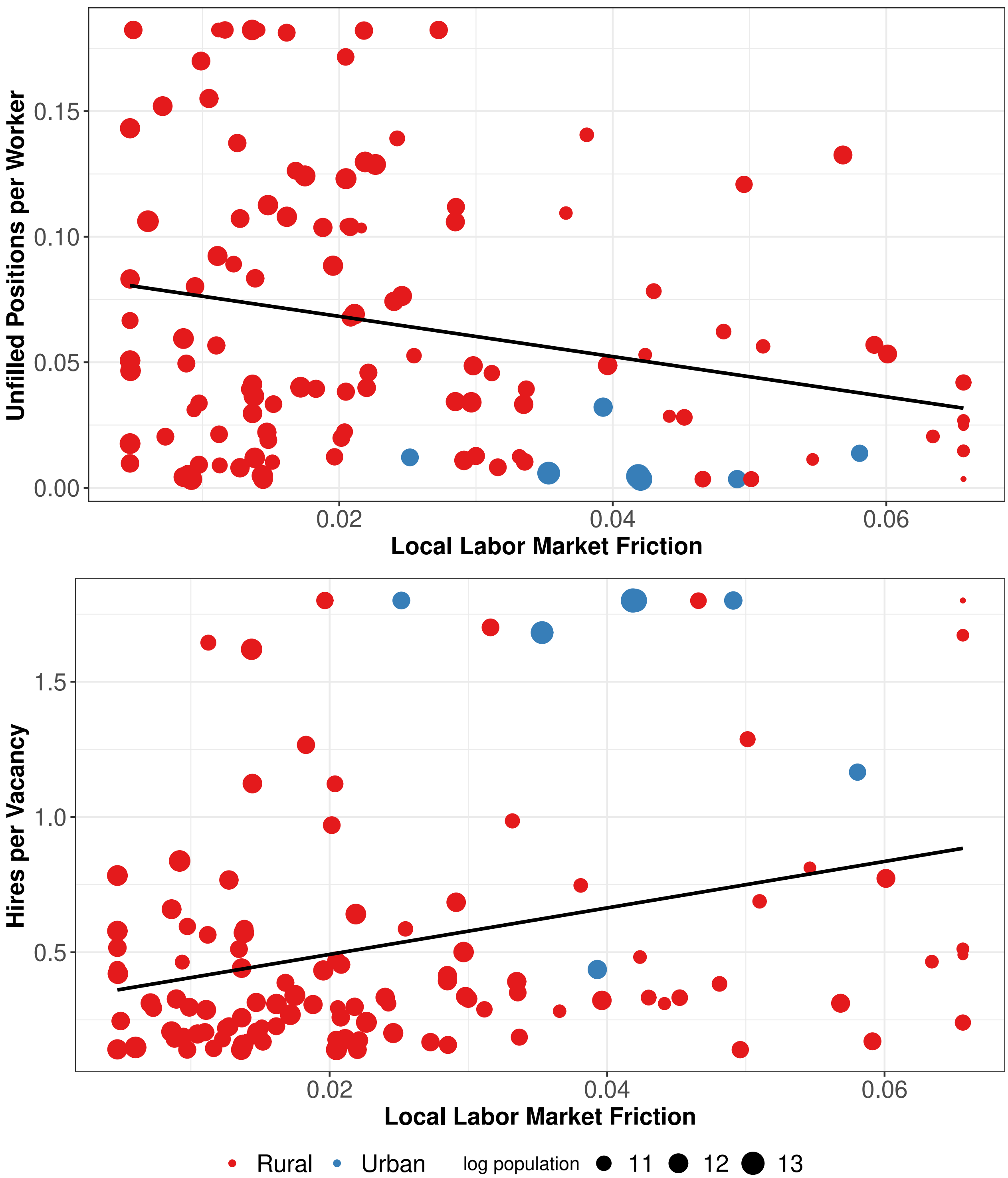
Notes: The figure plots the equilibrium wage markdown (μ) as a function of firm employment share (s_f). The markdown curves are calibrated with the average values of s_d and s_m in urban and rural markets. The two highlighted points represent the average values for s_f in urban and rural markets.

Figure 8: Estimated Average District Migration Cost



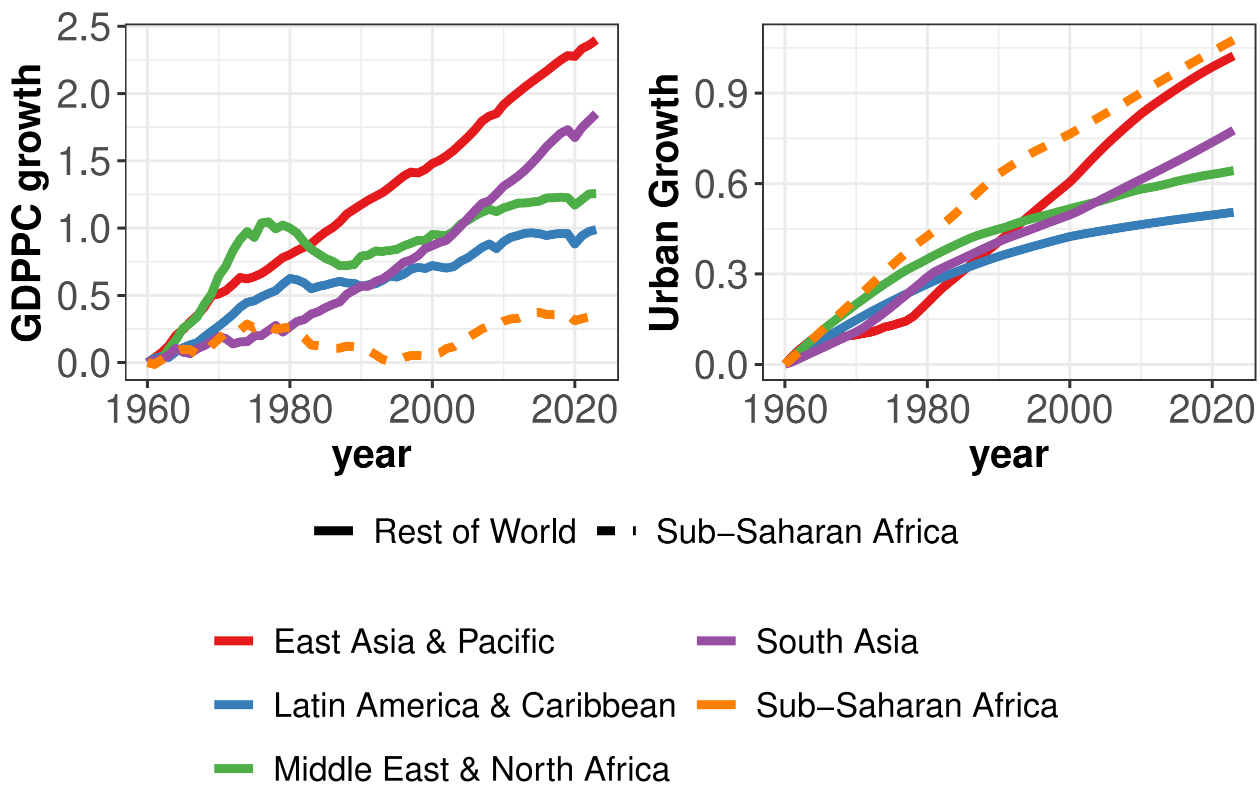
Notes: Displaying the estimated migration cost in each district using one-year migration flows among individuals aged 15-65 in the 2012 census. Migration rates are winsorized at the 5th and 95th percentiles to show variation.

Figure 9: The Correlation between Search Costs and Hires and Vacancies



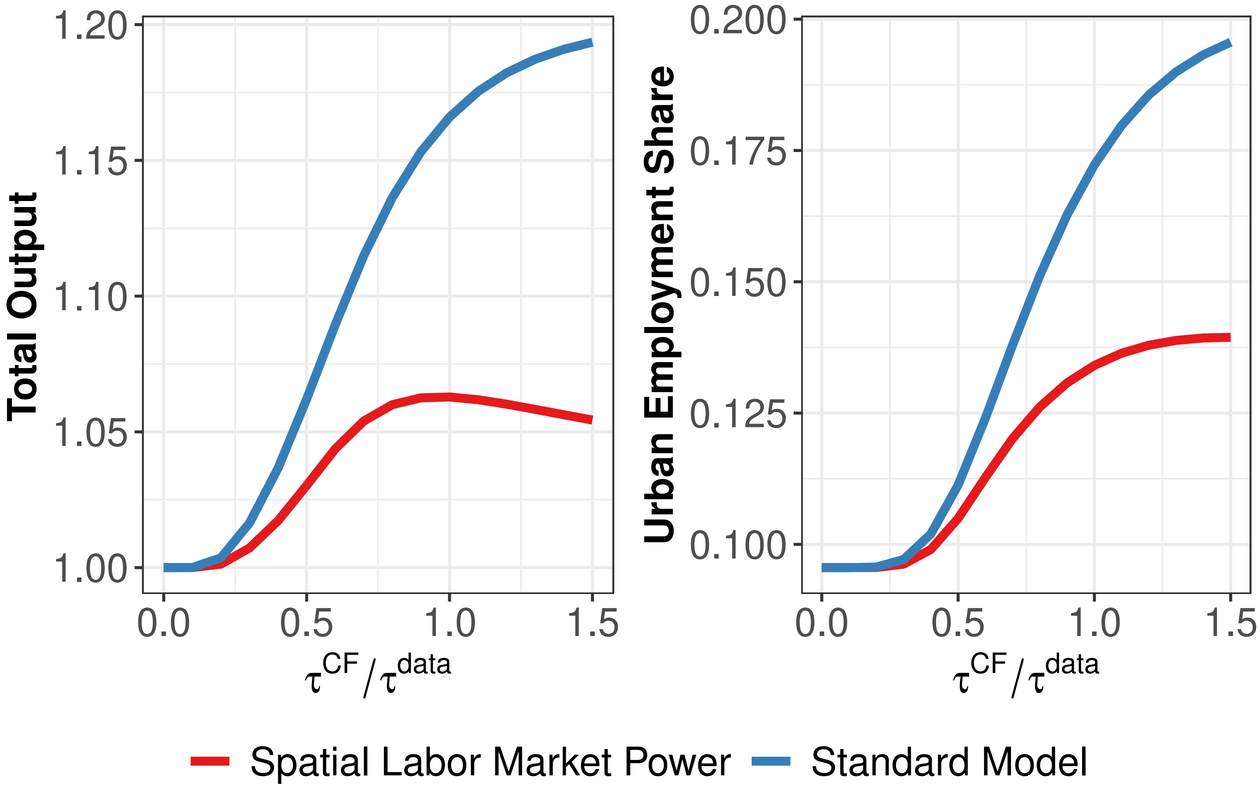
Notes: Displaying the average number of vacancies per worker by district (top) and the number of hires per posted vacancy (bottom) against the estimated search cost δ. Vacancy and hires are averaged over the period 2010-2017 and are winsorized at the 5th and 95th percentile.

Figure 10: Urbanization and Growth Across Regions



Notes: The figure plots the log change in real GDP per capita relative to the base year (left) and the log change in the urban population share (right). North-America and Europe are not pictured. *Source:* World Bank Development Indicators.

Figure 11: Urbanization and Growth in Two Models



Notes: The figure plots total output per worker relative to the baseline case with infinite migration costs (left) and the share of the population living in Dar es Salaam (right). The initial labor distribution corresponds to the labor shares in each district in 2000. The competitive markets model has firms pay workers their marginal product and has location amenities normalized to one across locations.

A Additional Figures and Tables

Table 16: First Stage Estimation Results for the Between-Firm Elasticity

	log wage			log wage $\times s_f$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{GAP}$	-0.215***		0.201***	0.002		-0.004
	(0.046)		(0.049)	(0.001)		(0.002)
$s_f \times \widehat{GAP}$	-0.698		-1.530	-0.427*		-0.133
	(1.758)		(2.209)	(0.252)		(0.461)
$s_f\, s_d \times \widehat{GAP}$	0.566		32.843	-3.659*		6.513*
	(18.520)		(28.162)	(2.033)		(3.847)
s_f	0.641*	0.825	0.715	12.490***	12.869***	12.887***
	(0.324)	(0.519)	(0.529)	(0.100)	(0.147)	(0.151)
$s_f\, s_d$	-0.457	-0.439*	-0.409	0.030**	0.011	0.008
	(0.276)	(0.258)	(0.273)	(0.012)	(0.011)	(0.011)
$s_f \times \widehat{ENC}$		-0.495	1.037		-1.289***	-1.226*
		(1.722)	(2.329)		(0.470)	(0.654)
$\widehat{ENC}$		-0.711***	-1.009***		0.003*	0.007*
		(0.090)	(0.116)		(0.002)	(0.004)
$s_f\, s_d \times \widehat{ENC}$		-7.538	-26.486*		-3.028*	-7.077*
		(7.406)	(15.918)		(1.803)	(3.891)
Firms	3880	3880	3880	3880	3880	3880

Notes: The table presents the first stage estimation results for the between-firm elasticity η. s_f is the firm's share of firm-employment in the local market,s_d is the share of employment in the local market that is engaged in wage-work, $\widehat{GAP}$ is the predicted exposure to the minimum wage, and $\widehat{ENC}$ is the predicted share of employees paid below the minimum wage, log employment is the log of total employment in the district, inclusive of self-employment. Columns (1) and (2) predict the GAP and ENC instruments using observations from 2005-2007 in the EES. Columns (3) and (4) restrict the prediction period to 2007. Robust standard errors clustered by district in parenthesis. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 16 Continued: First Stage Estimation Results for the Between-Firm Elasticity

	Dependent Variable: log wage $\times s_f s_d$		
	(1)	(2)	(3)
$\widehat{GAP}$	-0.007***		0.003*
	(0.001)		(0.002)
$s_f \times \widehat{GAP}$	-2.489***		0.077
	(0.583)		(0.507)
$s_f\, s_d \times \widehat{GAP}$	31.622***		0.627
	(4.489)		(6.752)
s_f	0.976***	1.063***	1.067***
	(0.116)	(0.132)	(0.136)
$s_f\, s_d$	-0.005	-0.005	-0.004
	(0.011)	(0.007)	(0.007)
$s_f \times \widehat{ENC}$		-2.341***	-2.385***
		(0.539)	(0.803)
$\widehat{ENC}$		-0.010***	-0.015***
		(0.002)	(0.003)
$s_f\, s_d \times \widehat{ENC}$		25.680***	25.309***
		(5.051)	(8.948)
Firms	3880	3880	3880

Notes: See Table 16 footnote for details.

Table 17: Berger et al. (2022) Between-Firm Elasticity Estimation Method

	(1) log wage	(2) log empl	(3) log wage	(4) log empl
$\widehat{ENC}$	-0.711***	-1.944***		
	(0.090)	(0.281)		
$s_f \times \widehat{ENC}$	-0.495	-2.286		
	(1.722)	(12.896)		
$s_f\, s_d \times \widehat{ENC}$	-7.538	336.738***		
	(7.406)	(93.952)		
s_f	0.825	0.882	0.782**	5.440***
	(0.519)	(3.418)	(0.362)	(2.003)
$s_f\, s_d$	-0.439*	2.717***	-0.456	2.525***
	(0.258)	(0.866)	(0.280)	(0.943)
$\widehat{GAP}$			-0.207***	-1.008***
			(0.049)	(0.069)
$s_f \times \widehat{GAP}$			-1.227	-25.282**
			(1.723)	(9.889)
$s_f\, s_d \times \widehat{GAP}$			1.450	572.645***
			(18.115)	(77.508)
Firms	3880	3880	3880	3880
Region FE			Y	Y
$\hat{\eta}$		2.736		4.859
std. err.		(0.379)		(1.011)
Instruments				

Notes: Robust standard errors clustered by district in parenthesis. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 18: Sensitivity of the Migration Elasticity θ to Alternative Estimating Strategies

	Dependent Variables: n_{od}					
	Census Migration			Survey Migration		
	(1)	(2)	(3)	(4)	(5)	(6)
W_d/W_o	1.398***	1.399***	0.857***	2.200***	2.014***	1.287***
	(0.116)	(0.116)	(0.207)	(0.214)	(0.205)	(0.332)
$\log \tau_{od}$	0.422***	0.421***		0.492***	0.413***	
	(0.015)	(0.016)		(0.024)	(0.026)	
log distance			-0.835***			-0.620***
			(0.088)			(0.145)
log migrant stock			0.195***			0.443***
			(0.029)			(0.071)
log employment ratio			1.049***			0.711***
			(0.112)			(0.179)
F-statistic	1345.804	1346.053	604.535	1345.804	1378.737	604.535
Region Pairs	552	546	552	552	335	552
Origin FE	Y	Y	Y	Y	Y	Y
η	2.5	2.5	2.5	2.5	2.5	2.5
γ	1.5	1.5	1.5	1.5	1.5	1.5
Restriction	None	Migrating Pairs	None	None	Migrating Pairs	None

Notes: The table presents IV-Poisson estimation results for the migration elasticity θ at the district level. See Table 9 for variable definitions. Reporting the results using spatially deflated wages in each district only. Robust standard errors in parenthesis. Reporting the Kleibergen and Paap (2006) weak identification F-statistic. $^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Table 19: District Level estimation of the Migration Elasticity θ

	Dependent Variable: n_{od}					
	(1)	(2)	(3)	(4)	(5)	(6)
	n_{od}	n_{od}	n_{od}	n_{od}	n_{od}	n_{od}
n_{od}						
$W_d(\underline{w})/W_o(\underline{w})$	0.211***			0.106***		
	(0.013)			(0.016)		
W_d/W_o		0.379***	0.692***		0.131***	0.301***
		(0.032)	(0.051)		(0.034)	(0.049)
$\log \tau_{od}$	0.483***	0.493***	0.486***			
	(0.005)	(0.006)	(0.005)			
log distance				-0.769***	-0.753***	-0.764***
				(0.024)	(0.024)	(0.024)
log migrant stock				0.187***	0.196***	0.190***
				(0.006)	(0.006)	(0.006)
log employment ratio				0.716***	0.774***	0.737***
				(0.028)	(0.027)	(0.028)
F-statistic			1.2e+04			1.1e+04
District Pairs	14042	14042	14042	14042	14042	14042
Origin FE	Y	Y	Y	Y	Y	Y
η	2.5	2.5	2.5	2.5	2.5	2.5
γ	1.4	1.4	1.4	1.4	1.4	1.4
Estimation	Poisson	Poisson	IV-Poisson	Poisson	Poisson	IV-Poisson

Notes: The table presents Poisson and IV-Poisson estimation results for the migration elasticity θ at the district level. See Table 9 for variable definitions. Reporting the results using spatially deflated wages in each district only. Robust standard errors in parenthesis. Reporting the Kleibergen and Paap (2006) weak identification F-statistic. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 20: Simulated Reduction in Migration Costs without Amenities

	No Location Amenities		No Job Amenities		No Location or Job Amenities	
	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs
	(1)	(2)	(3)	(4)	(5)	(6)
Total Output	1.000	1.010	1.000	0.951	1.000	1.005
Output per Worker						
Rural	0.916	0.921	0.853	0.857	0.878	0.886
Urban	1.442	1.508	1.831	2.141	1.604	1.680
Welfare	1.000	2.213	1.000	2.185	1.000	2.164
Average Wage	1.054	1.076	1.008	0.972	1.087	1.106
Urban-Rural Gap	1.212	1.258	1.546	1.793	1.322	1.368
Self-Emp Income Gap	1.349	1.326	1.845	2.000	1.540	1.510
Markdown						
Rural	0.705	0.705	0.713	0.713	0.712	0.712
Urban	0.714	0.714	0.714	0.714	0.714	0.714
Employment Share						
Firm	0.153	0.146	0.137	0.119	0.157	0.149
Urban	0.160	0.135	0.150	0.111	0.168	0.143

Notes: Columns (1) and (2) equate location amenities across space, *i.e.* $B_d = 1 \,\forall\, d$. Columns (3) and (4) equate job amenities, *i.e.* $b_f = 1 \,\forall\, f$. Columns (5) and (6) equate both. Even numbered columns report the simulated values turning off each type of amenity. Odd numbered columns report the counterfactual when migration costs are reduced by 10%, $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ for all o, d.

Table 21: Simulated Reduction in Migration Costs by Assumed Labor Market Competition

	Spatial Oligopoly		Monopsonistic Competition		Competitive Equilibrium	
	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs
	(1)	(2)	(3)	(4)	(5)	(6)
Total Output	1.000	0.958	1.000	0.957	1.000	0.962
Output per Worker						
Rural	0.894	0.894	0.893	0.893	0.918	0.914
Urban	1.643	1.916	1.647	1.922	1.497	1.738
Welfare	1.000	2.231	1.000	2.235	1.000	2.365
Average Wage	0.978	0.945	0.978	0.945	0.978	0.945
Urban-Rural Gap	1.414	1.646	1.414	1.646	1.414	1.646
Self-Emp Income Gap	1.579	1.677	1.579	1.676	1.579	1.676
Markdown						
Rural	0.705	0.704	0.714	0.714	1.000	1.000
Urban	0.714	0.714	0.714	0.714	1.000	1.000
Employment Share						
Firm	0.139	0.125	0.139	0.125	0.139	0.125
Urban	0.142	0.104	0.142	0.104	0.142	0.104

Notes: Even numbered columns report the simulated values under each type of competition. Odd numbered columns report the counterfactual when migration costs are reduced by 10%, $\tau_{od} = 0.9 * \tau_{od}^{data} + 0.1 * 1$ for all o, d. Columns (1) and (2) report the results under spatial oligopoly. Columns (3) and (4) report the results under monopsonistic competition $(\mu = \eta \, \forall \, f)$. Columns (5) and (6) report the results under perfect competition $(\mu = 1 \, \forall \, f)$. The labor supply curve is unchanged under all specifications, but the location and firm amenity values, and productivities are re-estimated to match the labor share under each type of competition. Welfare and total output are normalized to one in the baseline under each type of competition.

Table 22: Simulated Reduction in Migration Costs Ignoring the Self-Employment Margin

	Spatial Oligopoly		Monopsonistic Competition		Competitive Equilibrium	
	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs	Baseline	Reduced Migration Costs
	(1)	(2)	(3)	(4)	(5)	(6)
Total Output	1.000	0.986	1.000	0.986	1.000	0.986
Output per Worker						
Rural	0.996	0.972	0.988	0.966	0.988	0.966
Urban	1.006	1.042	1.017	1.052	1.017	1.052
Welfare	1.000	2.021	1.000	2.020	1.000	2.081
Average Wage	1.019	1.006	1.019	1.004	1.019	1.004
Urban-Rural Gap	1.030	1.087	1.030	1.090	1.030	1.090
Markdown						
Rural	0.701	0.704	0.714	0.714	1.000	1.000
Urban	0.714	0.714	0.714	0.714	1.000	1.000
Employment Share						
Urban	0.415	0.397	0.415	0.398	0.415	0.398

Notes: Even numbered columns report the simulated values under each type of competition. Odd numbered columns report the counterfactual when migration costs are reduced by 50%, $\tau_{od} = 0.5 * \tau_{od}^{data} + 0.5 * 1$ for all $o,\, d$. Columns (1) and (2) report the results under spatial oligopoly. Columns (3) and (4) report the results under monopsonistic competition $(\mu = \eta \,\forall\, f)$. Columns (5) and (6) report the results under perfect competition $(\mu = 1 \,\forall\, f)$. The labor supply curve is unchanged under all specifications, but the location and firm amenity values, and productivities are re-estimated to match the labor share under each type of competition. Welfare and total output are normalized to one in the baseline under each type of competition.

B Mathematical Appendix

Labor Supply Under the distributional assumptions on the amenities, the probability that a worker from o chooses firm f or self-employment s in market d, can be expressed as

$$\begin{aligned}\Pr(w_{f(d)}|\omega, o) =& \frac{(b_f w_f)^\eta}{\sum_{f\in d}(b_f w_f)^\eta} \times \frac{\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}}}{\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma} \\ & \times \frac{(B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}}{\sum_d \left((B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}\right)}\end{aligned}$$

$$\Pr(w_{ad}|\omega, o) = \frac{w_{ad}^\gamma}{\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma} \times \frac{(B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}}{\sum_d \left((B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}\right)}$$

Total labor supply is then found by aggregating

$$n_{fdo} = \int_{\omega\in o} \Pr(w_{f(d)}|\omega, o)\, \partial F(L)$$

$$n_{fdo} = \frac{(b_f w_f)^\eta}{\sum_{f\in d}(b_f w_f)^\eta} \times \frac{\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}}}{\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma} \times \frac{(B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}}{\sum_d \left((B_d\tau_{od})^\theta \left(\left(\delta_d^\eta \sum_{f\in d}(b_f w_f)^\eta\right)^{\frac{\gamma}{\eta}} + w_{ad}^\gamma\right)^{\frac{\theta}{\gamma}}\right)} L_o$$

$$n_{fdo} = \left(\frac{b_f w_f}{W_{fd}}\right)^\eta \left(\frac{\delta_d W_{fd}}{W_d}\right)^\gamma \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^\theta L_o$$

Where the aggregates W_{df}, W_d and $\mathbf{W_o}$ are those defined in the main text. Following the same logic, it can be shown that the labor supply to self-employment in d from o is given by:

$$n_{ado} = \left(\frac{w_a}{W_d}\right)^\gamma \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^\theta L_o$$

Firm's problem Firms choose wages taking the labor supply curve as given to maximize profits

$$\Pi = \max_{w_f} \quad A_f n_f^\alpha - w_f n_f \quad ; \quad n_f = \sum_o n_{fdo} \quad ; \quad n_{fdo}(w) = \left(\frac{b_f w_f}{W_{fd}}\right)^\eta \left(\frac{\delta_d W_{fd}}{W_d}\right)^\gamma \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^\theta L_o$$

The FOC is

$$\sum_o \left(\alpha A_f n_f^{\alpha-1}\frac{\partial n_{fdo}}{\partial w_f} - n_{fdo} - w_f\frac{\partial n_{fdo}}{\partial w_f}\right) = 0$$

For ease of notation, define the market share quantities

$$\begin{aligned} s_f = \left(\frac{b_f w_f}{W_d}\right)^{\eta} \quad &; \quad s_d = \left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma} \\ s_{od} = \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta} \quad &; \quad s_m = \sum_o \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta} L_o = \sum_o s_{od} L_o \end{aligned}$$

It is also useful to define the following partials

$$\begin{aligned} \frac{\partial W_{fd}}{\partial w_f} =& \frac{b_f^{\eta} w_f^{\eta-1}}{W_{fd}^{\eta-1}} = \frac{W_{fd}}{w_f} s_f \\ \frac{\partial W_d}{\partial W_{fd}} =& \frac{\delta_d^{\gamma} W_{fd}^{\gamma-1}}{W_d^{\gamma-1}} = \frac{W_d}{W_{fd}} s_d \\ \frac{\partial \mathbf{W_o}}{\partial W_d} =& \frac{(\tau_{od} B_d)^{\theta} W_d^{\theta-1}}{\mathbf{W_o}^{\theta-1}} = \frac{\mathbf{W_o}}{W_d} s_{od} \end{aligned}$$

The partial is

$$\begin{aligned} \frac{\partial n_{fdo}}{\partial w_f} =& \underbrace{\eta\left(\frac{b_f w_f}{W_{fd}}\right)^{\eta-1}\left(\frac{W_{fd} b_f - b_f w_f \frac{\partial W_{fd}}{\partial w_f}}{W_{fd}^2}\right) s_d s_{od} L_o}_{\text{firm partial}} \\ &+ \underbrace{s_f \gamma\left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma-1}\left(\frac{W_d \delta_d \frac{\partial W_{fd}}{\partial w_f} - \delta_d W_{fd}\frac{\partial W_d}{\partial W_{fd}}\frac{\partial W_{fd}}{\partial w_f}}{W_d^2}\right) s_{od} L_o}_{\text{market partial}} \\ &+ \underbrace{s_f s_d \theta\left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}}\right)^{\theta-1}\left(\frac{\mathbf{W_o}\tau_{od} B_d \frac{\partial W_d}{\partial W_{fd}}\frac{\partial W_{fd}}{\partial w_f} - \tau_{od} B_d W_d \frac{\partial \mathbf{W_o}}{\partial W_d}\frac{\partial W_d}{\partial W_{fd}}\frac{\partial W_{fd}}{\partial w_f}}{\mathbf{W_o}^2}\right) L_o}_{\text{aggregate partial}} \end{aligned}$$

Appealing to the wage index partials defined above, the firm partial can be reduced as

$$\begin{aligned}
&\eta\left(\frac{b_f w_f}{W_{fd}}\right)^{\eta-1}\left(\frac{b_f}{W_{fd}}-\frac{b_f w_f}{W_{fd}^2}\frac{W_{fd}}{w_f}s_f\right)s_d s_{od}L_o\\
=&\eta\frac{1}{w_f}\left(\frac{b_f w_f}{W_{fd}}\right)^{\eta}\left(1-s_f\right)s_d s_{od}L_o=\eta\frac{s_f}{w_f}\left(1-s_f\right)s_d s_{od}L_o\\
=&\eta\frac{n_{fdo}}{w_f}\left(1-s_f\right)
\end{aligned}$$

Where the final expression follows from the fact that $n_{fdo} = s_f\, s_d\, s_{do} L_o$. Similarly, the market partial reduces to

$$\begin{aligned}
&s_f\gamma\left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma-1}\left(\frac{\delta_d}{W_d}\frac{W_{fd}}{w_f}s_f-\frac{\delta_d W_{fd}}{W_d^2}\frac{W_d}{W_{fd}}s_d\frac{W_{fd}}{w_f}s_f\right)s_{od}L_o\\
=&s_f\gamma\left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma}\left(\frac{s_f}{w_f}-s_d\frac{s_f}{w_f}\right)s_{od}L_o\\
=&\gamma s_f s_d s_{od}L_o\frac{s_f}{w_f}\left(1-s_d\right)\\
=&\gamma\frac{n_{fdo}}{w_f}s_f\left(1-s_d\right)
\end{aligned}$$

Finally, the aggregate partial reduces to

$$\begin{aligned}
&s_f s_d\theta\left(\frac{\tau_{od}B_d W_d}{\mathbf{W_o}}\right)^{\theta-1}\left(\frac{\tau_{od}B_d}{\mathbf{W_o}}\frac{W_d}{W_{fd}}s_d\frac{W_{fd}}{w_f}s_f-\frac{\tau_{od}B_d W_d}{\mathbf{W_o}^2}\frac{\mathbf{W_o}}{W_d}s_{od}\frac{W_d}{W_{fd}}s_d\frac{W_{fd}}{w_f}s_f\right)L_o\\
=&s_f s_d\theta\left(\frac{\tau_{od}B_d W_d}{\mathbf{W_o}}\right)^{\theta}\left(\frac{s_f s_d}{w_f}-s_{od}\frac{s_f s_d}{w_f}\right)L_o\\
=&\theta s_f s_d s_{od}L_o\frac{s_f s_d}{w_f}\left(1-s_{od}\right)\\
=&\theta\frac{n_{fdo}}{w_f}s_f s_d\left(1-s_{od}\right)
\end{aligned}$$

Returning to the first order condition, we have

$$\frac{\partial n_{fdo}}{\partial w_f}=\frac{n_{fdo}}{w_f}\left(\eta\left(1-s_f\right)+\gamma s_f\left(1-s_d\right)+\theta s_f s_d\left(1-s_{od}\right)\right)$$

And the labor supply elasticity to firm f from o can be expressed as

$$\varepsilon_{of}=\frac{\partial n_{fdo}}{\partial w_f}\frac{w_f}{n_{fdo}}=\left(\eta\left(1-s_f\right)+\gamma s_f\left(1-s_d\right)+\theta s_f s_d\left(1-s_{od}\right)\right)$$

The first order condition for the firm can then be rewritten to express wages as a share of marginal product

$$
\begin{aligned}
&\sum_o \left(\alpha A_f n_f^{\alpha-1}\, \varepsilon_{of} \frac{n_{fdo}}{w_f} - n_{fdo} - \varepsilon_{of}\, n_{fdo} \right) = 0 \\
&\frac{\alpha A_f n_f^{\alpha-1}}{w_f} \left(\sum_o \varepsilon_{of} n_{fdo} \right) = n_f + \left(\sum_o \varepsilon_{of}\, n_{fdo} \right) \\
&w_f = \alpha A_f n_f^{\alpha-1} \left(\frac{\sum_o \varepsilon_{of} n_{fdo}}{n_f + \left(\sum_o \varepsilon_{of}\, n_{fdo}\right)} \right) \\
&w_f = \alpha A_f n_f^{\alpha-1} \left(\frac{\sum_o \varepsilon_{of} n_{fdo}/n_f}{1 + \left(\sum_o \varepsilon_{of}\, n_{fdo}/n_f\right)} \right)
\end{aligned}
$$

Then we can express the aggregate labor supply elasticity to the firm as

$$
\begin{aligned}
\varepsilon_f =& \sum_o \frac{n_{fdo}}{n_f} \left(\eta \left(1 - s_f\right) + \gamma s_f \left(1 - s_d\right) + \theta s_f s_d \left(1 - s_{od}\right) \right) \\
\varepsilon_f =& \eta \left(1 - s_f\right) + \gamma s_f \left(1 - s_d\right) + \theta s_f s_d \left(1 - \frac{\sum_o s_{od}^2 L_o}{s_d} \right)
\end{aligned}
$$

Which yields the main expression in the text $w_f = mrpl_f \frac{\varepsilon_f}{1+\varepsilon_f}$

B.1 Alternative Assumptions on Labor Competition

In Appendix B.1, I derive the expressions for wages under alternative assumption on the competition between firms.

Perfect Competition Under perfect competition, firms are price takers. This is a less natural analogue since firms are choosing wages in each of the other cases. In this form of competition, firms choose employment and the firms' wage can be expressed as

$$
w_f = \alpha A_f n_f^{\alpha-1}
$$

Firms will pay workers their marginal product, but there will still be variation in wages across firms on account of each firm facing its own labor supply curve.

Monopsonistic Competition Under monopsonistic competition, firms internalize the effect of their wages on labor supply to their own firm, but do not internalize the effect on the local market, *i.e.*

$\partial W_{fd}/\partial w_f = 0$. This implies that

$$\begin{aligned}
\frac{\partial n_{fdo}}{\partial w_f} =& \eta \left(\frac{b_f^{\eta} w_f^{\eta-1}}{W_{fd}^{\eta}} \right) s_d s_{od} L_o \\
\frac{\partial n_{fdo}}{\partial w_f} =& \frac{\eta}{w_f} \left(\frac{b_f w_f}{W_{fd}} \right)^{\eta} s_d s_{od} L_o \\
\frac{\partial n_{fdo}}{\partial w_f} =& \eta \frac{n_{fdo}}{w_f} \\
\rightarrow \varepsilon_f = \eta &
\end{aligned}$$

That is, the markdown is constant.

Intermediate Local Oligopoly Under this form of competition, firms internalize the effect of their wage on the wages in other firms, but not how that affects total labor supply to firms. That is, $\partial W_{fd}/\partial w_f \neq 0$ but $\partial W_d/\partial W_{fd} = 0$. Following the same logic as above, the labor supply elasticity here is

$$\varepsilon_f = \eta \left(1 - s_f\right) + \gamma s_f$$

Local Oligopoly A more sensible version of local oligopoly is that in which firms internalize the effect of changes in their wage on each of the aggregates except for the aggregate market indices, That is, $\partial \mathbf{W_o}/\partial W_d = 0$. This is consistent with each market being small relative to the entire market. Here the labor supply elasticity is

$$\varepsilon_f = \eta \left(1 - s_f\right) + \gamma s_f \left(1 - s_d\right) + \theta s_f s_d$$

B.2 Additional Results

Employment Elasticity Total firm-employment in each market can be expressed as the sum of employment in each firm

$$n_d = \left(\frac{\delta_d W_{fd}}{W_d} \right)^{\gamma} \left[\sum_o \left(\frac{\tau_{od} B_d W_d}{\mathbf{W_o}} \right)^{\theta} L_o \right]$$

The employment elasticity can be found by differentiating n_d with respect to W_{fd}

$$\begin{aligned}\frac{\partial n_d}{\partial W_{fd}} =&\gamma\left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma-1}\left(\frac{W_d\delta_d-\delta_d W_{fd}\frac{\partial W_d}{\partial W_{fd}}}{W_d^2}\right)s_m(d)\\ &+\sum_o s_d\theta\left(\frac{\tau_{od}B_dW_d}{\mathbf{W_o}}\right)^{\theta-1}\left(\frac{\mathbf{W_o}\tau_{od}B_d\frac{\partial W_d}{\partial W_{fd}}-\tau_{od}B_dW_d\frac{\partial \mathbf{W_o}}{\partial W_d}\frac{\partial W_d}{\partial W_{fd}}}{\mathbf{W_o}^2}\right)L_o\\ =&\frac{\gamma}{W_{fd}}\left(\frac{\delta_d W_{fd}}{W_d}\right)^{\gamma}(1-s_d)\,s_m(d)+\sum_o s_d\theta\left(\frac{\tau_{od}B_dW_d}{\mathbf{W_o}}\right)^{\theta}\left(\frac{s_d}{W_{fd}}-\frac{s_d}{W_{fd}}s_{od}\right)L_o\\ =&\gamma\frac{n_d}{W_{fd}}(1-s_d)+\theta\sum_o\frac{s_ds_{od}L_o}{W_{fd}}s_d(1-s_{od})\end{aligned}$$

For s_{od} small, this can be approximated as

$$\begin{aligned}\frac{\partial n_d}{\partial W_{fd}} =&\gamma\frac{n_d}{W_{fd}}(1-s_d)+\theta\frac{n_d}{W_{fd}}s_d\\ \rightarrow \varepsilon_d =&\gamma(1-s_d)+\theta s_d\end{aligned}$$

Migration Appealing to the expression for s_{od}, total migration from $o-d$ is given by

$$n_{od}=s_{od}L_o=\left(\frac{\tau_{od}B_dW_d}{\mathbf{W_o}}\right)^{\theta}L_o$$

Then the ratio of migration from $o-d$ to non-migration is

$$\frac{n_{od}}{n_{oo}}=\left(\frac{B_d\tau_{od}W_d}{B_oW_o}\right)^{\theta} \tag{15}$$

Rearranging (15) yields the estimation equation in the main text

$$n_{od}=n_{oo}\exp\left[\theta\log\left(\frac{W_d}{W_o}\right)+\theta\log\tau_{od}+\theta\log\left(\frac{B_d}{B_o}\right)\right]$$

Where the migration cost terms are estimated via the existing migration network and the distance between pairs and the ratio of the amenities are the error terms. Returning to (15), the exact τs are finally calculated as

$$\frac{n_{od}}{n_{oo}}\frac{n_{do}}{n_{dd}}=\left(\frac{B_d\tau_{od}W_d}{B_oW_o}\frac{B_o\tau_{do}W_o}{B_dW_d}\right)^{1+\theta}=\left(\tau_{od}\tau_{do}\right)^{\theta}$$

By the assumption that $\tau_{od}=\tau_{do}$, the migration cost is

$$\tau_{od}=\left(\frac{n_{od}}{n_{oo}}\frac{n_{do}}{n_{dd}}\right)^{\frac{1}{2\theta}}$$

Ten-year Migration Flows Let π_{od} denote the share of migrants from o to d. The number of individuals remaining in o after one year is then given by $n_o * \pi_{oo}$. Appealing to this logic, the total number of migrants in a ten year period can be extrapolated from the number of migrants in a single year as

$$
\begin{aligned}
n_{od}^{10} =& n_o * \pi_{od} + n_o * \pi_{oo} * \pi_{od} + n_o * \pi_{oo}^2 * \pi_{od} + \ldots + n_o * \pi_{oo}^9 * \pi_{od} \\
n_{od}^{10} =& n_o * \pi_{od} \left(1 + \pi_{oo} + \pi_{oo}^2 + \ldots \pi_{oo}^9\right) \\
n_{od}^{10} =& n_o * \pi_{od} \left(\frac{1 - \pi_{oo}^{10}}{1 - \pi_{oo}}\right)
\end{aligned}
$$

Where the final line follows from the difference of the two infinite geometric sums. Then our decadal migration shares can be expressed as

$$
\pi_{od}^{10} = \pi_{od} \left(\frac{1 - \pi_{oo}^{10}}{1 - \pi_{oo}}\right) \quad ; \quad \pi_{oo}^{10} = \pi_{oo}^{10}
$$

C Minimum Wage Law

Tanzania's Employment and Labour Relations Act, 2004 set forth a broad set of regulations "to promote economic development through economic efficiency, productivity and social justice" (p. 6). Links to the source document held by the International Labor Organization online can be found here. The act applied to all laborers in the country, both public and private, except for those in the People's Defense force, the police force, prison services and the national service (Employment and Labour Relations Act, 2004, p. 5). The legislation made it illegal for children under the age of 14 to work (and under the age of 18 in hazardous sectors). The maximum number of usual hours that an employee could work were set at 9 hours per day, 6 days per week and 45 hours per week (p. 19). The penalty for violating these laws was up to one year in jail and a 5 million shilling fine (p. 79). The legislation did not set an official binding minimum wage, but made a provision to set one within three years, laying the groundwork for creating a national minimum wage (p. 84).

Progress continued with the 2007 Labour Institutions Act, which allowed for the creation of sectoral wage boards that would determine the minimum wage within their sector. In 2010, minimum wages were passed into law in eight sectors. I report the monthly minimum wages for these sectors in Column 3 of Table 23. Each minimum wage stipulated an hourly, daily, weekly, fortnightly, and monthly rate. Adding further complexity, several sectors provided different levels for subsectors, creating 20 total minimum wages. The differences within sectors could be large; for domestic and hospital services the monthly minimum wage ranged from 65,000 TSH for domestic workers to 150,000 TSH for tourist hotels. Employers in any sector not mentioned were required to pay all employees at least 80,000 TSH per month for full time work. The "all other sectors" minimum wage equaled that for health services and commerce, industry, and trade, and exceeded the minimum wage for agriculture (70,000 TSH per month) and domestic workers (65,000 TSH per month). These sectoral minimum wages remained in place until July 2013 when the 2010 Wage Order was repealed and replaced.

Table 23: 2010 Sectoral Minimum Wage Details

Sector	ISIC Code	Monthly Wage
Agricultural Services	1-2	70000
Marine and Fishing	3	165000
Mining Primary Licenses	5-9	150000
Mining License/ Prospecting licenses	NA	350000
Mining Dealers licenses	NA	250000
Mining Brokers licenses	NA	150000
Trade, Industry and Commerce	10-33	80000
Transport Services: Inland Transport	49, 491-493	150000
Transport Services: Aviation	51	350000
Transport Services: Clearing and Forwarding	52	230000
Hotels: Medium Hotels	55	100000
Hotels: Potential and Tourists hotel	NA	150000
Hotels: Restaurants, Guest Houses and Bars	56	80000
Telecommunication	61-63	300000
Private security: other	80	80000
Private security: International or potential security Companies	NA	105000
Health Services	86-88	80000
Domestic Services: Other	97 98	65000
Domestic Services: Diplomats	NA	90000
Domestic Services: Entitled Officers	NA	80000
Domestic Services: Other	97 98	65000
Other	35, 36-39, 50, 52, 53	80000

Notes: The table reports for each sector that was specified in the 2010 Wage Order, the corresponding ISIC codes used in the analysis and the monthly minimum wage.

Table 24: EES Firm Employment Summary Statistics

	min wage	wage	employees	casual	hires	female
	(1000 TSH)		% of total			(%)
	(1)	(2)	(3)	(4)	(5)	(6)
Agriculture						
2005-2007		123.3	4.2	23.8		30.0
2010-2013	70	317.4	3.9	31.9	2.6	30.2
2014 - 2017	100	413.2	4.1	25.1	3.9	22.0
Fishing						
2005-2007		105.9	0.2	0.5		18.3
2010-2013	165	262.2	0.3	0.4	0.0	13.5
2014 - 2017	200	173.7	0.2	0.4	0.0	9.7
Mining						
2005-2007		203.9	0.5	2.2		12.8
2010-2013	150	407.7	0.9	1.1	0.2	12.2
2014 - 2017	200	664.5	1.4	1.1	0.9	8.1
Manufacturing, Commerce, Trade						
2005-2007		195.8	13.0	42.2		22.5
2010-2013	80	236.2	21.7	41.2	16.6	24.2
2014 - 2017	115	356.0	25.1	37.2	13.7	19.3
Energy Services						
2005-2007		441.5	0.7	0.8		17.2
2010-2013	80	766.1	0.8	0.4	0.5	19.0
2014 - 2017	150	933.6	1.2	0.2	0.2	23.1
Construction						
2005-2007		248.9	1.5	5.7		15.4
2010-2013	80	331.0	2.1	5.0	1.2	15.9
2014 - 2017	250	493.9	2.2	7.0	1.6	8.4
Inland Transport						
2005-2007		199.8	1.2	0.3		9.3
2010-2013	150	350.1	1.5	0.9	1.1	13.0
2014 - 2017	200	419.2	1.5	1.4	1.8	7.8
Aviation Services						
2005-2007		323.9	0.2	0.1		24.6
2010-2013	350	534.6	0.1	0.1	0.0	38.9
2014 - 2017	300	1084.5	0.2	0.0	0.1	21.9
Clearing and Forwarding						
2005-2007		151.9	1.2	1.7		17.8
2010-2013	230	432.5	1.5	1.7	1.5	22.7
2014 - 2017	300	624.6	1.1	0.4	0.7	17.8
Hotels						
2010-2013	100	164.3	4.0	1.8	4.9	49.7
2014 - 2017	150	195.7	4.2	2.3	6.0	29.8

...Continued on next page

EES Firm Employment Summary Statistics – *continued from previous page*

	min wage	wage	employees	casual	hires	female
	(1000 TSH)		(%) of total			(%)
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurants						
2005-2007		88.6	3.9	3.1		50.9
2010-2013	80	163.1	3.7	2.5	4.0	56.8
2014 - 2017	130	188.8	2.7	1.4	4.4	27.1
Information Services						
2005-2007		267.2	1.4	1.0		37.4
2010-2013	80	447.0	0.8	0.2	0.3	38.7
2014 - 2017	150	585.8	1.0	0.6	0.6	25.1
Telecommunication Services						
2005-2007		263.3	1.2	1.0		38.7
2010-2013	300	795.1	0.6	0.0	0.4	32.0
2014 - 2017	400	824.3	1.1	10.6	0.9	24.7
Financial Services						
2005-2007		1542.1	3.2	0.2		40.8
2010-2013	80	842.7	2.1	0.1	2.6	42.9
2014 - 2017	400	1227.3	2.1	0.2	2.4	29.7
Private Security						
2010-2013	80	183.7	1.5	0.4	2.8	18.8
2014 - 2017	100	200.3	2.7	0.4	6.2	12.0
Education						
2005-2007		208.5	11.7	4.3		32.3
2010-2013	80	433.9	17.3	2.3	27.8	43.5
2014 - 2017	140	656.7	19.7	2.9	26.3	28.6
Health Services						
2005-2007		186.8	6.3	0.9		56.3
2010-2013	80	384.9	7.3	1.6	9.0	60.4
2014 - 2017	132	507.5	7.3	1.9	11.6	44.2
All Others						
2005-2007		288.8	49.6	12.1		31.5
2010-2013	80	436.2	29.7	8.3	24.4	34.9
2014 - 2017	100	640.6	22.2	7.1	18.6	24.4

Notes: Reporting the average values during each of the three periods. Columns(1) and (2) report the minimum wage and average monthly wage in thousands of Tanzanian Shillings. Columns (3)-(5) report the employment, casual employment, and hires as a percentage of the total. Column (6) reports the share of workers in that sector-period that are female.

Table 25: EES Firm Size Summary Statistics

	min wage	Firms	employees (% of total)		Private	Districts
	(1000 TSH)	Total	5-49	50+	(%)	
	(1)	(2)	(3)	(4)	(5)	(6)
Agriculture						
2005-2007		1722	59.6	6.0	96.7	105
2010-2013	70	1090	70.1	15.2	85.5	111
2014 - 2017	100	1584	69.3	19.0	95.0	100
Fishing						
2005-2007		205	47.7	2.3	73.6	38
2010-2013	165	240	93.4	2.5	93.1	29
2014 - 2017	200	382	98.5	1.2	98.8	12
Mining						
2005-2007		142	69.8	15.6	89.9	30
2010-2013	150	161	81.7	17.0	97.2	29
2014 - 2017	200	508	77.4	15.0	97.4	60
Manufacturing, Commerce, Trade						
2005-2007		7321	67.5	4.3	94.9	96
2010-2013	80	15735	62.9	3.8	98.7	114
2014 - 2017	115	23876	61.9	4.3	98.9	124
Energy Services						
2005-2007		317	71.9	9.0	5.9	74
2010-2013	80	138	55.9	37.1	23.1	69
2014 - 2017	150	130	33.1	64.2	41.6	39
Construction						
2005-2007		860	76.8	4.0	87.5	69
2010-2013	80	1422	80.1	5.7	95.9	88
2014 - 2017	250	2182	85.2	6.0	98.2	67
Inland Transport						
2005-2007		493	51.6	7.8	70.6	40
2010-2013	150	837	73.4	6.3	95.7	42
2014 - 2017	200	1231	78.2	8.8	93.6	55
Aviation Services						
2005-2007		114	69.2	3.6	57.5	12
2010-2013	350	62	74.1	5.9	58.5	14
2014 - 2017	300	154	44.6	6.1	70.2	20
Clearing and Forwarding						
2005-2007		632	67.0	4.4	82.5	21
2010-2013	230	585	77.2	8.8	80.5	41
2014 - 2017	300	683	79.5	7.8	86.2	31
Hotels						
2010-2013	100	4084	70.8	2.2	100.0	100
2014 - 2017	150	6707	68.5	2.7	100.0	113

...Continued on next page

EES Firm Size Summary Statistics – *continued from previous page*

	min wage (1000 TSH)	Firms Total	employees (% of total) 5-49	50+	Private (%)	Districts
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurants						
2005-2007		4899	68.7	0.9	99.3	83
2010-2013	80	5829	71.4	0.5	99.6	102
2014 - 2017	130	6994	74.5	0.4	100.0	105
Information Services						
2005-2007		738	66.8	5.7	56.3	49
2010-2013	80	502	52.5	6.2	51.9	95
2014 - 2017	150	768	81.5	9.2	76.2	51
Telecommunication Services						
2005-2007		2314	16.8	1.4	48.9	108
2010-2013	300	249	57.6	9.2	75.2	59
2014 - 2017	400	675	66.5	8.2	87.2	57
Financial Services						
2005-2007		2686	88.0	1.4	16.4	103
2010-2013	80	1269	66.3	4.5	59.4	89
2014 - 2017	400	2166	68.6	4.3	58.3	82
Private Security						
2010-2013	80	357	70.3	25.6	98.7	34
2014 - 2017	100	779	73.1	24.1	100.0	41
Education						
2005-2007		7092	82.5	3.0	15.3	116
2010-2013	80	3063	76.7	11.2	81.4	124
2014 - 2017	140	5080	76.9	13.9	81.3	123
Health Services						
2005-2007		2479	69.7	10.0	52.1	113
2010-2013	80	2480	64.6	13.7	86.6	123
2014 - 2017	132	3786	69.8	15.6	84.5	123
All Others						
2005-2007		25986	52.5	4.0	11.3	119
2010-2013	80	11230	57.6	8.6	50.0	123
2014 - 2017	100	14083	62.1	8.2	60.4	123
All						
2005-2007		58009	61.5	3.9	39.2	119
2010-2013		49332	65.5	6.1	83.6	124
2014 - 2017		71783	67.4	6.8	87.5	124

Notes: Reporting the average values during each year in each of the three periods. Column (2) reports the total number of firms in that sector-period. Columns (3)-(6) report the share of those firms with 5-49 employees, at least 50 employees and that are private, respectively. Column (6) reports the number of districts in which there are at least one firm in that sector operating in the sample.