Urbanization and the Routes to Structural Transformation*

Juan Soto-Díaz, Olga M. Vargas, and Julio A. Berdegué

[Preliminary Draft]

[Download the most recent version here]

June, 2024

Abstract

How urbanization shapes structural change out of agriculture? To what extent this is mediated by productivity shocks in agriculture? How transport infrastructure development can be better designed to account for these effects? This article estimates the effects of urbanization and road infrastructure development on the structural transformation of rural villages in Chile. Following a market access approach derived from a spatial quantitative model of structural transformation that is informed by the elasticities of urban market access on the population, and farm and non-farm employment of rural villages. The results support the hypothesis of the diversification of the rural economy that is also consistent with the intensification of agriculture. Moreover, the evidence suggests important heterogeneous effects across rural areas, that reveal that transport infrastructure development can be better designed to take advantages areas with better conditions for agricultural production.

*Soto-Díaz: School of Economics, The University of Queensland (UQ), St. Lucia 4072, Brisbane, Australia (email: juan.soto@uq.edu.au). Department of Geography and Environment, London School of Economics and Political Science London, UK. Vargas: Latin American Center for Rural Development -Rimisp-, Santiago, Chile. Berdegué: Latin American Center for Rural Development -Rimisp-, Santiago, Chile. The authors are grateful to ANID Fondecyt 1161424. Especially thanks to Guillermo and Jorge Soto-Díaz for excellent research assistance. We are grateful for insightful comments from: Hans Binswanger-Mkhize, Chiara Cazzuffi, Luc Christiaensen, Riccardo Crescenzi, Vernon Henderson, Carlo Pietrobelli, Andrés Rodríguez-Pose, Francisco Rowe, Daria Taglioni, Hanne Van Cappellen, and multiple participants of seminars at LSE, Brunel, UniOviedo, Rimisp, UAI, UTalca, and UCN-Chile.

...if we endow an underdeveloped country with a first-class highway network, with extensive hydroelectric and perhaps irrigation facilities, can we be certain that industrial and agricultural activity will expand in the wake of these improvements? Would it not be less risky and more economical first to make sure of such activity...

Albert O. Hirschman, *The Strategy of Economic Development* (1958)

1 Introduction

It is widely documented that modern economic growth has been largely explained by a move of workers from agriculture to more productive economic sectors (Herrendorf et al., 2014). Since the global spread of the industrial revolution, this process of structural change has been strongly correlated with urbanization rates, which increased living standards in urban areas (Bairoch, 1988). However, the opposite is not necessarily true, i.e., the extent to which urbanization induces growth and economic development in rural areas. Features such as the economic nature of small rural economies, the exposure of farms to exogenous climatic productivity shocks, and the different stages of development across the geographical space, suggest the existence of important heterogeneity and nonlinearities. Nonetheless, based on this relationship between the urban and the rural economy, over the last decades an increasing number of policies in the developing world promote the creation of the so-called *urban-rural linkages*, to foster trade integration, agricultural intensification, and the growth of non-farm activities in rural areas.[1] For this argument of urbanization with rural development to work, however, this would imply that agricultural productivity is an increasing function of access to urban markets and apart from migration to urban areas, there is a movement towards the non-agricultural sector within rural areas in order to keep rural population growing.[2]

The idea behind the promotion of *urban-rural linkages*, therefore, relies on this argument that the net effect of growth in urban areas in the growth and development of rural areas is positive. In other words, that the consumption and production linkages in the form of positive and negative spillovers from urban growth, that drive the pattern of rural-to-urban migration, employment, and productivity in rural areas, are virtuous, on average.[3] The mechanisms that define the sign and magnitude of these geographical spillovers from urbanization, are key in the current and future design of rural development policies (World Bank, 2008, 2009; United Nations, 2015, 2017; FAO, 2017), and therefore, crucial to identify. Without a clear understanding of how urbanization and transport infrastructure development connecting rural areas to cities, are actually generating economic development in rural areas, seems problematic to foster policies that eventually would have negligible impacts (see for example, Asher and Novosad, 2020 evaluation of the 1 billion rural roads program in India).

[1] For example, six points of the UN Sustainable Development Goals (United Nations, 2015) are dedicated to this topic, and also developed in World Bank (2008) and World Bank (2009).

[2] Note that this growth in rural population is still seeing in the developing world. Despite growth rates, in both, urban and rural areas has been declining over the last two decades.

[3] Also known as *spread* and *backwash effects* (Colby, 1933; Gaile, 1980; Hughes and Holland, 1994; Barkley et al., 1996; Chen and Partridge, 2013), or *urban-rural linkages* (Berdegué et al., 2014).

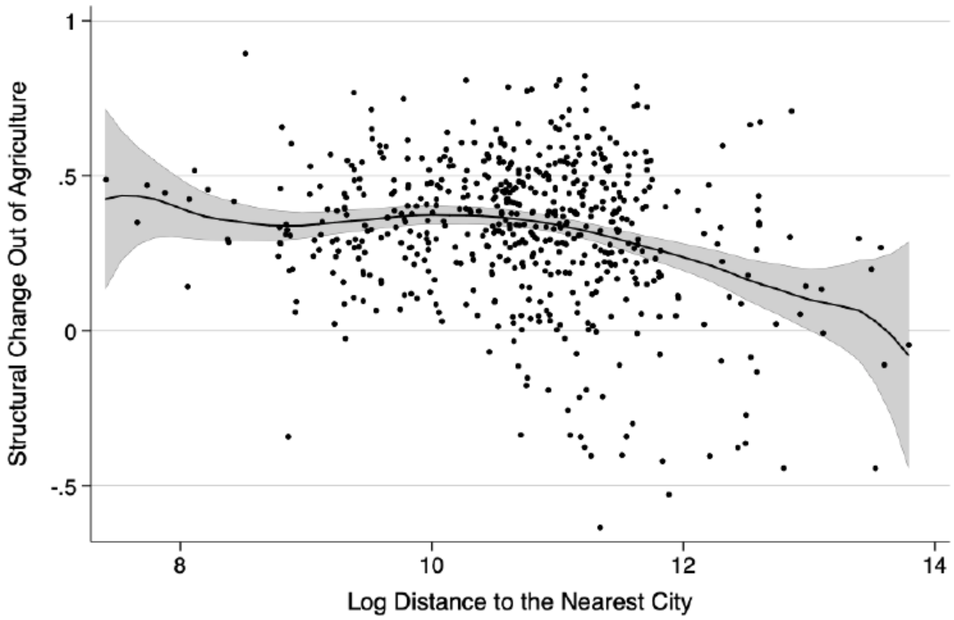
The notion that as small rural locations growth the agricultural sector will represent a lower size of their economies, is consistent with the existing evidence showing that in rural areas there is also a process of structural transformation out of farm activities (Reardon et al., 2001, 2009), which is underlying increases in agricultural productivity and more aggregate trends in structural transformation. The extent to which these issues are related to urbanization, however, requires a better understanding of the mechanisms that determine the degree of complementarity between urban and rural activities, and agricultural and non-agricultural activities in rural areas. Notwithstanding, these hypothesis has been challenging to test because of the lack of good data. In particular, an important limitation is that the spatial granularity of the different mechanisms at play and data restrictions, imply that many of these urban-to-rural spillovers are difficult to measure. To do a fair comparison, we have to define rural and urban areas as spatial entities comparable over a long-period of time, and find data about this process of structural transformation and the linkages between urban and rural areas that is also at this level of spatial granularity. These features make difficult to identify the gains for rural areas from policies that improve access to urban markets, with effects that are expected to be highly heterogeneous given the small size of most rural economies.[4]

To illustrate these arguments more clearly on the data, Fig. 1 shows for Chile, our case study, the relationship between the change in the share of non-agricultural employment in rural villages and the distance to the nearest city, a measure of access to urban markets. Consistent with the arguments made by agricultural economists (Berdegué et al., 2001; Reardon et al., 2001, 2009), the negative relationship between urbanization and the structural transformation of the rural economy seems to hold even at a small spatial scale. I.e, on average, as rural locations have more access to urban areas they tend to diversify by moving away from agricultural activities. However, the figure describes a negative weak correlation that is explained by the large heterogeneity in both, geographical location and specialization in the agricultural sector. Which have important implications for policy design on the arguments of transport infrastructure investment and rural development.

Motivated by these facts, in this paper we estimate the effects in the growth and structural change out of agriculture in rural villages from gains in access to urban markets derived from road infrastructure development and urbanization. Specifically, we construct a granular spatial dataset using population and agricultural censuses, roads networks, and remote sensing data. Then, we estimate the effects of access to urban markets —induced by urban growth and road infrastructure development— on the population, agricultural and non-agricultural employment, and agricultural potential (proxy of agricultural productivity), for more than 500 rural communities/villages in Chile over 25 years. Our results show an elasticity of market access for the population of rural communities of around 1.0 to 1.4 in our preferred estimations, and a positive significant effect in non-farm employment. These results are robust to different specifications for market access. [5]

[4]Such as for example the aggregate effects of integrating remote rural communities that might induce gains from trade but at the same time depopulation. Despite that even with rural decline can be individual welfare gains of rural population moving to cities. Even if it does in the average or it can generate welfare gains from rural to urban migrants.

[5]Variation in this market access variable comes from changes in urban growth and road networks.



Notes: The figure shows the association between the logarithm of the distance to the nearest city and the change in the share of non-farm employment for a sample of more than 500 rural villages/communities identified across the censuses of 1992 and 2017 in Chile. The distance to the nearest city is constructed computing the straight euclidean distance between each rural village and each one of the 170 urban areas in Chile. The figure displays the local polynomial fit of the scatter plot using an Epanechnikov kernel function with a second-order degree. *Source:* Own elaboration based on data from the Chilean National Office of Statistics (INE).

Figure 1: Access to Urban Markets and Structural Change Out of Agriculture

This paper builds on the shoulders of contributions aiming to understand the causes and consequences of spatial structural transformation (e.g., Michaels et al., 2012b), the local economic impacts of transport infrastructure improvements (e.g., Asher and Novosad, 2020; Storeygard, 2016; Jedwab and Storeygard, 2022), the consequences of urbanization to spatial development (e.g., Michaels et al., 2012b), and the long-term dynamics of economic development in rural areas (e.g., Christiaensen et al., 2013). This paper distinguishes from these literature, however, in that this study is an effort to give a comprehensive view on these topics, providing evidence on the mechanisms explaining the heterogeneity on the effects of urbanization and transport infrastructure development on the growth and economic development of rural areas for an emerging economy. This is important not only from a practical policy view, but also from a theoretical perspective. Because the existing evidence is ambiguous in establishing to what extent economic growth in more populated areas can also lead to spatial development, i.e. growth and economic development in less dense areas.

This article contributes to the literature in two different ways. First, we present a framework that allows us to understand the average and heterogeneous effects of urban growth on the economic development of rural areas, under the assumption of spatial equilibrium *á la Roback*, in which variations in population and employment of rural communities are informative of localization incentives of rural workers and firms that reallocates to compensate for differences in utility levels and costs across space. In addition, a wide existing empirical evidence on the impact of cities on the growth of rural areas are generalizations of the empirical partial adjustment models of Carlino and Mills (1987), which are usually not explicitly microfounded (e.g., Henry et al., 1999; Deller et al., 2001; Carruthers and Vias, 2005), with only few studies distinguishing between farm and non-farm rural employment. Notwithstanding, this distinction is more common in the literature on agricultural economics, where the evidence at household level suggest a positive impact on farm and non-farm income, and the intensification of the agricultural activity (for some recent evidence see: Binswanger-Mkhize and Savastano, 2017; Davis et al., 2017; Vandercasteelen et al., 2018).

Second, using a market access approach and following recent advances in measuring the impact of infrastructure on local economic development (Donaldson and Hornbeck, 2016; Jedwab and Storeygard, 2022), we present a methodology to estimate the influence of urban markets on the economic development of rural communities using the richness of population censuses, and remote sensing data. Following the recent increase in the use of satellite imagery data in the contexts of unavailable or less reliable official information (Henderson et al., 2012; Donaldson and Storeygard, 2016). In addition, this paper takes advantages of the remote sensing data to have a clear distinction between rural and urban areas, and in the same vein as Michaels et al. (2012b), we adopt a more disaggregated spatial unit of analysis (community/village/locality) that leads to a better identification of the heterogeneity on the effects of urban growth on the economic development of rural areas.[6]

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature. Section 3 presents the model. Section 4 details the econometric specifications. Section 5 describes the data and estimation issues. Section 6 shows the results and, finally, Section 7 summarizes main results and concludes the paper.

2 Literature Review

The agricultural economics literature proposes multiple causal mechanisms to understand the relationship between urbanization and the economic development of rural areas.[7] First, rural locations close to and connected to a densely populated area are more likely to receive more gains of trade than remote ones. This is because cities represent large markets for the commercialization of agricultural goods produced in rural areas, which would raise the prices of these goods, increasing farmers profits from trade (Jacoby, 2000; Donaldson and Hornbeck, 2016), and intensifying agricultural production (Binswanger-Mkhize and Savastano, 2017; Damania et al., 2017; Davis et al., 2017). Trade between rural communities and cities could be one of the main channels for rural economic development (Fleming

[6]Counties, districts, or municipalities are usually classified as rural when a large percentage of their population is living in rural villages. However, these spatial units are not usually entirely rural. As such, researchers cannot distinguish if growth is occurring only in rural communities, thus overestimating or underestimating the impact of urban growth on the growth and economic development of rural communities. In this study, rural areas are defined by the census as an agglomeration of rural villages.

[7]For a good critical summary on this relationship, see Wu et al. (2016).

and Abler, 2013) and can contribute to its productive specialization, as well as to the diversification of its economy.

The economic specialization of rural areas tends to be concentrated toward agricultural products that require lower travel times from the production places to the market (Beckmann, 1972; Costinot and Donaldson, 2012). This specialization could lead to higher wages in the agricultural sector, derived from the division of labor (Yang and Liu, 2012). Nonetheless, the increasing demand for jobs in activities auxiliary to agricultural production (e.g., transportation, manufacturing, or sales) could also foster the economic diversification of rural communities through an increase in the number of non-farm activities (Foster and Rosenzweig, 2004). This diversification may also raise wages in the non-farm sector (Berdegué et al., 2001). However, this is not the rule, since some workers in non-farm jobs in rural or even urban areas may earn less than the average wage in the agricultural sector (Perloff, 1991; Lanjouw and Lanjouw, 2001).

The increase in rural non-farm employment could have two different origins. The first is the demand for jobs that are auxiliary to the agricultural sector due to the increasing number of non-farm activities in growing agricultural markets (Foster and Rosenzweig, 2004). The second is the demand for jobs in cities, which might increase the number of commuters living in rural communities and working in cities (Renkow, 2003). Both causes are likely to be related to low-skill jobs from rural communities in proximity to cities (So et al., 2001). Moreover, rural workers may also be working in remote cities or highly productive extractive places by long-distance commuting; however, this would typically be associated with medium-skill jobs or seasonal workers (Paredes et al., 2018).

The number of amenities and employment opportunities in cities could also influence the localization incentives of rural households, causing important migration flows from rural areas to cities, which lead to the decline of the population in rural communities (Goetz and Debertin, 2001). However, in theory, this process may also have a positive effect on the labor productivity of the rural place of origin.[8] This is because rural-urban migration reduces the number of agricultural workers and, since the agricultural production is subject to constant returns to scale, it increases the marginal productivity of each worker in that agricultural area (Harris and Todaro, 1970), accompanying a process of structural change in the economy (Alvarez-Cuadrado and Poschke, 2011; Michaels et al., 2012b). Similarly, the migration of rural workers to cities could, at the same time, increase the amount of remittances that families in the rural area of origin receive (Banerjee, 1984).

Another potential channel of the impact of urbanization on the development of rural communities is the rent of land (Thunen, 1826). Rural households and landowners have access to lower consumer prices and land rent than city residents, allowing a lower cost of living (Kurre, 2003; Loveridge and Paredes, 2016), and consequently a higher quality of life (Roback, 1982; Deller et al., 2001), which is also due to natural amenities such as open spaces (Klaiber and Phaneuf, 2009) or the access to services

[8]This holds only in theory because the empirical findings on this topic are ambiguous at this point. Some evidence of a more rapid growth in agriculture than in the non-agriculture sector are: Griliches (1957); Rasmussen (1962); Maddison (1980).

that nearby cities might offer. Such price advantages in rural areas are widely exploited in urban processes, such as the relocation of the manufacturing industry (Lonsdale and Browning, 1971) and suburbanization (Lopez et al., 1988; Burchfield et al., 2006). The price of the land in rural locations near cities may also be affected by these factors, leading to price increases as market access improves.

The above review suggest that the positive effects derived from the trade between rural communities and cities may have a large spatial scope compared to the other mechanisms, which usually have a reduced spatial scope and ambiguous outcomes (Irwin et al., 2009; Castle et al., 2011). In fact, empirical works relating the impacts of urban growth on rural development at county or municipality level found ambiguous results for the effect of this relationship on employment (e.g., Chen and Partridge, 2013 found a non-significant effect on employment in rural areas in China and a negative effect when only capital districts are considered), but also identified consistent positive effects on the population (Partridge et al., 2009). However, these results have to be carefully interpreted due to aggregation in their units of analysis (Briant et al., 2010). On the other hand, evidence at the household level indicates a positive association between market access and rural wages in the farm and non-farm sector (Binswanger-Mkhize et al., 2016). Therefore, this article uses a market access approach on measuring and understanding the role of urbanization and transport infrastructure development in the growth and structural transformation of rural communities.

3 Theory

The idea that access to urban markets is associated with a process of structural change out of agriculture in rural areas seems to be supported in Fig. 1 for the case study, but also suggest important heterogeneity. These facts are consistent with a Ricardian world in which different locations trade based on comparative advantages but are also affected by exogenous productivity shocks as driving technological differences and inducing the observed heterogeneity. With population and employment growing accordingly to this and respective migration frictions. Consequently, to inform our empirical analysis, we rely on a canonical new Ricardian trade model that builds over Eaton and Kortum (2002), and follows closely the work of Donaldson (2018) and Donaldson and Hornbeck (2016). The model will describe theoretically the mechanisms and provide a simple guide to the empirical identification.

3.1 Model Environment

The intuition of the model is that changes in urbanization and transport infrastructure development are captured by a market access effect. An increase in market access would generate a positive demand shock in a discrete number of locations that will move activities from a basic food sector producing a homogeneous non-tradable good, towards a tradable manufacturing sector producing a continuum of differentiated varieties. There is no a priory discrete distinction between rural and urban areas. The conceptual difference between both is that cities are large locations more specialized in the production of a wide range of manufacturing varieties, while rural areas are small locations that rely more on the production of a non-tradable food sector.[9] Then, structural transformation will imply a move of workers towards the production of manufacturing varieties, which for the case of rural areas can also be thought as agricultural products with more value added or less complex manufacturing goods than those produced in cities.[10] The idea is that as small rural locations get more connected to large agglomerations, this increase in market access would imply that they would start to rely more on the production and trade of manufacturing varieties.[11]

On the other hand, the heterogeneous effects of market access among different locations are explained by exogenous productivity shocks. These shocks would determine the degree of specialization in the tradable manufactured good, and therefore affect the pattern of structural transformation induced by an increase in market access. Changes in market access would also affect technological adaptation in each location. These effects would induce a complementarity between productivity shocks and market access that is consistent with existing evidence documenting heterogeneous effects of transportation infrastructure investments and a growing literature on the effects of climate change in rural areas and agricultural production (Barwick et al., 2021).[12] The final equilibrium in rural areas would be determined by the relative strength of each one of these forces.

3.2 Demand

Let assume a small open economy consisting of a discrete number of locations $\Im$ indexed by an origin i and a destination j. Preferences for a representative individual in a given destination j are represented by

$$U_j = n_j C_j^{\delta} F_j^{1-\delta} \tag{1}$$

[9] This form of modelling structural transformation in rural areas is flexible enough to be consistent with situations in which the move from farm to non-farm employment might happen in activities that are related with the agricultural sector but rely more on some manufacturing processing, or activities related to the distribution and commercialization of those manufactured agricultural goods. These activities have shown to be important for rural economic development (Reardon et al., 2001, 2009).

[10] As in Eaton and Kortum (2002), the specific goods that are traded are not relevant to the model. This is because one of the features of working with an stochastic component of productivity imply that all the adjustments are through the extensive margin of trade, given that the conditional distribution of prices are going to be equal to its unconditional distribution disregarding the origin location of those traded goods. In our context, this imply that smaller rural locations would have lower productivity and a smaller number of varieties produced there. However, more recently, Moneke (2020) worked a three sector model extension of Michaels et al. (2012a) to make more clear the distinctive patterns of structural transformation by understanding better the complementarity/substitution between sectors.

[11] For our purposes there is no much gain in distinguishing a priory between rural and urban locations in a discrete form, although we made this separation in the empirics. As described in Donaldson and Hornbeck (2016), the distinction between rural and urban locations would lead to four different types of isomorphic market access measures that are empirically equivalent to the derived in this model. Therefore, to preserve parsimony in the theory, we instead consider that rural areas are locations less specialized in the production of manufacturing varieties, while urban areas are places more specialized in the production of a large variety of complex manufacturing tradable goods. In consequence, structural transformation in this paper is the transition towards a larger share of the workforce specialized in these complex tradable goods. This is consistent with our empirics in which we are particularly interested in small communities/villages that move from farm to non-farm employment due to increases in access to large markets.

[12] These ideas are also consistent with Gollin et al. (2016) and recent insights from the work of Vandercasteelen et al. (2018) and Damania et al. (2017) on the role of cities in agricultural intensification and transformation. These models summarize a large body of literature on structural change and agriculture (see Herrendorf et al., 2014 for a detailed survey).

where $\delta, \in (0,1)$. The terms C_j and F_j describe the amount of consumption of a composite manufacturing and an homogeneous agricultural good respectively. n_j is an exogenous preference shifter representing local amenities. Individuals maximize their utility subject to a budget constraint given by $P_jC_j + F_j = Y_j$, where the agricultural good is the numeraire. In consequence, the optimal demand for the agricultural good is $(1-\delta)Y_j$ and the share of income spent in the composite good is δY_j, which is allocated according to the following CES preferences, $C_j = \left(\sum_{i \in \Im} \int_{\Omega} c_{ij}(\omega)^{\rho} d\omega\right)^{\frac{1}{\rho}}$. Consequently, the demand for varieties of the composite good is given by $c_j(\omega) = p_j(\omega)^{-\sigma}\delta Y_j P_j^{\sigma-1}$, where $p_j(\omega)$ is the price of a variety ω, and P_j is the Dixit-Stiglitz price index, defined by

$$P_j \equiv \left(\sum_{i \in \Im} \int_{\Omega_i} p_{ij}(\omega)^{1-\sigma} d\omega\right)^{\frac{1}{1-\sigma}} \tag{2}$$

where Ω_j is the set of varieties available in location j. Then, total welfare in location j is $U_j = Y_j / P_j^{\delta}$, and the real wage can also be expressed as $w_j = W_j / P_j^{\delta}$, where W_j is the nominal wage. The total value of composite goods traded between origin i and destination j is $X_{ij} = p_j(\omega)c_j(\omega)$, which given the CES demand for varieties of the composite good, is equivalent to

$$X_{ij} = p_{ij}(\omega)^{1-\sigma}\delta Y_j P_j^{\sigma-1} \tag{3}$$

while the trade of the agricultural good is given by the demand for this good from the optimization problem, i.e. $F_j = (1-\delta)Y_j$.

3.3 Supply

On the supply side, it is assumed that the production of the homogeneous agricultural good is under perfect competition and constant returns to scale, with a Cobb-Douglas technology $F_j = A_j L_j^{1-\alpha}$, $\alpha \in 0,1$, in which the farmer produce one unit of F_j with one unit of labor L_j^F. For the case of the manufactured good, the price of a manufactured variety ω produced at origin location i and sell it in destination j, is

$$p_{ij}(\omega) = \frac{c_i}{z_i(\omega)}\tau_{ij} \tag{4}$$

where c_i is the unit cost measured in terms of labour, τ_{ij} are "iceberg" transport costs, and $z_i(\omega)$ are productivity independently draws from a Frechet distribution. The "iceberg" trade cost for the manufacturing varieties of selling each good from an origin i to a given destination location j, implies that $\tau_{ij} > 1$ units of the good must be send to reach one unit at the destination, $\tau_{ii} = 1 \;\forall\; i \in \Im$, and it is assumed that the no arbitrage condition holds, i.e. $\forall\; i,j,k \in \Im : \tau_{ij}\tau_{jk} \geq \tau_{ik}$. Additionally, consumers only purchase from the location with the lowest price, which implies that

$$p_j(\omega) \equiv \min_{i \in \Im} p_{ij}(\omega) = \min_{i \in \Im} \frac{c_i}{z_i(\omega)}\tau_{ij} \tag{5}$$

where the cumulative distribution that characterizes the productivity in each location is $F_{i(z)} \equiv \Pr\{z_i(\omega) \leq z\}$, which is assumed to be Fréchet, so $\forall\ z \geq 0$, $F_{i(z)} = exp\{-T_i z^{-\theta}\}$, where T_i is a measure of aggregate productivity in the location, and θ is assumed constant across locations. An intuitive interpretation of this function, is that the scale parameter of the Fréchet distribution T_i is measuring the local comparative advantage, while the dispersion parameter θ, measure the gains from trade on those goods. To the extent that θ is lower, i.e. a lower dispersion of the distribution, greater the gains from trading those goods. The idea is that any exogenous productivity shock in location i is captured by T_i, while transport infrastructure improvements would make the trading of those goods easier, which would imply a lower θ.

3.4 Equilibrium

The model is solved by following a probabilistic formulation. Where the probability that a location $i \in \Im$ offer to a location $j \in \Im$ a particular good $\omega \in \Omega$ for a particular price less than p, i.e. $\Pr\{p_{ij}(\omega) \leq p\}$, is

$$G_{ij}(p) \equiv 1 - \exp\{-T_i \left(\frac{c_i}{p}\tau_{ij}\right)^{-\theta}\}$$

Equivalently, the probability that a location $j \in \Im$ pays for a good $\omega \in \Omega$, a price less than p, i.e. $\Pr\{p_j(\omega) \leq p\}$, considering that consumers in location j only buys from the least cost location, is $G_j(p) = 1 - \exp\{-p^{\theta}\Phi_j\}$, where $\Phi_j \equiv \sum_{i\in\Im} T_i(c_i\tau_i)^{-\theta}$, which is the distribution of prices of each variety for a destination j. Then, considering these conditions, the equilibrium price index P_j in location $j \in \Im$ is

$$P_j = C\left(\sum_{i\in\Im} T_i(c_i\tau_{ij})^{-\theta}\right)^{-\frac{1}{\theta}} \iff P_j^{-\theta}C^{\theta} = \sum_{i\in\Im} T_i(c_i\tau_{ij})^{-\theta} = \Phi_j \tag{6}$$

where $C \equiv \Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)^{\frac{1}{1-\sigma}}$ and $\Gamma(t) \equiv \int_0^{\infty} x^{t-1}e^{-x}dx$. A remarkable property of this price index, is that is not only dependent on the bilateral trade between an origin i and destination location j, but also depends on all the other locations, as the sum over $i \in \Im$ indicates. This implies that the cost of those traded goods is not only dependent on the technology productivity draws T_i in each location, but is lower as the distance —or transport costs— to all other locations is low τ_{ij}, or equivalently, as the wages are low (given that perfect competition implies $w_i = c_i$). This useful feature will be important to derive an empirical measure of the market access variable that captures these multilateral trade frictions, as in Anderson and van Wincoop (2003).

3.5 Trade

Bilateral trade on the composite good between two pair of locations would be determined by the probability that a particular location $i \in \Im$ would be the least cost provider of a good ω to a destination

$j \in \Im$, which is equal to the fraction of goods i sells to j, i.e.,

$$\pi_{ij} = \frac{T_i(c_i \tau_{ij})^{-\theta}}{\Phi_j}$$

or equivalently, the proportion of goods that $j \in \Im$ buys from $i \in \Im$. The intuition behind this remarkable feature of this model, is that bilateral trade between an origin i and destination j would depend in a source of absolute advantage in technology T_i, and local comparative advantages given by marginal costs c_i and bilateral distance τ_{ij}, relative to all other locations measured in the denominator as $\Phi_j \equiv \sum_{i \in \Im} T_i(c_i \tau_i)^{-\theta}$. Where high values on these two sources of comparative advantages are penalized by θ. Due that θ is capturing the dispersion of the Fréchet distribution, it governs the role of comparative advantages, which imply that low penalties on θ (or equivalently low trade costs) would lead to a higher probability of trade between the pair of locations ij. Given the Fréchet distribution, this is also equal to the proportion of income that consumers in j spend on manufacturing goods from location i, specifically $\Lambda_{ij} \equiv x_{ij}/\delta Y_j$. In consequence, we can write the trade of manufacturing varieties that a destination location j spent in goods from an origin i as

$$X_{ij} = \pi_{ij} \delta E_j = \frac{T_i(c_i \tau_i)^{-\theta}}{\Phi_j} \delta E_j = C^{-\theta} \tau_{ij}^{-\theta} w_i^{-\theta} T_i P_j^{\theta} \delta E_j$$

where $\Phi_j \equiv \sum_{i \in \Im} T_i(c_i \tau_i)^{-\theta}$, and using the fact that these goods are produced under perfect competition, i.e. $c_i = w_i$.[13] This equation is also known as the Eaton-Kortum gravity equation, in this case for the manufacturing good.

3.5.1 Market Access

In general equilibrium, it would be the case that the value of goods purchased by a location i is equal to the income that consumers spent on that good, i.e., $\delta Y_j = \sum_{i \in \Im} X_{ij}$. In consequence and assuming that the income received at the sourcing location i has to be equal to the expenditure spent in destination location j, i.e., $Y_i = E_j$. The Eaton-Kortum gravity equation can be expressed as $P_j^{-\theta} = C^{-\theta} \sum_{i \in \Im} T_i w_i^{-\theta} \tau_{ij}^{-\theta} \equiv CMA_j$, which is defined as consumer market access, as in Donaldson and Hornbeck (2016). Substituting the consumer market access into the Eaton-Kortum gravity equation and assuming that goods markets clear, implies

$$\delta Y_i = C^{-\theta} T_i w_i^{-\theta} \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i \iff \delta Y_i = C^{-\theta} T_i w_i^{-\theta} FMA_j \tag{7}$$

where $FMA_j \equiv \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i$ is the firm market access in destination location $j \in \Im$. Under spatial equilibrium, workers perfectly mobile across locations equalize utility in space, which implies

[13] This is thanks to the convenient feature of the model, although strong, that $X_{ij}/X_j = \pi_{ij}$.

that real wages are also equalized across locations, i.e., $\overline{U} = \frac{w_i}{P_i} = \frac{w_j}{P_j}$.[14] In consequence, by substituting the nominal wage w_i in Eqn. 18, we obtain $\delta Y_i = C^{-\theta} T_i \overline{U}^{-\theta} CMA_j FMA_j$. Assuming symmetric transport $\tau_{ij} = \tau_{ji}$, and that firm and consumer market access satisfy a proportionality condition defined as $MA_i \equiv FMA_i = \rho CMA_i \ \forall i \in \Im$, with $\rho > 0$. Which implies that $FMA_i \equiv \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i \iff MA_i = \sum_{j \in \Im} \tau_{ij}^{-\theta} \rho MA_j^{-1} w_i N_i$, where $\delta Y_i = w_i N_i$. Substituting $w_i = \overline{U} P_i = \overline{U} \left(CMA_i\right)^{-\frac{1}{\theta}} = \overline{U} \rho^{\frac{1}{\theta}} MA_i^{-\frac{1}{\theta}}$, yields

$$MA_i = \sum_{j \in \Im} \tau_{ij}^{-\theta} \rho MA_j^{-1} \left(\overline{U} \rho^{\frac{1}{\theta}} MA_i^{-\frac{1}{\theta}} \right) N_i \iff MA_j = \overline{U} \rho^{\frac{1+\theta}{\theta}} \sum_{j \in \Im} \tau_{ij}^{-\theta} MA_i^{-\frac{(1+\theta)}{\theta}} N_i \qquad (8)$$

which is a convenient expression for the market access that can be directly approximately with available data.

3.6 Predictions

The model establishes a direct mapping between market access and local population and sectoral employment for each location. To yield more convenient expressions for this relationship, we assume labor market clearing $\delta Y_i = \delta w_i L_i$, together with the condition that $w_i = \overline{U} P_i = \overline{U} \left(CMA_i\right)^{-\frac{1}{\theta}}$, which allows us to formulate the following predictions to be taken to the data.

3.6.1 Market Access and Population/Employment

Market access increases local population. These effects are mediated by exogenous productivity shocks and changes in non-farm employment. Solving for the total number of workers (population) N_i and taking the logs gives

$$\log N_i = k_2 + \left(\frac{\theta}{1 - 2\theta} \right) \log T_i + \left(\frac{2 - \theta}{\theta} \right) \log MA_j + \left(\frac{\theta + 1}{2\theta - 1} \right) \log L_i^C \qquad (9)$$

where $k2 = \log \left(C^{\frac{\theta^2 - \theta - 1}{2\theta - 1}} \overline{U}^{\frac{\theta(\theta - 3)}{2\theta - 1}} \rho^{\frac{\theta(\theta - 1) + (2\theta - 1)(1 + \theta)}{\theta(2\theta - 1)}} \right)$. Note here that T_i, MA_j and L_i^C are interrelated by the parameter θ. Each coefficient latter estimated in regression would be a proportion of θ. The condition for a positive effect of the market access is that $\theta < 2$.

3.6.2 Market Access and Non-Farm Employment

Market access increases non-farm employment. Labor market clearing and spatial equilibrium implies

$$L_i^C = C^{-\theta} T_i \overline{U}^{-\theta} CMA_j FMA_j w_i^{-1} \iff L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} CMA_j^{\frac{\theta - 1}{\theta}} FMA_j$$

[14] Note that the assumption of spatial equilibrium is only used here at the end avoid the necessity to observe wages in the data. However, we also show results using nominal wages to see the robustness of this result to this assumption.

where $\delta L_i = L_C$ is the labor in the manufacturing sector. Using Eqn. 22 and taking the logs gives the equation of the manufacturing employment

$$\log L_i^C = k_1 + \log T_i + \left(\frac{2\theta - 1}{\theta}\right) \log MA_j \tag{10}$$

where $k_1 = \log\left(C^{-\theta}\overline{U}^{1-\theta}\rho^{\frac{1-\theta}{\theta}}\right)$. Here the condition for a positive effect of the market access is $\theta > 1/2$.

3.6.3 Market Access, Productivity Shocks, and Structural Transformation

Substituting Eqn. 22 in Eqn. 23 yields

$$L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \rho^{\frac{1-\theta}{\theta}} \left(\overline{U}\rho^{\frac{1+\theta}{\theta}} \sum_{j \in \Im} \tau_{ij}^{-\theta} MA_i^{-\frac{(1+\theta)}{\theta}} N_i\right)^{\frac{2\theta-1}{\theta}}$$

Then, with the proportionality condition $MA_i \equiv FMA_i = \rho CMA_i$ we can express Eqn. 21 as

$$L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \left(\rho^{-1} MA_j\right)^{\frac{\theta-1}{\theta}} MA_j \iff L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \rho^{\frac{1-\theta}{\theta}} MA_j^{\frac{2\theta-1}{\theta}} \tag{11}$$

Market access induces structural transformation but these changes are mediated by the exogenous productivity shocks. Which also can be expressed as employment share in manufacturing, as

$$\log\left(\frac{L_i^C}{N_i}\right) = k_3 + \left(\frac{\theta}{\theta+1}\right) \log T_i + \left(\frac{2(\theta-2)(\theta-\frac{1}{2})}{\theta(\theta+1)}\right) \log MA_j \tag{12}$$

where $k_3 = \log\left(C^{-\frac{\theta^2}{\theta+1}} \overline{U}^{\frac{-\theta^2-3\theta}{\theta+1}} \rho^{\frac{\theta(\theta-1)+(2\theta-1)(1+\theta)}{\theta(\theta+1)}}\right)$. As in the previous case, in this equation productivity shocks T_i and changes in market access MA_j, would have proportional effects on the structural transformation of each location, a complementarity that is mediated by transportation costs θ. At this point a θ in the interval (0.5, 2) is consistent with the idea of spatial development, in the sense that market access will lead to increases in population but also structural transformation. Therefore, identifying θ would be key in the empirical strategy.

4 Empirics

4.1 Data

4.1.1 Rural Communities/Villages

We use data from the Chilean censuses of population and housing of 1992, 2002 and 2017 (INE). The Chilean statistical office defines a rural community/village as any spatial entity with less than 3,000 inhabitants.[15] We select all those rural communities defined in 1992 with a population greater than 100 inhabitants for a total of 536 observations in each year.[16] This corresponds to approximately one third of the total number of rural communities in the census of 1992 (see Table A5.1 for summary statistics of the sample and Figure A5.2 shows the spatial distribution of rural communities and cities in the country). Overall, the changes across census years of the main variables are reported in Figure 2. Changes in population are mainly observed for the last census of 2017, as well as overall employment. Notwithstanding, non-farm employment in rural communities has been regularly increasing over the census years. Overall farm employment, on the other hand, has decreased between 1992 and 2002, but remains relatively stable between 2002 and 2017.

4.1.2 Roads

In order to properly compute the market access variable for each year, we use road network for each year of the census. The road network for Chile, however, is only available since 2011. Thus, we digitize information of the road network from previous years from the Ministry of Public Infrastructure (MOP), which are publicly available for years close to census periods (2003, 1999, 1986 and 1980). To have information for the specific census years, we complement this information using satellite images. To compute the least-cost route, we use the "fast-marching method", an algorithm introduced for the same purposes in Allen and Arkolakis (2014), and designed by Sethian (1996).

4.1.3 Urban Growth

For a proxy of the economic activity in cities, we use stable satellite nighttime light data, from the NASA Operational Line Scan Defense Meteorological Satellite Program (OLS-DMSP).[17] In the absence of economic data, the nighttime light satellite data was considered a good proxy of economic activity of the subnational units when used appropriately (Chen and Nordhaus, 2011; Henderson et al., 2012; Donaldson and Storeygard, 2016).[18] We process the nighttime light satellite images at one kilometer of spatial resolution for 1992, 2002, and 2017. For each year, we compute the sum of the nighttime light contained within the official urban boundaries defined for the census of 2002 and for areas considering buffers of 2km from the urban boundaries (for similar applications, see Binswanger-Mkhize and Savastano 2017; Henderson et al. 2017).[19] Fig. 3 compares the urban growth using nighttime lights

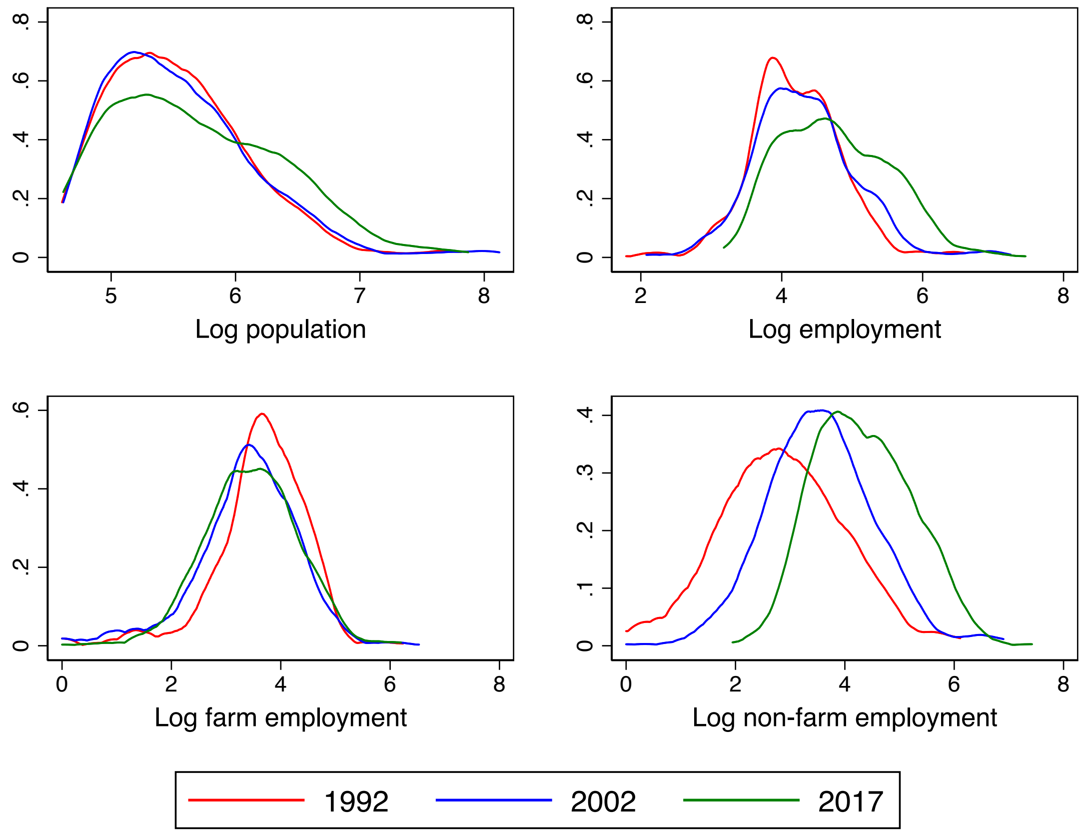
[15] This definition was established in 1992. Since 2002, a rural community is defined a spatial entity with less than 5,000 inhabitants and more than 35% of the labor force working in the agricultural sector. A rural community can be very heterogeneous, and can be a: *Town*, which is defined as a place between 1,000 and 5,000 inhabitants, or composed by more than 250 houses. It also can be an *Aldea*, which is a place between 300 and 1,000 inhabitants, or between 75 to 250 houses, without considering autonomous indigenous localities or agriculture communities. Also it can be a *Caserio*, which are places with less than 300 inhabitants. Within the category of places with less than 300 inhabitants, there can be mining campaments, a ranch, small human settlements withing agricultural land, or indigenous communities. These categories and the form of georeferencing these places is documented in INE (1995), INE (2005), and Carvajal et al. (2012).

[16] We also perform robustness checks of our estimations modifying this threshold.

[17] Chile, as many developing countries, does not provide data on the economic activity at city level.

[18] Some other applications are those of Beakley and Lin (2012); Michalopoulos and Papaioannou (2014); Storeygard (2016); Axbard (2016); Pinkovsky and Sala-Martin (2016); Henderson et al. (2017)

[19] We also perform robustness checks to buffers of 5, 10 and 15km from the urban boundaries.



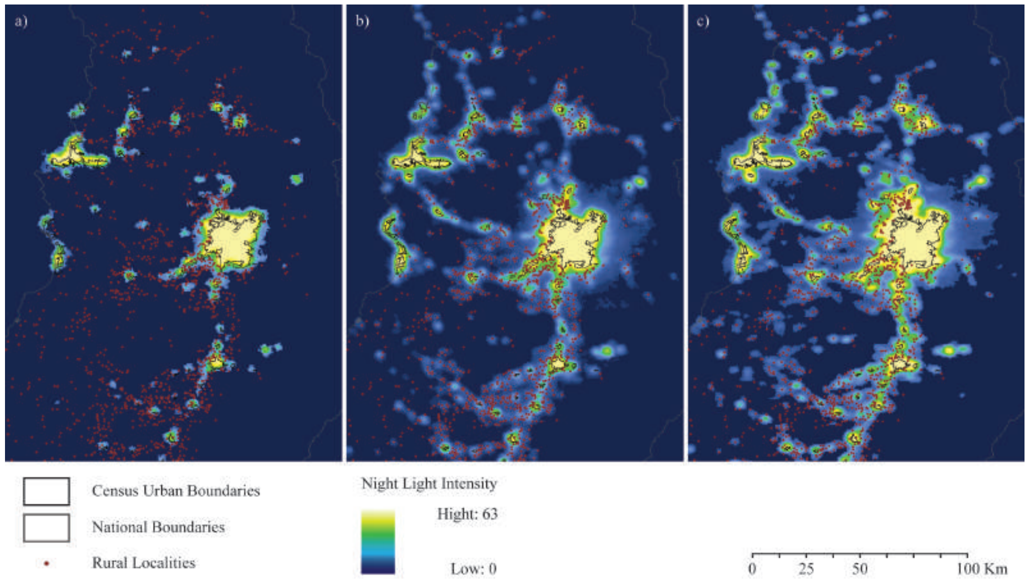
The figure shows the kernel (epanechnikov) density graphs for the log of population, employment, farm employment and non-farm employment in rural communities for 1992, 2002 and 2017.

Figure 2: Changes in Population and Employment in Rural Communities

between 1992, 2002 and 2013 (from left to right), near the metropolitan area of Santiago de Chile.[20] Rural communities are represented by red points. The black lines are the census urban boundaries at 2002, and nighttime light intensity is represented by a color scale from dark-blue (low values) to light yellow (high values).

The literature on remote sensing had identified two main problems related with the use of nighttime light data, namely the "blooming" and "saturation" effects (Imhoff et al., 1997). These effects are distortions in the satellite image due to the high intensity of light is some places. A simple analogy would be taking a photography with a source of light just in front of the camera. This would cause a saturation on the values of pixels from the direct source of light (saturation effect) and a light blurring to other pixels in the image (blooming effect). The main problem for our purposes is the blooming effect, due that overestimate the size of urban areas and, consequently, the sum of lights from cities. This is a problem that, until recent years, had been scarcely accounted for in the applied economics

[20] Time-series comparison on satellite nighttime lights only allows reliables comparisons between 1992 and 2013 after the calibrations that we implemented.



The figure describes the spatial distribution of rural communities (red points) near the metropolitan area of Santiago de Chile. Panel a) shows the image for 1992, panel b) for 2002 and panel c) for 2013. Urban boundaries (in black) are the official boundaries for the national census of 2002. The economic activity of urban areas was approximated for each year using nighttime light. Nighttime light intensity goes from zero (in dark-blue) to 63 (in light-yellow) and has a spatial resolution of one pixel per kilometer.

Figure 3: Urban Growth and Rural Communities

literature using nighttime light data (Abrahams et al., 2018). Hence, we clean nighttime lights images following remote sensing literature, specifically Su et al. (2015).[21] The basic idea of this cleaning process is that allow us to identify thresholds in the distribution of nighttime lights and extract built-up urban areas.

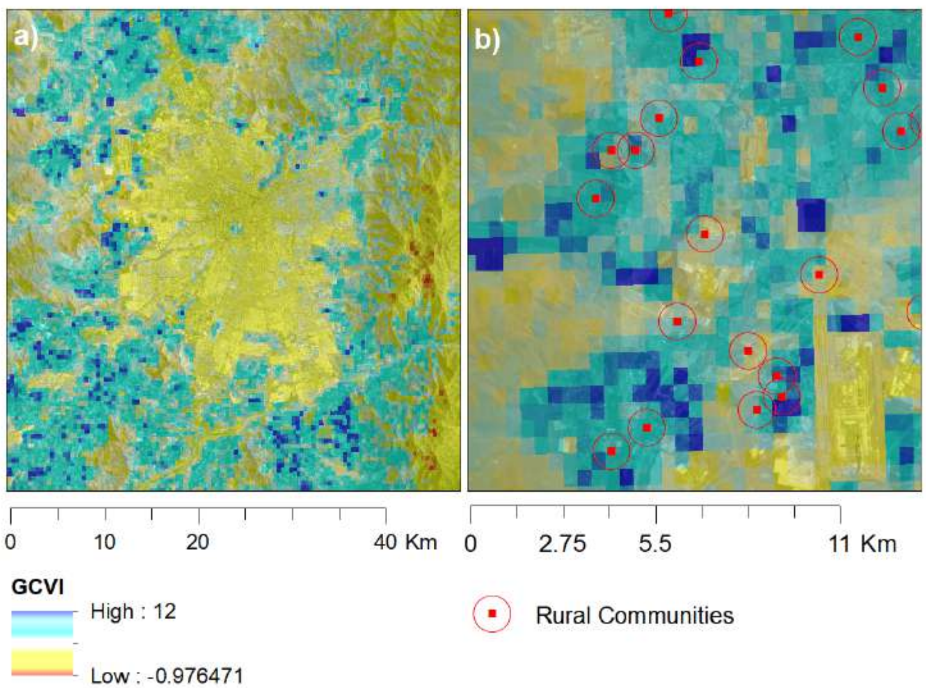
Fig. A5.6 describes the spatial distribution of rural communities near the metropolitan area of Santiago de Chile and the cleaning process of nighttime lights for 2002. Rural communities are represented by red points that correspond to the centroid of the area. The black lines represent the urban boundaries in 2002. The nighttime light information of each pixel is represented by a color scale from dark-blue (low values) to light yellow (high values). The maximum nighttime light intensity is 63 and the minimum nighttime light intensity is 0. The image on the left displays nighttime lights without blooming correction and the image on the right is corrected. White lines are used to represent the urban buffers of 2, 5, 10 and 15km from the urban boundaries. Since rural communities do not have boundaries delineating their areas, the sum of the nighttime light cannot be accurately computed for rural communities.[22] The distances between rural communities and cities were computed using the Euclidean distance between the centroids of rural communities and the centroids of cities.

[21] See Abrahams et al. (2018) for a more recent approach.

[22] We also extract the value of the pixel interpolated with surrounding pixels. This reaches an area of approximately two square kilometers (for a similar computation of interpolated values of satellite images, see Nunn and Puga, 2012). However, only 99 rural communities had nighttime light information in 1992 mainly due to the low access to electricity in rural communities during that census.

4.1.4 Agricultural Productivity

We follow recent studies that use remote sensing data to measure the agricultural potential of farms by computing vegetation indices (Costinot and Donaldson, 2012; Donaldson and Storeygard, 2016; Costinot et al., 2016; Costinot and Donaldson, 2016; Burke and Lobell, 2017). Since true agricultural productivity is unobserved, both survey-based methods and remote sensing indicators represent limited but informative approximations of the agricultural productivity of farms (Burke and Lobell, 2017). Specifically, we use NASA's Landsat-5 and Lansat-7 satellite data to compute the vegetation indices. We compute three different vegetation indices for robustness check, namely the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Green Chlorophyll Vegetation Index (GCVI), as proxies of the agricultural potential in rural villages.[23] Due to the satellite coverage of the NASA Landsat-5, these indices cannot be computed for the entire country for almost the entire duration of the 1990s.



The figure describes the agricultural potential near the metropolitan area of Santiago de Chile in 2015. Agricultural potential is represented with the GCVI, with a spatial resolution of 500 meters per pixel using data from the NASA Landsat-7. Panel a) shows the overall hinterland of the city of Santiago, where high values of the GCVI are displayed with more intense blue colors. Panel b) displays a zoomed image of the hinterland of the city, in which rural communities are represented by a red circle.

Figure 4: Cities and Agricultural Productivity

Consequently, we estimate the elasticities of market access for only 189 rural communities that have

[23]We describe in the Figure A8 the correlation between the vegetation index proxy of agricultural potential and the observed agricultural productivity using the agricultural census at 2007 at municipality level.

this information. The three vegetation indices differ in the spectral bands of the sensors of the satellite used for their computations, the GCVI being the most accurate to capture the agricultural productivity of farms (for a detailed discussion of applicability of these indices for similar purposes, see Burke and Lobell, 2017).[24] A cloud-free annual median of each pixel at 30 meters of spatial resolution was used for each year to compute the vegetation indices using Google Earth Engine (Gorelick et al., 2017). Subsequently, the resulting raster images of the vegetation indices averaged to a spatial resolution of 500 meters per pixel were processed in Google Earth Engine platform and interpolated with surrounding pixels to impute a value of agricultural productivity considering a representative area of one square kilometer. Fig. 4 shows the GCVI index for 2015 in the hinterland of the metropolitan area of Santiago de Chile.

4.2 From Theory to Estimation

The following equations came as a first-order approximations to the gravity model presented in the previous section.

4.2.1 Market Access

Prices and productivity in rural areas are rarely observed for a large number of rural locations, goods, and services. However, as specified in the theoretical section, we assume that market access captures those effects described in the model conditional on location specific fixed effects interacted with time dummies. Empirically, the influence of urban market access on the outcomes of rural areas is given by k cities influencing nearby rural communities through potential demand. Market access is a measure that captures this potential demand. The market access of a rural community c, located at a distance $d_{c,k}$ from a city with C_k income, is given by

$$MA_{ct} = \sum_k C_{kt} e^{-\theta d_{ck}} \tag{13}$$

where θ is a parameter that defines the shape of the distance discount over C_{kt}. This specific empirical approximation to the market access can be found in Hanson (2005). This imposes a negative exponential functional form structure to the variable.[25] The market access variable, is computed using stable satellite nighttime lights to approximate the values of C_{kt}, and the road distance between the centroids of rural communities and cities were used to compute d_{ck}. Since we are interested in estimating the impact of cities on the growth and development on rural areas, our relevant measure is the access to urban markets for each rural community/village.

[24]The NDVI is computed as $NDVI = (NIR - Red)/(NIR + Red)$, EVI is computed as $EVI = 2.5 * (NIR - red)/(NIR + 6 * Red + 7 * Blue + 1)$, and GCVI as $GCVI = (NIR/Green) - 1$, where NIR is the Near Infrared Band and each color represents a different wavelength band of the satellite sensors.

[25]Which implies a faster decay rate of nighttime lights with distance in comparison to the polynomial functional form frequently used (Harris, 1954). This decay rate is more intensified when the distance variable is to the power of 2.

Additionally, we are aware that the values of θ could affect the estimated elasticity of market access. Therefore, we follow a different approach to observe how the elasticity of market access can be affected by different values of θ, since we do not have precise information on transportation costs and flows of goods between rural communities and cities to an appropriate estimation of θ. Consequently, we operate under three different scenarios to check the robustness of our estimations. The estimation of θ is important because this parameter summarize trade relations that are implicit in the market access equation (Donaldson and Hornbeck, 2016; Donaldson, 2018). And our different scenarios give an insight on how much of the variation in θ can influence the market access elasticities of population and employment.[26]

Under the first scenario, we compute market access following an approximation of Harris (1954). This is the baseline scenario, as the case in which $\theta = 1$ and the square of $d_{j,k}$ is used, which is also the most widely used empirical approximation for market access. For the second scenario, we rely on the literature on planning and transportation, based on the early works of Carrothers (1956), Hansen (1959), and Weibull (1976) and recently used for similar applications in agricultural economics by Binswanger-Mkhize et al. (2016) and Binswanger-Mkhize and Savastano (2017). In this scenario, we use a standard negative exponential distance decay function, in which $\theta = 1/2a^2$ and the square of $d_{j,k}$ is used, where a is a parameter that represents the distance to the point of inflection of the distance decay function. We estimate it by using a representative value for $a = 50$ and $a = 100$, as in Binswanger-Mkhize et al. (2016).

Therefore, in addition we estimate parameter θ of the market access function adjusted by using three different distance decay functions according to the size of cities (a similar approach has been explored by Halas et al., 2014).[27] This allows us to account for the fact that the spatial distribution of medium- and small-sized cities is more spread and easier to access by the population of rural communities, together with the fact that large cities are usually more concentrated in some areas and have less spatial scope over the entire spatial distribution of rural communities.[28]

4.2.2 Market Access and Rural Population and Employment Growth

From Eqn. 9, we can derived an empirical approximation as:

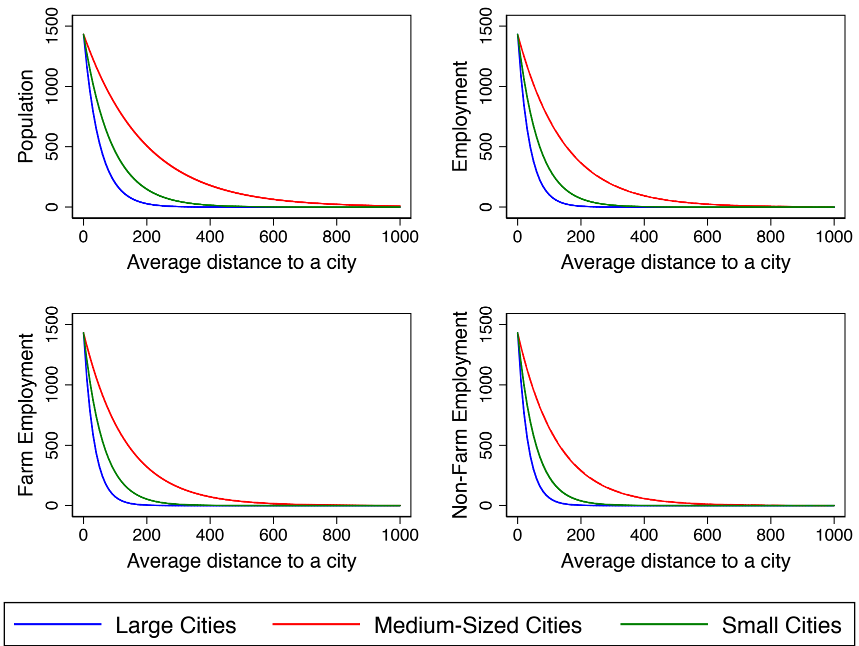
$$\ln Y_{ct} = \beta \ln MA \left(\mathbf{C}_k, \mathbf{d}_{ck}, \hat{\theta}\right)_{ct} + \delta_m + \delta_t + \delta_{mt} + f(x_c, y_c) + \epsilon_{ct} \tag{14}$$

where Y_{ct} = is population, employment (farm and non-farm) in rural community j in year t, MA_{ct} = access to urban markets of the rural community j in year t, $MA_{ct} = \sum_k C_{kt} e^{-\theta d_{c,k}}$, δ_m = municipality

[26]Usually, this parameter is estimated using trade flows, but in this case we assume that the structure of the trade relationship depends on the outcome in the destination locality (Donaldson, 2018). Therefore, the parameter that describe contemporaneous population is the similar to the parameter that describe flows (Donaldson and Hornbeck, 2016).

[27]This mimic the approach of computing different θ for different transport modes with the distinction that here we assume that transport infrastructure might vary depending on the size of cities. Considering that usually better infrastructure is founded to access to larger cities.

[28]This stylized fact can be inferred from Fig. A5.2 and Fig. A5.8, and confirmed in Fig. 5 for the case of study.



The figure shows the market access predictions for population, total employment, farm and non-farm employment by city-size at 1992 using non-linear least squares. The estimated parameters of these distance decay functions were used to construct the city-size adjusted market access variable. In all cases, there are larger effects of medium-sized cities (the red line) over rural population and employment. These effects are because the aggregate wealth of medium-sized cities is larger and they are more equally distributed in space, having a larger spatial scope of their positive effects, compared to big or small-sized cities.

Figure 5: Market Access Functions Adjusted by City-Size

fixed effect, δ_t = year fixed effect, δ_{mt} = municipality-year fixed effect, $f(x_c, y_c)$ = cubic polynomial in latitude and longitude of the rural community. Standard errors are clustered at the rural community level. Due that the effects of urban growth on the population and employment of rural communities would take time to manifest, we estimates these equations using a one-year lagged market access variable. There are demographic and cultural factors such as social norms, as well as other externalities affecting local population. If these are specific to municipalities, then the time by municipality fixed effects can capture those effects.

According to our framework, the effects of market access on farm and non-farm employment may be associated with two main mechanisms: the demand for agricultural goods and changes in agricultural productivity. Both of these mechanisms are captured by market access. On one hand, a higher access to urban markets is associated with a higher demand for agricultural products and, consequently, a higher employment in the farm sector and also in the non-farm sector but in activities auxiliary to agriculture, such as the commerce and transportation of agricultural goods. On the other hand, rural communities with better connectivity to cities may have more access to technology for agricultural production. This would increase the agricultural productivity of these rural communities, diminishing the price of agricultural goods in relation to manufactured goods, leading to a movement of workers from the farm to the non-farm sector. However, even in this scenario, is likely that some non-farm jobs would be in activities that are auxiliary to agriculture. Location specific fixed effects interacted with year dummies would take variation caused by local differences in technology and prices and left the variation that is induced due to economic growth in cities and improvements in access to urban markets.

The effects of market access on non-farm employment of rural communities may also impact activities not related to agricultural production. For example, if the opportunity cost of producing agricultural goods is higher in the hinterland of cities due to the reallocation of some urban activities such as manufacturing production, this would lead to an increase in the employment of rural communities in the non-farm sector in activities not necessarily related to agricultural production. However, this effect is likely to be found only in the hinterland of cities. This implies that if prices and productivity of the non-tradable sector increases more than prices and productivity in the farm sector, which is in theory expected as higher the market access, then the impact in the farm employment would be negative.[29] However, in the case of the non-farm employment, the effect of the size of the non-tradable sector has a major weight masking negative effects in the tradable sector.

From our setting and assuming that workers are indifferent to working in either sector and have free mobility, the farm employment of a rural community j depend on the way farm and non-farm workers substitute agricultural and manufactured goods. If the demand for agricultural goods is inelastic (as suggested by Tobin, 1950; Tolley et al., 1969; Van-Driel et al., 1997), any improvement in agricultural productivity could induce a movement of farm workers to the non-farm sector since an increase in agricultural productivity decreases the price of the agricultural good. However, since a proportion of non-farm employment is generated in rural communities, this migration from the farm to the non-farm sector does not always imply a reduction in the total employment level of the rural community. This allows for the possibility that the diversification of the employment in rural communities could also generate growth in total population and offset the decline in the farm employment. This also is consistent with the literature of agricultural intensification and can allow rural growth with structural change, or a decline in rural growth if the increase in the non-farm sector do not offset the decrease in the farm-employment.

The population of rural communities, is an important indicator of their development and also determines total employment (Carlino and Mills, 1987; Henry et al., 1999; Partridge and Rickman, 2003; Hoogstra et al., 2011; Chen and Partridge, 2013; Olfert et al., 2014), considering both the number of workers in the farm and non-farm sector. Despite the migration effects of workers to cities, on average, this may be offset by the benefits of higher market access. From our setting, the expected effect of market access on the non-farm employment of rural communities is positive and if this offset any

[29] Empirically we do not observe prices and productivity for each rural community and time, however it is an, arguably, plausibly assumption that productivity and prices in the agricultural sector are lower than in the non-tradable and tradable non-agricultural sector.

decline in the farm employment in these communities, a positive effect would be expected for the population of these communities. An increase in the non-farm employment, however, may also induce a decrease in the farm employment. Notwithstanding, this only would affect negatively the average population of rural communities if this migration is to urban areas.

However, only rural communities that are very close to cities are likely to be affected by both of these factors. Following a similar approach as Donaldson and Hornbeck (2016), we account for this issue by constructing buffers/polygons of 10 and 15 kilometers from the boundaries of cities and select those rural communities outside these areas. Farther away from these thresholds, the average effect of market access on the population and farm and non-farm employment of rural communities is expected to be positive. Moreover, we also estimate the results with and without small rural communities to observe the robustness of our results to different cut-off points for the inhabitants of rural communities.

Additionally, we instrument the market access using the historical census of population in cities in $t = 1865$, 1889, and 1900 $\left(G_{j,t} = \sum_k Pop_{k,t} e^{-d_{j,k}}\right)$.[30] Most of places where rural communities are nowadays where inhabited in those early years, just fifty years after the Chilean independence, making the instrument relevant. In the sense that is a good predictor of the location and growth of rural communities only because it explains the location and growth of cities. Identification comes from historical population influencing contemporaneous population in cities.

4.2.3 Market Access, Agricultural Potential and Farm Employment Growth

It might be the case that the effects of the market access depends on the heterogeneity in specialization on agricultural activities in rural communities. This is the hypothesis that market access have a positive effect on agricultural intensification. It is important to understand this because in areas that are more specialized in agricultural activities, it would be more convenient to invest on the development of the agricultural sector, but in areas that are more diversified it might be better to support non-farm activities. Consequently, we estimate the heterogeneous effects of market access across rural communities with different levels of agricultural potential. This would be given by

$$\ln L^F_{ct} = \alpha A^F_{ct} \ln MA\left(\mathbf{C}_k, \mathbf{d}_{ck}, \hat{\theta}\right)_{ct} + \delta_m + \delta_t + \delta_{mt} + \delta_l + f(x_c, y_c) + \epsilon_{ct} \tag{15}$$

with L^F_{ct} = farm employment in rural community c in year t, A^F_{ct} = agricultural potential in rural community c in year t, Normalized Difference Vegetation Index (**NDVI**), Enhanced Vegetation Index (**EVI**), Green Chlorophyll Vegetation Index (**GCVI**), δ_c = municipality fixed effect, δ_t = year fixed effect, δ_{mt} = municipality-year fixed effect, δ_l = land-cover fixed effect, $f(x_c, y_c)$ = cubic polynomial in latitude and longitude of the rural community, Identification comes from: exogenous geography influencing agricultural productivity. Standard errors clustered at the rural community level. Therefore, we also estimate the elasticities of the market access incorporating in the farm employment equation

[30] This approximate of the market access in 1865 is expected to only have an effect on contemporaneous population and employment growth in rural communities through contemporaneous market access. The urban system at 1865 was composed of 30 cities, while at 1992 it was composed of 160 cities.

the agricultural potential proxy with the vegetation index. The square of the agricultural potential is also incorporated in this set of equations, following the evidence in Damania et al. (2017), that suggest non-linear effects of agricultural potential on production, due to increasing or diminishing returns of crop suitability that different crops on modern and traditional production might have.

4.2.4 Farm to Non-Farm Employment Multipliers

The growth in the agricultural sector, still can be one of the main drivers for rural economic development in some rural areas. And, given that market access have a positive effect for rural communities that have more agricultural potential, then it is useful to understand better to what extent farm employment generate employment in the non-farm sector, by activities that can be auxiliary to agriculture or that are associated with the intensification of agricultural activities in those areas.

$$\ln L_{ct}^{NF} = \gamma L_{ct}^{F} + \delta_m + \delta_t + \delta_{mt} + f(x_c, y_c) + \epsilon_{ct} \tag{16}$$

L_{ct}^{NF} = non-farm employment in rural community c in year t, L_{ct}^{F} = farm employment in rural community c in year t, δ_m = municipality fixed effect, δ_t = year fixed effect, δ_{mt} = municipality-year fixed effect, $f(x_c, y_c)$ = cubic polynomial in latitude and longitude of the rural community. Identification comes from exogenous geography influencing agricultural productivity, standard errors clustered at the rural community level.

4.3 Evidence

4.3.1 General Effects

Table A5.3 reports the results of the impact of market access on the log of population, farm employment, and non-farm employment of rural communities. The table describes the different scenarios used to compute the market access variable. Panel a) shows the baseline case, which assumes that the parameter $\theta = 1$ (i.e., the market access variable used in these estimations), which is approximately similar to that of Harris (1954). We choose this as the baseline scenario because is the most widely used measure of market access. All regressions include fixed effects at the rural community level and standard errors are clustered at the municipality level.[31] The first three columns of Table A5.3 show the Ordinary Least Squares (OLS) estimations of the impact of market access on the log of population, farm employment, and non-farm employment of rural communities. The last three columns of the table present the Instrumental Variable (IV) estimations to account for the endogeneity of the market access variable. The observations in this set of regressions are 1,451, that correspond to rural communities with a population between 100 and 3,000 inhabitants at 1992. This rural villages have an average population of 500 inhabitants (see Table A5.1). Across all estimations elasticities might be particularly high, and this is expected due to the small size of the rural communities, but also an important

[31]Each municipality has several rural communities or villages. There are 346 municipalities in Chile, but approximately 90 have rural communities. Most of the municipalities at the extreme north and south of the country do not have agro-ecological conditions favourable for agricultural production.

degree of heterogeneity is expected, as some rural communities are growing while other are loosing population as a function of the market access and the importance of the agricultural sector.

Across all IV estimates in Table A5.3, conditional on local demographic trends captured by locality fixed effects, the impact of the market access on the population of rural communities is positive and statistically significant. Although the significance is not too strong, that might reveal the low gains in average due that the heterogeneity that is underlying the estimates and that the overall population in rural communities just has an important increase during the last 15 years, as it is shown in 2.

The IV estimated 25-year elasticity of market access on population in Panel a), suggests that a 10% more market access led to an increase of approximately 3% in the population of rural communities, while this effect is between 1% and 4% for the low- to high-distance friction scenarios of the distance decay market access function (Panel b and c), and approximately 3% for the city-size adjusted market access function (Panel d). The assumption that rural communities face similar distance frictions to get to large and small size cities is difficult to maintain because infrastructure is usually better for accessing large cities. Therefore, Panel d) present the results using a market access variable calculated to allow rural communities to face different distance frictions to reach different types of cities. Consequently, in this scenario, we do not assume a particular value of θ for the computation of the market access variable, but use different values of θ to adjust the market access variable for different distance decay functions (these functions are presented in Figure 5). Given the small population size of rural communities, these are not too important effects, due that double the population of the largest rural community according to the census definition at 1992 just would mean to reach almost 6,000 inhabitants. However, the effects on population tend to be robust across the different market access functions.

The impact of market access on the change of non-farm employment of rural communities is also positive and stronger than the effect on the rural population, with particularly high elasticities that tend to vary in an important magnitude between the different market access functions. Specifically, for the baseline scenario in Panel a), the IV estimate is of nearly 1, however, these elasticities vary between 0.3 and 1.2 between the low- and high- distance frictions scenarios of the distance decay market access function. The city-size adjusted market access function, however, gives a more intermediate elasticity of approximately 0.7. In consequence, it is more likely that non-farm employment effects tend to be more heterogeneous and depend more on the accessibility than the population that is probably more stable. It might be the case that many of these jobs require of certain low commuting time to urban center to be created, and then estimations are more sensible to the selection of the parameter of distance friction.

On the other hand the market access elasticity on farm employment is non-significant in all the periods, both in OLS and IV estimations. This higher elasticity of market access for non-farm employment relative to farm employment is robust across all estimations and is likely to be associated with the process of structural change in the rural economy (Irwin et al., 2009; Castle et al., 2011). The important differences between the OLS and IV estimates are likely to be explained because these effects tend to

Table 1: Market Access, Population and Agricultural and Non-Agricultural Employment

	OLS			2SLS		
	Log Population	Log Agricultural Employment	Log Non-Agricultural Employment	Log Population	Log Agricultural Employment	Log Non-Agricultural Employment
Panel a): Baseline Market Access with $MA_c = \sum_k C_k e^{-d_{c,k}}$						
Log Market Access	0.105	0.093	0.178*	0.316**	-0.134	0.951***
	(0.069)	(0.083)	(0.102)	(0.163)	(0.212)	(0.254)
R^2	0.470	0.528	0.646			
F-stat				114.561	114.485	122.253
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel b): Distance Decay Market Access Function with $MA_c = \sum_k C_k e^{-\frac{1}{2(50)^2} d_{c,k}}$						
Log Market Access	0.021	0.012	0.030	0.097*	-0.042	0.301***
	(0.018)	(0.020)	(0.025)	(0.163)	(0.066)	(0.086)
R^2	0.470	0.527	0.646			
F-stat				50.103	47.874	50.195
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel c): Distance Decay Market Access Function with $MA_c = \sum_k C_k e^{-\frac{1}{2(200)^2} d_{c,k}}$						
Log Market Access	0.107	0.064	0.151	0.396*	-0.170	1.222***
	(0.078)	(0.084)	(0.105)	(0.205)	(0.271)	(0.343)
R^2	0.471	0.528	0.646			
F-stat				53.674	51.885	54.801
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel d): City-Size Adjusted Market Access Function with $MA_c = \sum_k Y_k e^{-\hat{\theta}_m d_{j,k}}$						
Log Market Access	0.153*	0.095	0.167**	0.326*	-0.099	0.693***
	(0.080)	(0.066)	(0.082)	(0.167)	(0.156)	(0.183)
R^2	0.473	0.529	0.647			
F-stat				120.152	136.927	147.121
Observations	1,451	1,451	1,451	1,451	1,451	1,451

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument historical urban market access.

be more local. The instruments are taking the variation of the market access variable that is being explained by historical population in cities and then explaining the contemporary patterns of rural population and employment.

The elasticity of market access for the population is similar to that in the literature (Jedwab and Storeygard, 2022). However, the elasticity is particularly high for farm and non-farm employment, but not much other empirical evidence exists to compare with these results. However, the distinction between farm and non-farm employment in rural communities is crucial. When no distinction is made and the elasticities of market access are computed for the total employment of the rural areas, the results may not be well understood, in the sense that, in many locations, particularly near cities, the behavior of farm employment and non-farm employment in rural communities may follow different patterns, leading to negative effects of market access on farm employment. This could also be the case in some articles that argue a negative or non-significant effect of market access on the total employment of rural areas but a positive and significant effect on the population, as in Chen and Partridge (2013) for China. The high values of the elasticities of market access for the non-farm employment of rural communities in relationship to the elasticities for the farm employment of rural communities are also indicative of the growing importance of the non-farm sector for rural areas, something that has been widely studied in agricultural economics (Berdegué et al., 2001). However, there is an important heterogeneity in rural communities, evidenced when our proxy of agricultural productivity was included in estimations.

A movement of workers from agriculture to the non-farm sector in rural communities may explain the lack of a significant effect of the market access on farm employment, and the high elasticities in the non-farm employment. If this movement, on average, was inside rural communities would also explain a consistently positive market access effect in the population of rural. Otherwise, we should observe a negative effect on rural population.

4.3.2 Mechanisms of Structural Change

One of the key elements of heterogeneous impacts of the market access on rural population and employment is agricultural productivity. Consequently, in order to provide more evidence on the mechanisms behind the growth of rural communities, Table A5.4 shows the regressions including our proxy of agricultural potential in the farm-employment equation. For this purpose, we use remote sensing data to construct a proxy of agricultural productivity, which as the evidence suggests, might be as informative as survey-based methods (Burke and Lobell, 2017). We estimate this relationship using three indices. Namely, the NDVI, EVI, and GCVI. The different market access variables are used to show the robustness of the results. Due to the availability of daytime cloud-free satellite data for the entire country, the sample is reduced. In Table A5.4, we also allow the market access vary with the agricultural potential to observe if market access have a positive effect in farm employment in rural communities that have better conditions for agricultural production. This would be more in line with the hypothesis of agricultural intensification. To do this we interact the market access with the agricultural potential.

Table 2: Market Access, Agricultural Potential and Agricultural Employment

	Log Agricultural Employment					
	OLS			2SLS		
	GCVI	NDVI	EVI	GCVI	NDVI	EVI
$A^F_{jt} \ln MA$	-0.034**	-0.033**	-0.028**	-0.492***	-0.365***	-0.296***
	(0.015)	(0.014)	(0.012)	(0.167)	(0.112)	(0.088)
A^F_{jt}	0.169***	0.066	0.082	0.876***	0.254**	0.183**
	(0.066)	(0.065)	(0.063)	(0.275)	(0.100)	(0.078)
R^2	0.562	0.547	0.547			
Observations	1,021	1,008	1,010	1,021	1,008	1,010
F-stat				15.675	20.379	23.066
P-val LM stat				0.000	0.000	0.000

$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument historical urban market access.

Our estimations show significant negative elasticities of the market access interacting the agricultural potential farm employment of rural communities and are robust across the different proxies of agricultural productivity. Therefore, is likely that rural communities highly specialized in agriculture, or at least with good agro-ecological conditions for the production of agriculture (probably captured by our vegetation indices), experiences decreases in the farm employment, that to some extents can be explained by technological productivity improvements, or diversification. Taking large advantages of the access to urban markets with also a growth in the non-farm sector, probably in activities that are auxiliary to agriculture, such as the transportation of agricultural goods or commerce, rather than in activities that are not related to agriculture. This is consistent with the hypothesis of the intensification of agricultural activities in places with more market access. The elasticity of agricultural labor supply to market access is an important indicator on how the impact of increasing demand generated in urban areas would impact rural inhabitants Jayachandran (2006).

To understand to what extent the growth in non-farm employment is related with agricultural production due to the requirements of activities that are auxiliary to agriculture, or might be explained by other activities that are not related even indirectly with agriculture. Table A5.5 shows the multipliers effects of farm employment on non-farm jobs. The IV estimate reveal an employment multiplier of approximately 0.9. This is a high elasticity that shows that it might be the case that the growth in the agricultural sector might induce growth in non-farm sector in activities that are not directly related. Then for some rural communities with better conditions for agriculture or more specialized on it, public investments in the sector might have a high social return.

Table 3: Agricultural to Non-Agricultural Employment Multipliers

	Log Non-Agricultural Employment	
	OLS	2SLS
Log Agricultural Employment	0.330***	0.865***
	(0.050)	(0.303)
R^2	0.672	
Observations	1,451	1,451
F-stat		17.491
P-val LM stat		0.000

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument agricultural potential captured with satellite images.

5 Conclusion and Policy Implications

In an increasingly urbanized world, the idea of cities as engines of economic growth and development has gained a large number of supporters. Even for rural development policies, it is argue that the positive effects derived from urbanization that spill over to the countryside such as trade integration or the growth of non-farm activities, are a potential source to outweigh the economic decline that is induced at the same time by rural-to-urban migration. In fact, far from being new, the idea that cities can shape the development of rural areas has been on the mind of scholars for a long time. It is not entirely clear, however, to what extent the positive spillover effects would offset the negative ones.

To provide evidence on the dynamics of urbanization for rural areas, we follow a market access approach to develop a framework to better understand how urban growth and transport infrastructure development influence the growth and structural change out of agriculture of rural communities through market access. Consequently, we estimate the effect of cities on the development of rural communities as the effect of the market access on the population, farm and non-farm employment of rural communities in Chile for a 25-year period.

The results can be summarized by the following four main points. First, we found, in our preferred estimations, that a 10% higher market access led to an average 25-year increase of nearly 3% in the population of rural communities in Chile. This elasticity is robust across all estimations. Second, a high access to urban markets is in the short-term generally associated with increasing both farm employment and non-farm employment in rural communities. Third, higher elasticities of market access were found for non-farm employment rather than in the farm sector. However, rural communities more specialized in agriculture might experience higher levels of growth induced by cities in the farm employment rather than in the non-farm sector.

Various policy implications can be derived from this work. First, improvements in the infrastructure connecting rural communities and cities could lead to important effects in the growth and development of rural communities, thus accelerating the process of structural change for those rural communities nearby cities. Second, these improvements in infrastructure can be prioritized for rural communities that are near medium-sized cities, since they can reach a larger number of rural communities. Third, for rural communities that are farther away from cities, policies can be oriented to develop both the farm and non-farm sector, but should carefully consider the conditions for agricultural production in those areas. Finally, for rural communities highly specialized in agriculture, rural development policies may be created to provide farmers with a better comprehension of the products that are highly demanded in cities, together with increased access to farming technology. These policies aiming rural-urban linkages are at the hearth of the objectives of the UN sustanaible development goals to promote rural development.

More research could be conducted to better understand the dynamics of rural communities near the boundaries of cities. The effects associated with the changes in land use, rural non-farm employment that is generated in the city, and other commuting opportunities for rural inhabitants could be leading to different patterns of development for those rural communities located farther away from cities. Despite that, on average, the effect of cities is positive and significant for population, the dynamics of nearby communities could be playing an important role and there is limited evidence available in this respect. On the other change, analyzing the distributional impacts on welfare from urban growth on rural households would help to clarify many of the mechanism that are included in the model and were not treated in this paper due to data limitations.

References

ABRAHAMS, A., ORAM, C., AND LOZANO-GRACIA, N. 2018. Deblurring DMSP Nighttime Lights: A New Method Using Gaussian Filters and Frequencies of Illumination. *Remote Sensing of Environment* 210:242–258.

ALLEN, T. AND ARKOLAKIS, C. 2014. Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics* 129:1085–1139.

ALVAREZ-CUADRADO, F. AND POSCHKE, M. 2011. Structural Change Out of Agriculture: Labour Push versus Labour Pull. *American Economic Journal: Macroeconomics* 3:127–158.

ANDERSON, J. E. AND VAN WINCOOP, E. 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93:170–192.

ASHER, S. AND NOVOSAD, P. 2020. Rural Roads and Local Economic Development. *American Economic Review* 110:797–823.

AXBARD, S. 2016. Income Opportunities and Sea Piracy in Indonesia: Evidence from Satellite Data. *American Economic Journal: Applied Economics* 8:154–194.

BAIROCH, P. 1988. Cities and Economic Development. The University of Chicago Press, Chicago.

BANERJEE, B. 1984. The Probability, Size, and Uses of Remittances from Urban to Rural Areas in India. *Journal of Development Economics* 16:293–311.

BARKLEY, D., HENRY, M., AND BAO, S. 1996. Identifying "Spread" Versus "Backwash" Effects in Regional Economic Areas: A Density Function Approach. *Land Economics* 72:336–357.

BARWICK, P. J., DONALDSON, D., LI, S., LIN, Y., AND RAO, D. 2021. Improved Transportation Networks Facilitate Adaptation to Pollution and Temperature Extremes.

BEAKLEY, H. AND LIN, J. 2012. Portage and Path Dependence. *Quarterly Journal of Economics* 127:587–644.

BECKMANN, M. 1972. Von Thunen Revisited: A Neoclassical Land Use Model. *The Swedish Journal of Economics* 74:1–7.

BERDEGUÉ, J., PROCTOR, F., AND CAZZUFFI, C. 2014. Inclusive Rural-Urban Linkages. *Working Paper Series of the Latin American Center for Rural Development* 123.

BERDEGUÉ, J., RAMIREZ, E., REARDON, T., AND ESCOBAR, G. 2001. Rural Non-Farm Employment and Incomes in Chile. *World Development* 29:441–425.

BINSWANGER-MKHIZE, H. AND SAVASTANO, S. 2017. Agricultural Intensification: The Status in Six African Countries. *Food Policy* 67:26–40.

Binswanger-Mkhize, H. P., Johnson, T., Samboko, P., and You, L. 2016. The Impact of Urban Growth on Agricultural and Rural Non-Farm Growth in Kenya.

Briant, A., Combes, P., and Lafourcade, M. 2010. Dots to Boxes: Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations? *Journal of Urban Economics* 67:287–302.

Burchfield, M., Overman, H., Puga, D., and Turner, M. 2006. Causes of Sprawl: A Portrait from Space. *Quarterly Journal of Economics* 121:587–633.

Burke, M. and Lobell, D. B. 2017. Satellite-Based Assessment of Yield Variation and its Determinants in Smallholder African Systems. *Proceedings of the National Academy of Sciences* 114:2189–2194.

Carlino, G. and Mills, E. 1987. The Determinants of County Growth. *Journal of Regional Science* 27:39–54.

Carrothers, G. 1956. An Historical Review of the Gravity and Potential Concepts of Human Interaction. *Journal of the American Institute of Planners* 22:94–102.

Carruthers, J. and Vias, A. 2005. Urban, Suburban, and Exurban Sprawls in the Rocky Mountain West: Evidence from Regional Adjustment Models. *Journal of Regional Science* 45:21–48.

Carvajal, L., Poch, M., and Osorio, R. 2012. Estudio de Identificación de Localidades en Condiciones de Aislamiento 2012. Subsecretaría de Desarrollo Regional y Administrativo (SUBDERE), Santiago, Chile.

Castle, E., Wu, J., and Weber, B. 2011. Place Orientation and Rural-Urban Interdependence. *Applied Economic Perspectives and Policy* 33:179–204.

Chen, A. and Partridge, M. 2013. When are Cities Engines of Growth in China? Spread and Backwash Effects Across the Urban Hierarchy. *Regional Studies* 47:1313–1331.

Chen, X. and Nordhaus, W. 2011. Using Luminosity Data as a Proxy for Economics Statistics. *Proceedings of the National Academy of Sciences* 108:8589–8594.

Christiaensen, L., Weerdt, J. D., and Todo, Y. 2013. Urbanization and Poverty Reduction: The Role of Rural Diversification and Secondary Towns. *Agricultural Economics* 44:435–447.

Colby, C. 1933. Centrifugal and Centripetal Forces in Urban Geography. *Annals of the Association of American Geographers* 23.

Costinot, A. and Donaldson, D. 2012. Ricardo's Theory of Comparative Advantage: Old Idea, New Evidence. *American Economic Review: Papers & Proceedings* 102:453–548.

Costinot, A. and Donaldson, D. 2016. How Large Are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-1997. Working Paper 22946, National Bureau of Economic Research.

COSTINOT, A., DONALDSON, D., AND SMITH, C. 2016. Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields Around the World. *Journal of Political Economy* 124:205–248.

DAMANIA, R., BERG, C., RUSS, J., BARRA, A., NASH, J., AND ALI, R. 2017. Agricultural Technology Choice and Transport. *American Journal of Agricultural Economics* 99:265–284.

DAVIS, B., GIUSEPPE, S., AND ZEZZA, A. 2017. Are African Households (Not) Leaving Agriculture? Patterns of Households' Income Sources in Rural Sub-Saharan Africa. *Food Policy* 67:153–174.

DELLER, S., TSAI, T., MARCOUILLER, D., AND ENGLISH, D. 2001. The Role of Amenities and Quality of Life in Rural Economic Growth. *American Journal of Agricultural Economics* 83:352–365.

DONALDSON, D. 2018. Railroads of the Raj: Estimating the Impact of Transportation infrastructure. *American Economic Review* 108:899–934.

DONALDSON, D. AND HORNBECK, R. 2016. Railroads and American Economic Growth: A Market Access approach. *Quarterly Journal of Economics* 131:799–858.

DONALDSON, D. AND STOREYGARD, A. 2016. The View from Above: Application of Satellite Data in Economics. *Journal of Economic Perspectives* 30:171–198.

EATON, J. AND KORTUM, S. 2002. Technology, Geography, and Trade. *Econometrica* 70:1741–1779.

FAO 2017. The State of Food and Agriculture 2017. Leveraging Food Systems for Inclusive Rural Transformation. United Nations, New York.

FLEMING, D. A. AND ABLER, D. G. 2013. Does Agricultural Trade Affect Productivity? Evidence from Chilean Farms. *Food Policy* 41:11–17.

FOSTER, A. AND ROSENZWEIG, M. 2004. Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970-2000. *Economic Development and Cultural Change* 52:509–542.

GAILE, G. 1980. The Spread-Backwash Concept. *Regional Studies* 14:15–25.

GOETZ, S. AND DEBERTIN, D. 2001. Why Farmers Quit: A County Level Analysis. *American Journal of Agricultural Economics* 83:1010–1023.

GOLLIN, D., JEDWAB, R., AND VOLLRATH, D. 2016. Urbanization With and Without Industrialization. *Journal of Economic Growth* 21:35–70.

GORELICK, N., HANCHER, M., DIXON, M., ILYUSHCHENKO, S., THAU, D., AND MOORE, R. 2017. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sensing of Environment* .

GRILICHES, Z. 1957. Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica* 25:501–522.

HALAS, M., KLAPKA, P., AND KLADIVO, P. 2014. Distance-Decay Functions for Daily Travel-to-Work Flows. *Journal of Transport Geography* 35:107–119.

HANSEN, W. G. 1959. How Accessibility Shapes Land Use. *Journal of the American Institute of Planners* 25:76–73.

HANSON, G. 2005. Market Potential, Increasing Returns and Geographic Concentration. *Journal of Economic Geography* 11:281–294.

HARRIS, C. 1954. The Market Access as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers* 44:315–348.

HARRIS, J. AND TODARO, M. 1970. Migration, Unemployment and Development: A Two-Sector Analysis. *American Economic Review* 60:126–142.

HENDERSON, J., STOREYGARD, A., AND WEIL, D. 2012. Measuring Economic Growth from Outer Space. *American Economic Review* 102:994–1028.

HENDERSON, J. V., SQUIRES, T., STOREYGARD, A., AND WEIL, D. 2017. The Global Distribution of Economic Activity: Nature, History, and the Role of Trade. *Quarterly Journal of Economics* 133:357–406.

HENRY, M., KRISTENSEN, B., BARKLEY, D., AND BAO, S. 1999. Extending Carlino-Mills Models to Examine Urban Size and Growth Impacts on Proximate Rural Areas. *Growth and Change* 30:526–548.

HERRENDORF, B., ROGERSON, R., AND VALENTINYI, A. 2014. Growth and Structural Transformation, pp. 855–941. *In* P. Aghion and S. N. Durlauf (eds.), Handbook of Economic Growth, chapter 6. Elsevier, NY.

HOOGSTRA, G., VAN-DIJK, J., AND FLORAX, R. 2011. Determinants of Variation in Population-Employment Interaction Findings: A Quasi-Experimental Meta-Analysis. *Geographical Analysis* 43:14–37.

HUGHES, D. AND HOLLAND, D. 1994. Core-Periphery Economic Linkage: A Measure of Spread and Possible Backwash Effects for the Washington Economy. *Land Economics* 70:364–377.

IMHOFF, M., LAWRENCE, W., STUTZER, D., AND ELVIDGE, C. 1997. A Technique for Using Composite DMSP-OLS City Lights Satellite Data to Map Urban Area. *Remote Sensing of Environment* 61:361–370.

INE 1995. Chile: Ciudades, Pueblos y Aldeas. Instituto Nacional de Estadísticas (INE), Santiago, Chile.

INE 2005. Chile: Ciudades, Pueblos, Aldeas y Caseríos. Instituto Nacional de Estadísticas (INE), Santiago, Chile.

IRWIN, E., BELL, K., BOCKSTAEL, N., NEWBURN, D., PARTRIDGE, M., AND WU, J. 2009. The Economics of Urban-Rural Space. *Annual Review of Resource Economics* 1:435–459.

JACOBY, H. 2000. Access to Markets and the Benefits of Rural Roads. *Economic Journal* 110:713–737.

JAYACHANDRAN, S. 2006. Selling Labor Low: Wages Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy* 114:538–575.

JEDWAB, R. AND STOREYGARD, A. 2022. The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010. *Journal of the European Economic Association* 20:1–38.

KLAIBER, H. AND PHANEUF, D. 2009. Do Sorting and Heterogeneity Matter for Open Space Policy Analysis? An Empirical Comparison of Hedonic and Sorting Models. *American Journal of Agricultural Economics* 91:1312–1318.

KURRE, J. 2003. Is the Cost of Living Less in Rural Areas? *International Regional Science Review* 26:86–116.

LANJOUW, J. AND LANJOUW, P. 2001. The Rural Non-Farm Sector: Issues and Evidence from Developing Countries. *Agricultural Economics* 26:1–23.

LONSDALE, R. E. AND BROWNING, C. E. 1971. Rural-Urban Locational Preferences of Southern Manufacturers. *Annals of the Association of American Geographers* 61:255–268.

LOPEZ, R., ADELAIA, A., AND ANDREWS, M. 1988. The Effects of Suburbanization on Agriculture. *American Journal of Agricultural Economics* 70:346–358.

LOVERIDGE, S. AND PAREDES, D. 2016. Are Rural Costs of Living Lower? Evidence from a Big Mac Index Approach. *International Regional Science Review* pp. 1–16.

MADDISON, A. 1980. Economic Growth and Structural Change in the Advanced Economies. *In* I. Leveson and J. W. Wheeler (eds.), Western Economies in Transition. Croom Helm, London.

MICHAELS, G., RAUCH, F., AND REDDING, S. 2012a. Technical Note: An Eaton and Kortum (2002) Model of Urbanization and Structural Transformation.

MICHAELS, G., RAUCH, F., AND REDDING, S. 2012b. Urbanization and Structural Transformation. *Quarterly Journal of Economics* 127:535–586.

MICHALOPOULOS, S. AND PAPAIOANNOU, E. 2014. National Institutions and Subnational Development in Africa. *Quarterly Journal of Economics* 129:151–213.

MONEKE, N. 2020. Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia.

NUNN, N. AND PUGA, D. 2012. Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics* 94:20–36.

OLFERT, M. R., PARTRIDGE, M. D., BERDEGUE, J., ESCOBAL, J., JARA, B., AND MODREGO, F. 2014. Places for Place-Based Policy. *Development Policy Review* 32:5–32.

PAREDES, D., SOTO, J., AND FLEMING, D. A. 2018. Wage Compensation for Fly-In/Fly-Out and Drive-In/Drive Out Commuting. *Papers in Regional Science* 97:1337–1353.

PARTRIDGE, M. AND RICKMAN, D. 2003. The Waxing and Waning of Regional Economies: The Chicken-Egg Question of Jobs versus People. *Journal of Urban Economics* 53:5–32.

PARTRIDGE, M., RICKMAN, D., ALI, K., AND OLFERT, M. 2009. Agglomeration Spillovers and Wage and Housing Costs Gradients Across the Urban Hierarchy. *Journal of International Economics* 78:126–140.

PERLOFF, J. 1991. The Impact of Wage Differentials on Choosing to Work in Agriculture. *American Journal of Agricultural Economics* 70:346–358.

PINKOVSKY, M. AND SALA-MARTIN, X. 2016. Lights, Camera... Income! Illuminating the National Accounts-Household Surveys Debate. *Quarterly Journal of Economics* 131:579–631.

RASMUSSEN, W. D. 1962. The Impact of Technological Change on American Agriculture, 1862–1962. *Journal of Economic History* 22:578–591.

REARDON, T., BARRET, C. B., BERDEGUÉ, J. A., AND SWINNEN, J. F. 2009. Agrifood Industry Transformation and Small Farmers in Develping Countries. *World Development* 37:1717–1727.

REARDON, T., BERDEGUÉ, J. A., AND ESCOBAR, G. 2001. Rural Nonfarm Employment and Incomes in Latin America: Overview and Policy Implications. *World Development* 29:395–409.

RENKOW, M. 2003. Employment Growth, Worker Mobility, and Rural Economic Development. *American Journal of Agricultural Economics* 85:503–513.

ROBACK, J. 1982. Wages, Rents, and the Quality of Life. *Journal of Political Economy* 90:1257–1278.

SETHIAN, J. 1996. A Fast Marching Level Set Method for Monotonically Advancing Fronts. *Proceedings of the National Academy of Sciences* 93:1591–1595.

SO, K., ORAZEM, P., AND OTTO, D. 2001. The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices. *American Journal of Agricultural Economics* 83:1036–1048.

STOREYGARD, A. 2016. Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *Review of Economic Studies* 83:1263–1295.

SU, Y., CHEN, X., WANG, C., ZHANG, H., LIAO, J., YE, Y., AND WANG, C. 2015. A New Method for Extracting Built-Up Urban Areas Using DMSP-OLS Nighttime Stable Lights: A Case Study in the Pearl River Delta, Southern China. *GIScience and Remote Sensing* 52:218–238.

THUNEN, J. V. 1826. The Isolated State. Perthes, Hamburg.

TOBIN, J. 1950. A Statistical Demand Function for Food in the U.S.A. *Journal of the Royal Statistical Society* 113:113–141.

TOLLEY, G. S., WANG, Y., AND FLETCHER, R. 1969. Reexamination of the Time Series Evidence on Food Demand. *Econometrica* 37:695–705.

UNITED NATIONS 2015. Transforming our World: The 2030 Agenda for Sustanaible Development. United Nations, New York.

UNITED NATIONS 2017. The New Urban Agenda. Habitat III. United Nations, New York.

VAN-DRIEL, H., NADALL, V., AND ZEELENBERG, K. 1997. The Demand for Food in the United States and the Netherlands: A Systems approach with the CBS Model. *Journal of Applied Econometrics* 12:509–523.

VANDERCASTEELEN, J., TAMRU, S., MINTEN, B., AND SWINNEN, J. 2018. Cities and Agricultural Transformation in Africa: Evidence from Ethiopia. *World Development* 105:383–399.

WEIBULL, J. W. 1976. An Axiomatic Approach to the Measurement of Accesibility. *Regional Science and Urban Economics* 6:357–379.

WORLD BANK 2008. Agriculture for Development: World Development Report 2008. World Bank, Washington, D.C.

WORLD BANK 2009. Reshaping Economic Geography: World Development Report 2009. World Bank, Washington, D.C.

WU, J., WEBER, B., AND PARTRIDGE, M. 2016. Rural-Urban Interdependence: A Framework Integrating Regional, Urban, and Environmental Economic Insights. *American Journal of Agricultural Economics* 99:464–480.

YANG, D. AND LIU, Z. 2012. Does Farmer Economic Organization and Agricultural Specialization Improve Rural Income? *Economic Modelling* 29:990–993.

A Appendix A

A.1 Model Appendix

The previous part of the mathematical model are standard to solve in the literature (see **?**). However, the derivation of the market access requires some additional particular assumptions. In general equilibrium, it would be the case that the value of goods purchased by a location i is equal to the income that consumers spent on that good, i.e.,

$$\delta Y_j = \sum_{i \in \Im} X_{ij}$$

In consequence (and assuming $Y_i = E_j$ the income received at the supply location i has to be equal to the expenditure spent in location j), then the gravity equation can be expressed as

$$\delta Y_j = \sum_{i \in \Im} C^{-\theta} \tau_{ij}^{-\theta} w_i^{-\theta} T_i \delta E_j P_j^{\theta} \iff 1 = C^{-\theta} \sum_{i \in \Im} w_i^{-\theta} T_i \tau_{ij}^{-\theta} P_j^{-\theta}$$
$$\iff P_j^{-\theta} = C^{-\theta} \sum_{i \in \Im} T_i w_i^{-\theta} \tau_{ij}^{-\theta} \equiv CMA_j$$

that we define as consumer market access, as Donaldson and Hornbeck (2016). Substituting the consumer market access into the gravity equation, and assuming again $Y_i = E_j$), yields

$$X_{ij} = C^{-\theta} T_i w_i^{-\theta} \tau_{ij}^{-\theta} P_j^{\theta} \delta E_j \iff X_{ij} = C^{-\theta} T_i w_i^{-\theta} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i. \tag{17}$$

Assuming that goods markets clear, so for manufacturing goods this means that $\delta Y_i = \sum_{j \in \Im} X_{ij}$, which implies that

$$\delta Y_i = C^{-\theta} T_i w_i^{-\theta} \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i \iff \delta Y_i = C^{-\theta} T_i w_i^{-\theta} FMA_j \tag{18}$$

where $FMA_j \equiv \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i$ is the firm market access in destination location $j \in \Im$. Under spatial equilibrium (workers perfectly mobile across locations), utility equalize across locations which implies that real wages also equalize across locations.[32] i.e.

$$\overline{U} = \frac{w_i}{P_i} = \frac{w_j}{P_j}. \tag{19}$$

In consequence, by substituting the nominal wage w_i in Eqn. 18, we obtain

$$\delta Y_i = C^{-\theta} T_i \overline{U}^{-\theta} P_j^{-\theta} FMA_j \iff \delta Y_i = C^{-\theta} T_i \overline{U}^{-\theta} CMA_j FMA_j \tag{20}$$

[32] Note that the assumption of spatial equilibrium is only used here at the end avoid the necessity to observe wages in the data. However, we also show results using nominal wages to see the robustness of this result to this assumption.

Another condition of the spatial equilibrium imposed in these types of models is that the labor market clears, i.e. $\delta Y_i = \delta w_i L_i$, together with the condition that $w_i = \overline{U} P_i = \overline{U}\left(CMA_i\right)^{-\frac{1}{\theta}}$ which implies that

$$L_i^C = C^{-\theta} T_i \overline{U}^{-\theta} CMA_j FMA_j w_i^{-1} \iff L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} CMA_j^{\frac{\theta-1}{\theta}} FMA_j \tag{21}$$

where $\delta L_i = L_C$ is the labor in the manufacturing sector. Taking the logs, we can express the previous equation as

$$\log L_i^C = -\theta \log C + \log T_i + (1-\theta)\log \overline{U} + \left(\frac{\theta-1}{\theta}\right)\log CMA_j + \log FMA_j$$

Assuming symmetric transport $\tau_{ij} = \tau_{ji}$, firm and consumer market access have to satisfy a proportionality condition $MA_i \equiv FMA_i = \rho CMA_i \; \forall i \in \Im$ with $\rho > 0$. Which implies that

$$FMA_i \equiv \sum_{j \in \Im} \tau_{ij}^{-\theta} CMA_j^{-1} \delta Y_i \iff MA_i = \sum_{j \in \Im} \tau_{ij}^{-\theta} \rho MA_j^{-1} w_i N_i$$

where $\delta Y_i = w_i N_i$. Substituting $w_i = \overline{U} P_i = \overline{U}\left(CMA_i\right)^{-\frac{1}{\theta}} = \overline{U} \rho^{\frac{1}{\theta}} MA_i^{-\frac{1}{\theta}}$, yields

$$MA_i = \sum_{j \in \Im} \tau_{ij}^{-\theta} \rho MA_j^{-1} \left(\overline{U} \rho^{\frac{1}{\theta}} MA_i^{-\frac{1}{\theta}}\right) N_i \iff MA_j = \overline{U} \rho^{\frac{1+\theta}{\theta}} \sum_{j \in \Im} \tau_{ij}^{-\theta} MA_i^{-\frac{(1+\theta)}{\theta}} N_i \tag{22}$$

Then, with the proportionality condition $MA_i \equiv FMA_i = \rho CMA_i$ we can express Eqn. 21 as

$$L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \left(\rho^{-1} MA_j\right)^{\frac{\theta-1}{\theta}} MA_j \iff L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \rho^{\frac{1-\theta}{\theta}} MA_j^{\frac{2\theta-1}{\theta}} \tag{23}$$

Taking the logs gives the equation of the manufacturing employment

$$\log L_i^C = k_1 + \log T_i + \left(\frac{2\theta-1}{\theta}\right)\log MA_j \tag{24}$$

where $k_1 = \log\left(C^{-\theta} \overline{U}^{1-\theta} \rho^{\frac{1-\theta}{\theta}}\right)$. Substituting Eqn. 22 in Eqn. 23 yields

$$L_i^C = C^{-\theta} T_i \overline{U}^{1-\theta} \rho^{\frac{1-\theta}{\theta}} \left(\overline{U} \rho^{\frac{1+\theta}{\theta}} \sum_{j \in \Im} \tau_{ij}^{-\theta} MA_i^{-\frac{(1+\theta)}{\theta}} N_i\right)^{\frac{2\theta-1}{\theta}}$$

Solving for the total number of workers (population) N_i and taking the logs gives

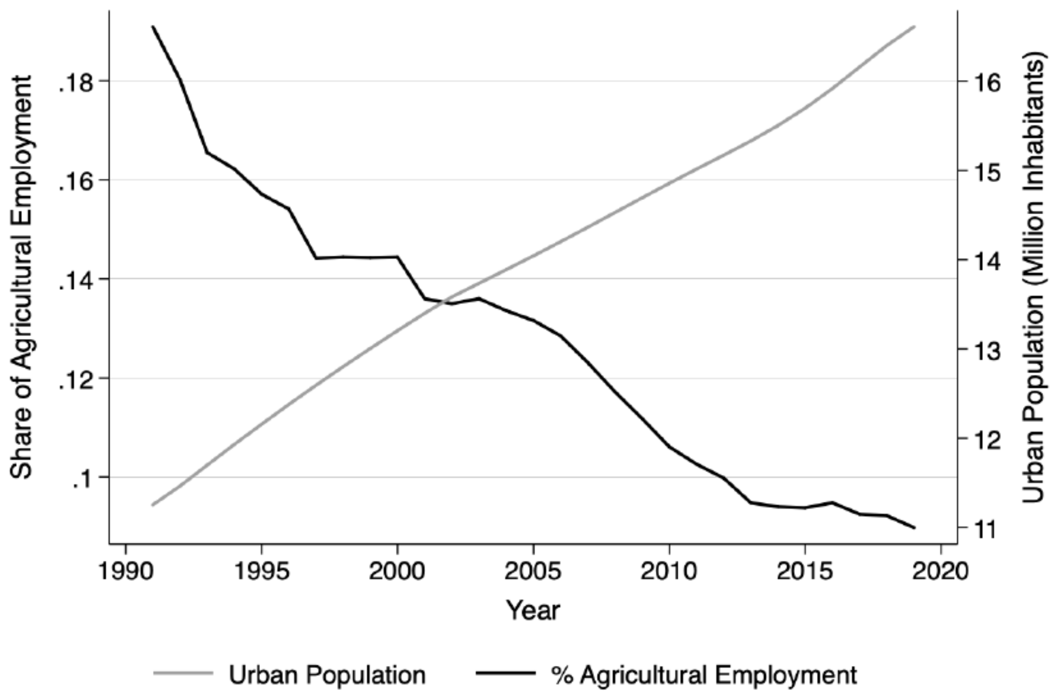
$$\iff \log N_i = k_2 + \left(\frac{\theta}{1-2\theta}\right)\log T_i + \left(\frac{2-\theta}{\theta}\right)\log MA_j + \left(\frac{\theta+1}{2\theta-1}\right)\log L_i^C \tag{25}$$

where $k2 = \log\left(C^{\frac{\theta^2-\theta-1}{2\theta-1}}\overline{U}^{\frac{\theta(\theta-3)}{2\theta-1}}\rho^{\frac{\theta(\theta-1)+(2\theta-1)(1+\theta)}{\theta(2\theta-1)}}\right)$. Which also can be expressed as employment share in manufacturing, as

$$\log\left(\frac{L_i^C}{N_i}\right) = k_3 + \left(\frac{\theta}{\theta+1}\right)\log T_i + \left(\frac{2(\theta-2)(\theta-\frac{1}{2})}{\theta(\theta+1)}\right)\log MA_j \tag{26}$$

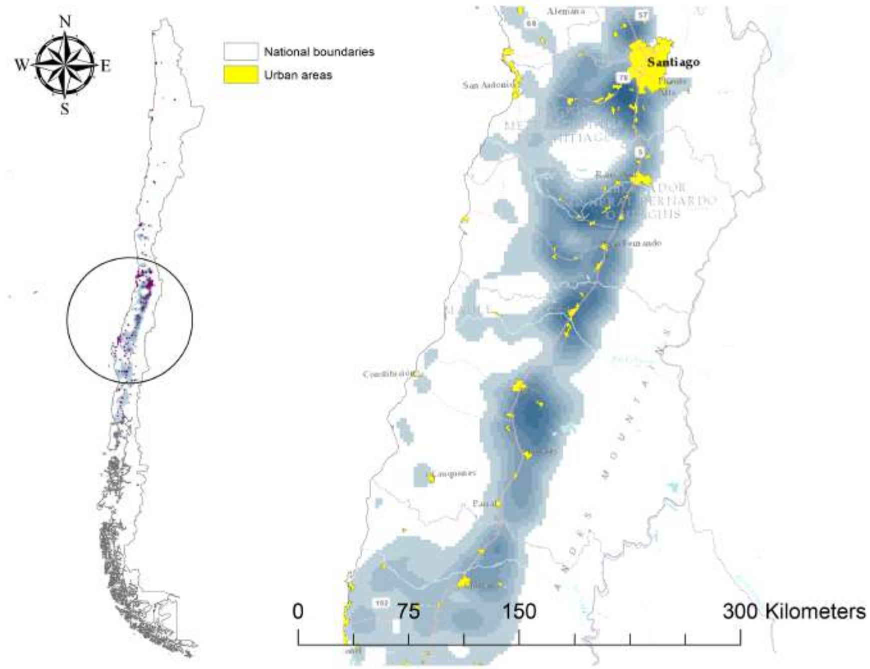
where $k_3 = \log\left(C^{-\frac{\theta^2}{\theta+1}}\overline{U}^{\frac{-\theta^2-3\theta}{\theta+1}}\rho^{\frac{\theta(\theta-1)+(2\theta-1)(1+\theta)}{\theta(\theta+1)}}\right)$.

A.2 Additional Figures



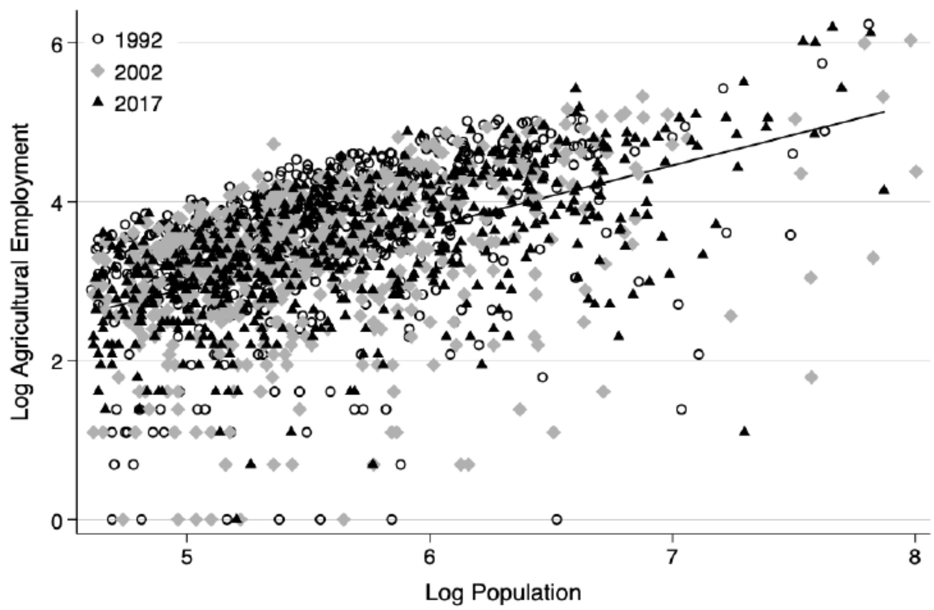
Notes: The figure displays the aggregate trends in agricultural employment and the urban population in Chile between 1991 and 2019. Over the same time the share of urban population increased by approximately 4%, from about 83.4% in 1991 to 87.6% by 2019. *Source:* Own elaboration based on data compiled by the World Bank from the International Labour Organization (ILOSTAT database), and the United Nations Population Division's World Urbanization Prospects: 2018 Revision.

Figure A5.1: Agricultural Employment and Urban Population

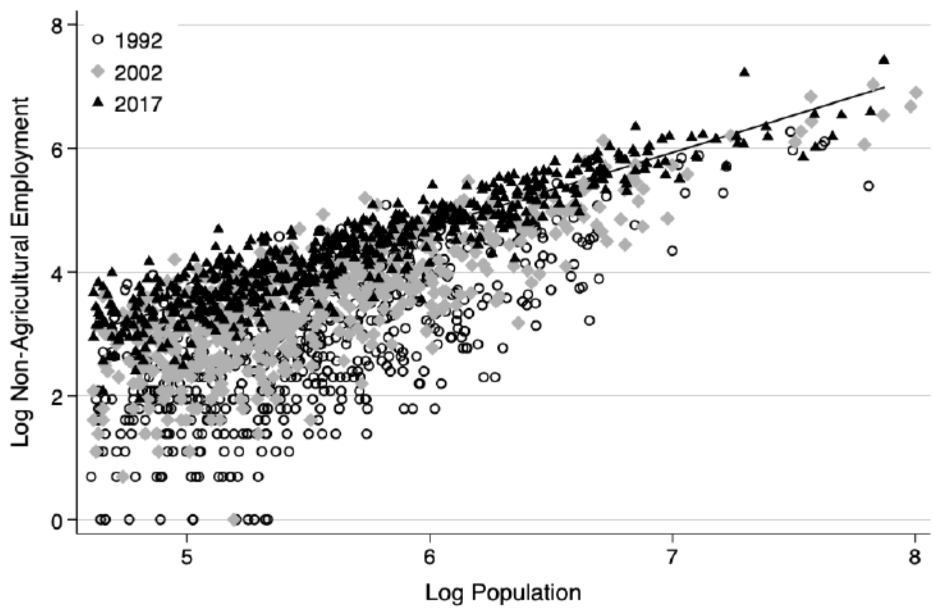


The figure describes the spatial distribution of cities and rural villages defined by the census in Chile. The yellow color represents urban areas using the administrative city borders at 2002. The blue scale represents the density of rural villages. These areas were defined as communities with less than 3,000 inhabitants. A spatial kernel density of the centroids of the rural villages weighted by their population in 2002 is used in the figure. The darker blue color represents more densely populated rural areas. Fig. A5.2 shows the spatial distribution of rural communities and cities in the country. Chile is 4,700 km long and 450 km wide at its widest point. However, agro-climatic conditions mean that most of the rural communities are located in the central-south part of the country, where the majority of the population is also concentrated as are the three most important cities in the country: the metropolitan areas of Santiago and Valparaiso—located at one and a half hours car-driving distance from each other—and Concepción—located on the coastal central south zone of Chile, at eight hours of car-driving distance from Santiago. In this area, between the cities Santiago-Valparaiso and Concepción, there are an important number of small- and medium-sized cities surrounded by a dense "green belt" of rural communities (represented by the blue-color gradient in the Fig. A5.2), where the majority of the rural population works and lives. Urban density is also particularly concentrated in this area.

Figure A5.2: The Spatial Distribution of the Urban and Rural Population



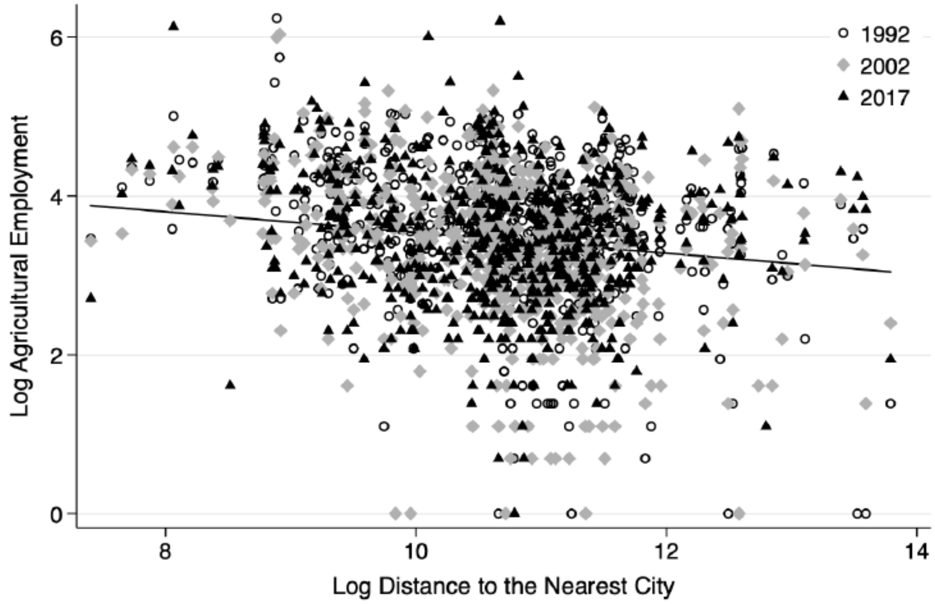
(a) Population and Agricultural Employment in Rural Villages



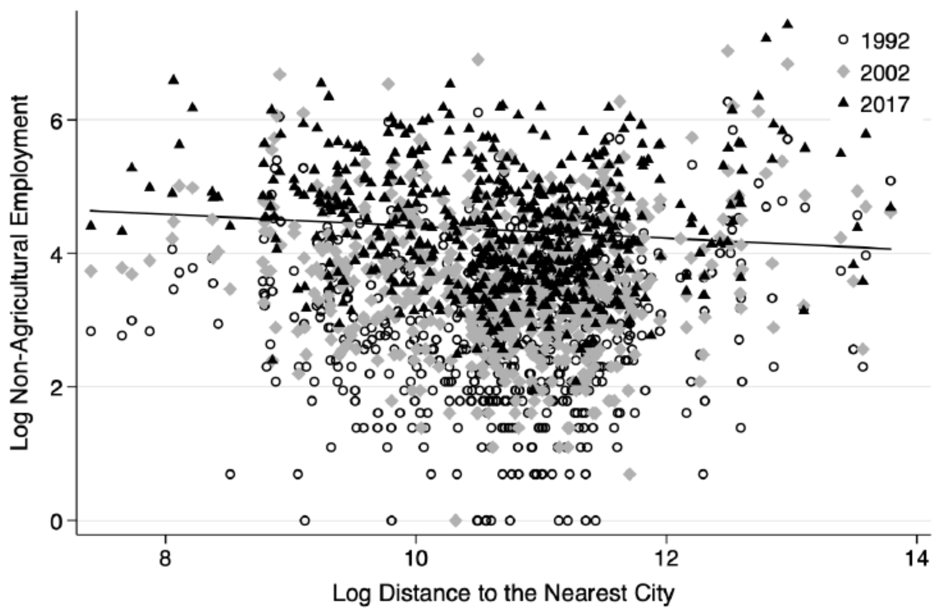
(b) Population and Non-Agricultural Employment in Rural Villages

The figure shows the linear relationship between the population and agricultural and non-agricultural employment in rural villages in Chile, Panel a) and b) respectively. The fitted line is constructed over the average values across the different years of the census. Larger rural locations are endowed with larger agricultural and non-agricultural employment. However, employment in the non-agricultural sector is larger than in the agricultural sector for more bigger rural areas. In addition, the non-agricultural sector has been increasing in time, as it is also shown in Fig. 1.

Figure A5.3: Population and Specialization in Agricultural and Non-Agricultural Activities



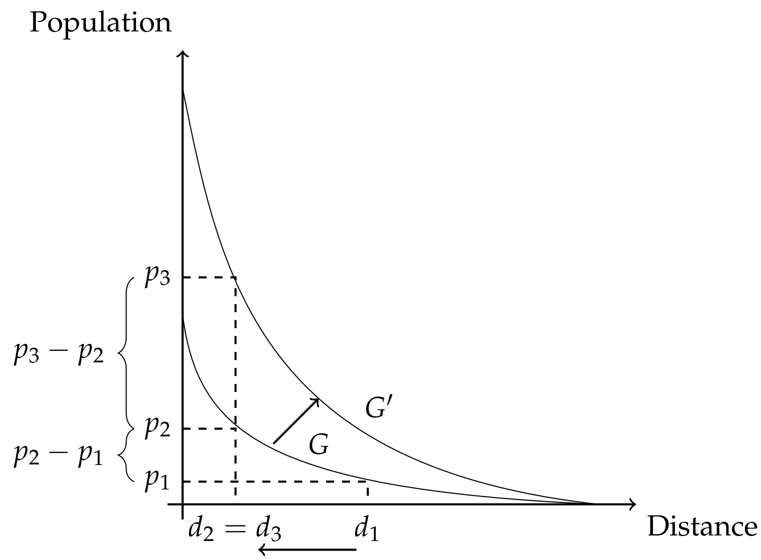
(a) Distance to the Nearest City and Agricultural Employment in Rural Villages



(b) Distance to the Nearest City and Non-Agricultural Employment in Rural Villages

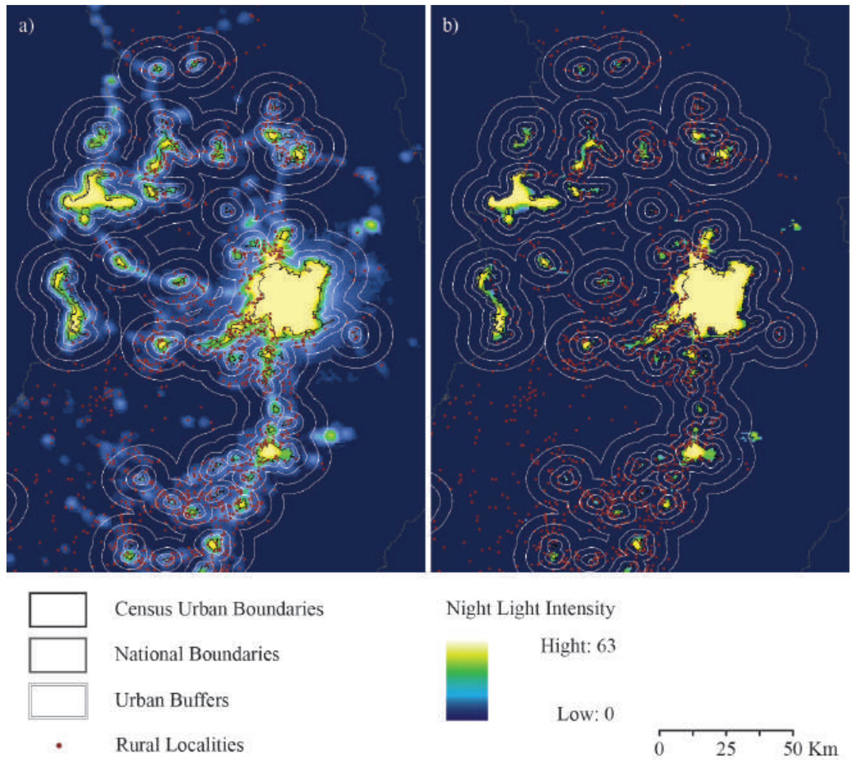
The figure shows the linear relationship between the distance to the nearest city and agricultural and non-agricultural employment in rural villages in Chile, Panel a) and b) respectively. The fitted line is constructed over the average values across the different years of the census. These values of sectoral employment for rural communities correspond to the levels over which Fig. 1 is constructed.

Figure A5.4: Access to Urban Markets and Specialization in Agricultural and Non-Agricultural Activities



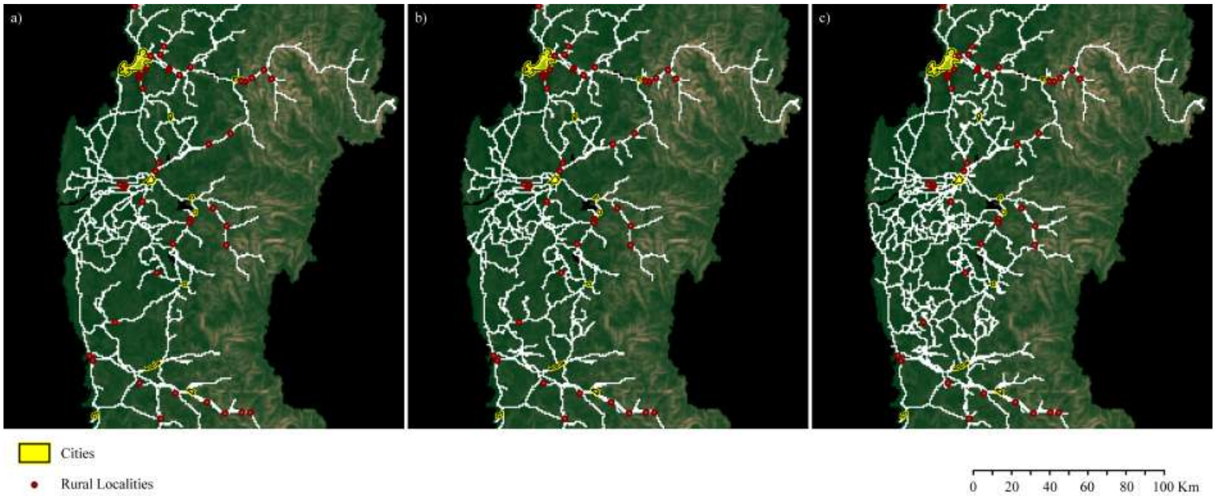
The figure compares the effect of the market access in the population of a rural community at an average distance d_1 of k urban centers, with another rural community at a distance d_2 of the same cities. Where G is the distance decay market access function, that for each rural community j is defined by $g_j = \sum_k Y_k e^{-\theta d_{j,k}}$, and $d_{j,k}$ is the distance from a city with Y_k income. The figure also illustrates the case in which the aggregated effect of medium- and small-sized cities are more important in terms of income than big cities. In such cases, the market access effect of second- and third-order cities (G') would have a higher impact than the effect of big cities in the population of rural communities (G).

Figure A5.5: Impact of Market Access on the Population of Rural Communities



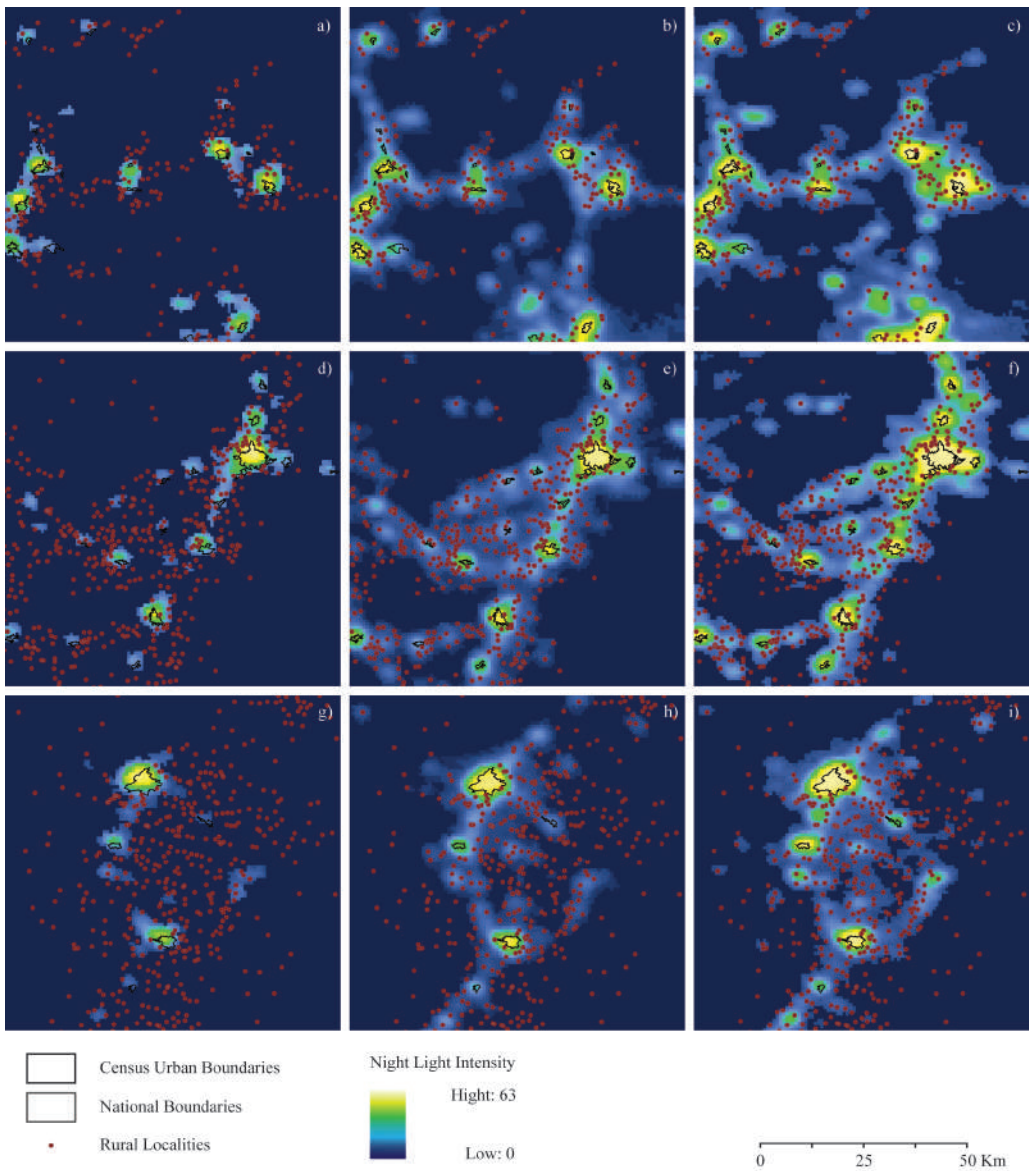
The figure describes the nighttime lights image correction for the blooming effect following (Su et al., 2015). Both images display nighttime lights in the area surrounding the metropolitan area of Santiago de Chile in 2002. Panel a) displays the nighttime lights image without correction and Panel b) shows the image corrected. In addition, the figure describes urban buffers for 2, 5, 10 an 15 kilometers from 2002 the census urban boundaries.

Figure A5.6: Nighttime Lights Image Correction



The figure shows the road network for the years 1992, 2002, and 2017, in the central-south part of Chile. Cities are displayed with a yellow polygon, while rural communities with a dark-red circle.

Figure A5.7: Road Infrastructure Development



The figures describes the spatial distribution of rural communities near medium-sized cities in Chile. The left-hand side panel of the figure shows the scenario for 1992, the center at 2002, and the right-hand side for 2013. Urban boundaries (in red), are the official boundaries delimited for the national census of 2002 by the Chilean National Statistical Office (INE).

Figure A5.8: Medium-Sized City Growth and Rural Communities

Figure shows the spatial coverage of LandSat L5 and LandSat L7 day-time satellite images for the census years of 1992 and 2002. Information is missing for most of the area of the south of Chile until 1998. These images were used to compute the vegetation indexes (proxies of agricultural productivity).

Figure A5.9: Spatial Coverage of Day-Time Satellite Images

A.3 Data Appendix

Table A5.1: Descriptive Statistics

	1992	2002	2017	$\geq 20km$	$\geq 50km$	$\geq 100km$
Population and Employment						
Population	311.41	328.15	391.29	329.03	394.79	669.76
	(360.52)	(420.43)	(347.17)	(378.77)	(490.14)	(707.70)
Employment	86.80	105.75	165.31	109.83	152.88	283.02
	(107.52)	(155.89)	(168.29)	(150.22)	(238.65)	(316.48)
Farm Employment	47.62	47.62	47.62	41.30	24.77	22.88
	(45.73)	(45.73)	(45.73)	(34.85)	(19.36)	(18.45)
Non-Farm Employment	39.18	39.18	39.18	48.44	83.60	187.06
	(84.00)	(84.00)	(84.00)	(99.06)	(138.41)	(218.42)
Market Access						
Distance to a City	20.95	20.95	20.95	39.52	83.25	160.64
	(22.39)	(22.39)	(22.39)	(28.88)	(46.88)	(54.32)
Baseline	-13.77	2.29	1.73	-4.06	-5.10	-6.52
	(1.56)	(1.43)	(1.23)	(7.62)	(7.74)	(7.78)
Distance Decay ($a = 50$)	-2.48	-2.88	-3.25	-5.28	-11.93	-34.68
	(6.59)	(5.30)	(7.16)	(7.91)	(14.68)	(19.13)
Distance Decay ($a = 200$)	-0.09	-0.13	-0.21	-1.01	-3.80	-12.29
	(3.92)	(1.95)	(2.49)	(3.27)	(6.67)	(9.72)
City Size Adjusted	0.47	0.40	-0.57	-0.85	-3.11	-8.28
	(1.25)	(1.30)	(1.75)	(2.01)	(3.51)	(3.92)
Observations	1292	1292	1292	1377	240	48
Agricultural Productivity						
GCVI	-0.14	-1.03	-0.30	-0.70	-1.14	-1.81
	(0.51)	(0.93)	(0.65)	(0.93)	(1.05)	(1.17)
NDVI	-1.41	-1.87	-1.36	-1.61	-1.95	-2.41
	(0.59)	(0.88)	(0.57)	(0.77)	(0.91)	(0.99)
EVI	-2.50	-2.80	-2.19	-2.53	-2.92	-3.40
	(0.66)	(0.93)	(0.63)	(0.87)	(0.97)	(1.01)
Observations	433	1152	1220	944	177	32

Mean and standard deviation (in parentheses).

Table A5.2: Data Description (GIS Variables)

Variable	Description	Primary Source
1. Baseline Market Access	$g_j = \sum_j Y_k e^{-d_{j,k}}$	NASA DMSP-OLS
2. Distance Decay Market Access with $a = 50$	$g_j = \sum_j Y_k e^{-\frac{1}{2(50)^2} d_{j,k}}$	NASA DMSP-OLS
3. Distance Decay Market Access with $a = 100$	$g_j = \sum_j Y_k e^{-\frac{1}{2(200)^2} d_{j,k}}$	NASA DMSP-OLS
4. City Size Adjusted Market Access (Population)	$g_j = \sum_j Y_k e^{-\theta_m d_{j,k}}$, $\theta_m =${0.0196(large-cities), 0.0052(medium-sized-cities), 0.0113(small-cities), 0.0095(towns)}	NASA DMSP-OLS
5. Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{(NIR-Red)}{(NIR+Red)}$, pixel by pixel annual average at 500m spatial resolution	NASA/USGS LandSat L5 and L7
6. Enhanced Vegetation Index (EVI)	$EVI = \frac{2.5(NIR-Red)}{(NIR+6*Red+7*Blue+1)}$, pixel by pixel annual average at 500m spatial resolution	NASA/USGS LandSat L5 and L7
7. Green Chlorophyll Vegetation Index (GCVI)	$GCVI = \frac{NIR}{Green} - 1$, pixel by pixel annual average at 500m spatial resolution	NASA/USGS LandSat L5 and L7
8. Ruggedness	Average (interpolated) terrain ruggedness in a neighborhood of 2km from the centroid of the rural community using terrain elevation data	NASA SRTM 90m DEM
9. Precipitation of driest month	Average precipitation (≈1960-1990) of driest month	WorldClim v1.4
10. Min temperature in coldest month	Average (interpolated) minimum temperature (≈1960-1990) in coldest month	WorldClim v1.4
11. Distance to minor water sources	Euclidean distance between the centroid of the rural community and the nearest minor source of water (polyline shape with water springs and wells)	SIIT-BCN-Chile

For 1–6, $d_{j,k}$ is the Euclidean distance between the centroid of a rural community j and the centroid of a city k, with Y_k sum of night-time light intensity. For 7–12, pixels were interpolated for an area of 2km from the centroid of each rural community.

A.3.1 Agricultural Activities and Occupations in Rural Areas

On the definition of farm employment, the following activities were considered:

- Agriculture, livestock, hunting and related activities
- Crops: market-oriented crops, horticulture
- Domestication and animal breeding
- Mixed exploitation of crops, domestication and animal breeding
- Agriculture and livestock services (except veterinary)
- Hunting and related activities
- Forestry, timber extraction and related activities
- Agroforestry
- Fishing and related activities

On the definition of farm employment by occupation, is considered the following activities, from the Census of 1992:

- Farmers and market oriented qualified agricultural workers and fisherman
- Farmers and market oriented qualified workers in agricultural farms, forestry, and fisherman
- Farmers and qualified workers in market-oriented crops
- Farmers and qualified workers in extensive crops
- Farmers and qualified workers in forest plantation and shrubs
- Farmers and qualified agricultural workers in orchards and greenhouses
- Farmers and qualified workers in mixed crops cultivation
- Market oriented breeders and livestock qualified workers
- Cattle and other domestic breeders, diary produces and subproducts
- Poultry farmers and qualified workers in poultry (eggs, meet, and feathers)
- Beekepers and sericulturists, and qualified workers in these items
- Breeders and qualified workers in breeders and other domestic animals
- Breeders and qualified livestock workers in breeding of wild animals for markets
- Market oriented producers and agricultural qualified workers

- Qualified workers in forest plantation and related tasks
- Three fellers and choppers, and other workers in forest related activities
- Charcoal workers and related activities
- Fishermen, hunters, and trappers
- Breeders of aquatic species
- Freshwater and coastal fishermen
- Deep sea fishermen (fishing and related tasks)
- Hunters and trappers
- Subsistence farming and fishing workers

Table A5.3: Market Access, Population and Farm and Non-Farm Employment

	OLS			2SLS		
	Log Population	Log Farm Employment	Log Non-Farm Employment	Log Population	Log Farm Employment	Log Non-Farm Employment
Panel a): Baseline Market Access with $MA_c = \sum_k C_k e^{-d_{c,k}}$						
Log Market Access	0.105	0.093	0.178*	0.316**	-0.134	0.951***
	(0.069)	(0.083)	(0.102)	(0.163)	(0.212)	(0.254)
R^2	0.470	0.528	0.646			
F-stat				114.561	114.485	122.253
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel b): Distance Decay Market Access Function with $MA_c = \sum_k C_k e^{-\frac{1}{2(50)^2} d_{c,k}}$						
Log Market Access	0.021	0.012	0.030	0.097*	-0.042	0.301***
	(0.018)	(0.020)	(0.025)	(0.163)	(0.066)	(0.086)
R^2	0.470	0.527	0.646			
F-stat				50.103	47.874	50.195
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel c): Distance Decay Market Access Function with $MA_c = \sum_k C_k e^{-\frac{1}{2(200)^2} d_{c,k}}$						
Log Market Access	0.107	0.064	0.151	0.396*	-0.170	1.222***
	(0.078)	(0.084)	(0.105)	(0.205)	(0.271)	(0.343)
R^2	0.471	0.528	0.646			
F-stat				53.674	51.885	54.801
Observations	1,451	1,451	1,451	1,451	1,451	1,451
Panel d): City-Size Adjusted Market Access Function with $MA_c = \sum_k Y_k e^{-\hat{\theta}_m d_{j,k}}$						
Log Market Access	0.153*	0.095	0.167**	0.326*	-0.099	0.693***
	(0.080)	(0.066)	(0.082)	(0.167)	(0.156)	(0.183)
R^2	0.473	0.529	0.647			
F-stat				120.152	136.927	147.121
Observations	1,451	1,451	1,451	1,451	1,451	1,451

$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument historical urban market access.

Table A5.4: Market Access, Agricultural Potential and Farm Employment

	Log Farm Employment					
	OLS			2SLS		
	GCVI	NDVI	EVI	GCVI	NDVI	EVI
$A^F_{jt} \ln MA$	-0.034**	-0.033**	-0.028**	-0.492***	-0.365***	-0.296***
	(0.015)	(0.014)	(0.012)	(0.167)	(0.112)	(0.088)
A^F_{jt}	0.169***	0.066	0.082	0.876***	0.254**	0.183**
	(0.066)	(0.065)	(0.063)	(0.275)	(0.100)	(0.078)
R^2	0.562	0.547	0.547			
Observations	1,021	1,008	1,010	1,021	1,008	1,010
F-stat				15.675	20.379	23.066
P-val LM stat				0.000	0.000	0.000

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument historical urban market access.

Table A5.5: Farm to Non-Farm Employment Multipliers

	Log Non-Farm Employment	
	OLS	2SLS
Log Farm Employment	0.330***	0.865***
	(0.050)	(0.303)
R^2	0.672	
Observations	1,451	1,451
F-stat		17.491
P-val LM stat		0.000

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. *Note*: All models include municipality fixed effects, year fixed effects, municipality-year fixed effects, and cubic polynomial in latitude and longitude of the centroid of the rural community. The 2SLS uses as an instrument agricultural potential captured with satellite images.

Table A5.6: Robustness Checks Using Alternative Measures of Agricultural Potential

Panel a): Including EVI	(1) Population	(2) Farm Employment	(3) Non-Farm Employment	(4) Population	(5) Farm Employment	(6) Non-Farm Employment	(7) Population	(8) Farm Employment	(9) Non-Farm Employment
1. Baseline Specification	1.183**	3.152***	1.959***	-0.060	2.938**	0.882*	1.844***	2.562***	1.995***
	(0.504)	(0.923)	(0.527)	(0.159)	(1.392)	(0.501)	(0.526)	(0.592)	(0.510)
2. Distance Decay Market Access with $a = 50$	0.140**	0.526***	0.458***	0.046*	1.036**	0.718***	0.077**	0.084**	0.085**
	(0.068)	(0.145)	(0.104)	(0.026)	(0.412)	(0.184)	(0.033)	(0.040)	(0.038)
3. Distance Decay Market Access with $a = 100$	0.583**	1.964***	1.579***	0.171*	4.305**	2.895***	0.383***	0.340**	0.383***
	(0.269)	(0.571)	(0.380)	(0.097)	(1.672)	(0.732)	(0.144)	(0.160)	(0.141)
4. City-size Adjusted Market Access	2.065***	1.101**	0.985***	0.367***	1.519**	1.278**	1.015**	1.204***	1.271***
	(0.432)	(0.464)	(0.317)	(0.133)	(0.752)	(0.546)	(0.416)	(0.370)	(0.383)
Panel b): Including NDVI	(10) Population	(11) Farm Employment	(12) Non-Farm Employment	(13) Population	(14) Farm Employment	(15) Non-Farm Employment	(16) Population	(17) Farm Employment	(18) Non-Farm Employment
1. Baseline Specification	1.262**	3.108***	1.867***	-0.003	3.145**	0.833*	1.715***	2.491***	2.022***
	(0.494)	(0.943)	(0.536)	(0.156)	(1.482)	(0.483)	(0.396)	(0.585)	(0.462)
2. Distance Decay Market Access with $a = 50$	0.112*	0.606***	0.514***	0.049*	1.248***	0.868***	0.074**	0.084**	0.086**
	(0.068)	(0.161)	(0.118)	(0.026)	(0.484)	(0.243)	(0.032)	(0.041)	(0.039)
3. Distance Decay Market Access with $a = 100$	0.446*	2.339***	1.841***	0.179*	5.150***	3.496***	0.368***	0.341**	0.388***
	(0.268)	(0.649)	(0.447)	(0.098)	(1.987)	(1.005)	(0.142)	(0.163)	(0.145)
4. City-size Adjusted Market Access	1.941***	1.035**	0.949***	0.349***	1.690**	1.460***	1.090**	1.152***	1.196***
	(0.394)	(0.474)	(0.333)	(0.130)	(0.846)	(0.559)	(0.430)	(0.377)	(0.383)

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. All models were estimated by 2SLS and include control variables and fixed effects at municipality level. The EVI (Panel a) and NDVI (Panel b) is used as a proxy of agricultural productivity and included in the Farm Employment Equation. The The EVI is computed as $EVI = 2.5 * (NIR - red)/(NIR + 6 * Red + 7 * Blue + 1)$, and NDVI is computed as $NDVI = (NIR - Red)/(NIR + Red)$, where NIR is the Near Infrared Band and each color represents a wavelength band of the satellite sensors. Images from LanSat L5 was used for 1992 and LanSat L7 for 2002. Due to spatial coverage of L5 this set of regression is estimated for 189 rural communities (see Fig. A5.9 for a visualization of the spatial coverage of daytime satellite images in years of censuses).