Weather Shocks and the Geography of Agriculture: Evidence from the 1975 Frost in Brazil*

Bruno Barsanetti†

EPGE Brazilian School of Economics and Finance

April 2021

Abstract

How do weather shocks affect the spatial distribution of employment in agriculture? I investigate this question by examining the 1975 frost that damaged coffee trees in the Brazilian state of Paraná. I find that the frost damages had persistent effects on the spatial distribution of employment in agriculture. I identify the effects of the frost damages by comparing changes in agricultural employment across local economies that had different coffee tree densities at the time of the frost and that were in states differently affected by it. The frost resulted in a large and persistent displacement of agricultural workers. Agglomeration economies within the coffee sector and changes in frost risk perceptions are plausible explanations for why the effects on employment were persistent.

JEL Classification Codes: R11, O12, E22.

*This work benefited from the advice of Cynthia Kinnan, Seema Jayachandran, and Chris Udry; from comments of Çagri Akkoyun, Emily Breza, Cláudia Bruschi, José Carreño, Eduardo Cenci, Janjala Chirakijja, Gaby Cugat, Kyle Emerick, Alípio Ferreira, Haritz Garro, Sara Hernandez, Yutaro Izumi, Kiminori Matsuyama, Martí Mestieri, Ameet Morjaria, Pinchuan Ong, Marcel Ribeiro, and Leslie Streicher; and from seminar participants at many institutions and conferences. Camila Barsanetti, Mateus Vencato and Simone Vencato provided much appreciated help with the data. I appreciate the hospitality of the *Arquivo Nacional*, Rio de Janeiro, and the *Museu do Café*, Santos. I also acknowledge research funding from the Center for the Study of Development Economics and from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. All errors are mine.

†E-mail: bruno.barsanetti@fgv.br

1 Introduction

Few economic sectors are as at-risk from large temporary shocks as agriculture. This is particularly true with respect to environmental shocks: droughts, frosts, floods, and pests often harm agricultural production. As climate change increases weather variability across the tropics (Bathiany *et al* 2018), shocks to agriculture tend to become more frequent. Given that many populations depend on the sector, particularly in the developing world, these shocks may result in the movement of workers away from affected economies. In fact, there is plenty of evidence that climate shocks cause the migration of agricultural workers; see, e.g., Gray and Mueller (2012a), Minale (2018), and Jessoe *et al* (2018). However, there is still little evidence on whether and under which circumstances such short-run migration flows persist and bring lasting changes to the spatial distribution of employment in agriculture. Although a large and growing literature has examined whether the spatial distribution of economic activity is persistently affected by large temporary shocks, it has mostly ignored agriculture; see, e.g., Davis and Weinstein (2002), Brakman *et al* (2004), Schumann (2014), Hanlon (2017), or Siodla (2021).

This paper documents short-run and long-run changes to the spatial distribution of agricultural employment after an extreme weather event. In 1975, a major frost hit southern Brazil, effectively damaging all coffee trees in the state of Paraná. Coffee trees are a durable and highly specific type of capital, so the frost may be interpreted as a shock to capital in the coffee sector. At the time, Paraná was the largest coffee producing region in the country, a position it lost after the frost. Relative to most alternative land uses, coffee uses more labor per hectare, so a decline of the coffee industry reduces the demand for farm labor. In fact, the empirical findings in this paper suggest that the 1975 frost caused the decline of agricultural employment in coffee producing regions of Paraná. Such decline seems to persist over time.

To identify the effects of the coffee capital destruction of the frost on the spatial distribution of agricultural employment, I estimate the association between the coffee tree density a municipality had right before the frost and changes in agricultural employment it experienced over time. I measure coffee tree density with information from the Brazilian agricultural census, which records the number of trees harvested in 1975. The frost occurred after the harvest, so the density of harvested trees measures coffee capital intensity right before the frost. Identification is challenging because other contemporary or future shocks that correlate with coffee tree density in 1975 could differentially affect municipalities with higher coffee tree densities. To address this challenge, I exploit the stronger impact of the frost in Paraná in comparison with other states. One of the worst frosts on record, the 1975 frost damaged essentially all coffee trees in Paraná, but its effects were weaker in other coffee producing states, such as São Paulo. I then identify the effects of the frost damages to coffee from the relation between changes in agricultural employment and pre-frost coffee tree density observed in the affected state of Paraná but not in the control state of São Paulo.

Note that the identification strategy is resembles a triple-differences model. The identification assumption is that, in the absence of the frost, the relative agricultural employment growth of coffee denser municipalities in the two states would be similar, reflecting common trends associated with specialization in coffee. Under this assumption, the difference between the coefficients on coffee tree density for Paraná and São Paulo after 1975 represents the effects of the more severe frost in Paraná. The fact that the coffee industries in Paraná and São Paulo were on similar trajectories since the mid-1960s lends credibility to the identification assumption. As further support for it, I show that estimates for the relative employment growth of coffee denser municipalities in the pre-frost period of 1970 to 1975 are the same in both states.

The empirical findings suggest that the coffee capital destruction in Paraná led to lower employment in agriculture. There were no clear effects on average worker earnings, nor on employment in the non-agricultural sector. Instead, agricultural workers responded by migrating to other regions. The effects on farm employment are due to changes right after the frost, from 1975 to 1980, which supports the interpretation that the estimates identify the causal effect of the frost. Moreover, subsequent agricultural employment growth was not faster in affected municipalities, suggesting that the effects of the frost on the spatial distribution of farm employment were persistent.

The main results survive a variety of robustness exercises. The use of an alternative control state leads to qualitatively similar results. The results are robust to the inclusion of initial economic conditions and pre-trends as control variables. The main conclusions also hold when using alternative inference methods, indicating that the results are not driven by spurious correlations generated by spatial correlation of the variables.

Two hypotheses, or a combination of them, are the most plausible explanations for why the 1975 frost brought persistent changes to the spatial distribution of agricultural employment. They are the only two hypotheses that seem fully consistent with the empirical findings. Both hypothesis explain the decline of agricultural employment as a consequence of the decline of the coffee industry after the frost.

First, the persistent effects of the frost are consistent with multiple spatial equilibria in agriculture, an explanation that depends on agglomeration economies being sufficiently strong to overcome dispersion forces (Proost and Thisse 2019). Under this explanation, coffee trees anchor the economy on a steady-state in which most farmers plant coffee. The frost destroys the coffee trees and agglomeration economies generate strategic complementarities that allow the economy to move away from coffee. Although there is still little empirical evidence for agglomeration economies in agriculture, some recent research suggests they may occur (Holmes and Lee 2012, Richards

2017). In fact, agglomeration economies within the coffee sector would be a consequence of farmers sharing supply and distribution channels, specialized labor and equipment, or information. To better evaluate if these agglomeration economies exist, I exploit the fact that the planting of coffee trees in Paraná should be avoided in south-facing terrains. Using high-resolution data on the modern location of coffee trees, I show that coffee farmers are more likely to plant coffee when the share of neighboring farms that face south is lower, conditional on a farmer's terrain own cardinal orientation. I interpret this finding as suggestive evidence of agglomeration economies within the coffee industry.

Second, the fear of future frosts could have shifted farmers away from coffee production. Although the technical opinion at the time was that the 1975 frost should not change the assessment of frost risk or coffee suitability, anecdotal evidence suggests that farmers' frost risk perceptions may have changed. In other settings, previous research has shown that disaster occurrence may affect disaster risk percepctions; see, e.g., Gallagher (2014) and Brown *et al* (2018). Although there is no comprehensive data on beliefs about frost risk to test this hypothesis directly, it is consistent with the negative effects of the frost on land prices observed in the data (as is the multiple spatial equilibria hypothesis).

I also examine if, in the absence of the explanations above, alternative hypotheses explain the empirical findings. Historical evidence points against explanations that the coffee industry declined as a consequence of changes in government policy toward the sector. The interpretation that the frost simply accelerated an ongoing decline of the coffee industry is inconsistent with investment before the frost and negative effects of the frost on land prices. Labor market effects show no evidence that the frost reduced the spatial misallocation of labor.

The rest of the paper is structured as follows: the remaining of this section presents the related literature; Section 2 discusses the historical background; Section 3 presents the identification strategy and the data; Section 4 presents the empirical findings; Section 5 discuss the robustness of the findings; Section 6 examines plausible explanations for the results; Section 7 concludes.

1.1 Related Literature

The main contribution of this paper is to a large literature that documents migration responses to temporary weather shocks or other natural disasters. Migration has been shown to be affected by a variety of shocks, including temperature shocks (Dillon *et al* 2011, Gray and Mueller 2012a, Jessoe *et al* 2018), rainfall (Munshi 2003, Bazzi 2017, Kleemans and Magruder 2018, Minale 2018), floods (Hornbeck and Naidu 2010), crop losses (Gray and Mueller 2012b), hurricanes (Gröger and Zyllberberg 2016, Mahajan and Yang 2020), and tornadoes (Boustan *et* *al* 2020). However, few papers examine whether the migration effects are permanent. Bohra-Mishra *et al* (2014) document that there is some heterogeneity with respect to which shocks cause long-run migration movements. Hornbeck and Naidu (2010) show that a flood in the United States persistently reduced agricultural employment due to the adoption of labor-saving techniques in agriculture. My paper examines a specific event in which persistent changes in the spatial distribution of agricultural employment are associated with a decline of a labor-intensive crop. The results suggest the importance of two new factors, beliefs about future risks and agglomeration economies within agriculture, as a reason for why some shocks to agriculture could cause persistent migration.

This paper also contributes to a literature that examines if and how temporary shocks can persistently affect the spatial distribution of economic activity. After many of these shocks, the spatial distribution of economic activity seems to return to what it was before; see Davis and Weinstein (2002,2008), Brakman *et al* (2004), Miguel and Roland (2011), and Waldinger (2016). In contrast, persistent effects were documented by Schumann (2014), Hanlon (2017), Juhász (2018), Ager *et al* (2020), and Siodla (2021). Relative to these events, the frost was a shock concentrated in a specific sector and type of capital, which allows a better understanding of why its effects were persistent.[1] The main difference of my paper in comparison to the literature is that most of the literature focuses on changes in the spatial distribution of non-agricultural activities, such as relative city sizes, while my main variable of interest is the distribution of agricultural employment. Feigenbaum *et al* (2020) documents the effects of a war destruction episode on both agriculture and manufacturing. The authors do not find effects on the spatial distribution of agricultural employment but find a persistent contraction of investment in the sector, which they explain as a result of underdeveloped credit markets.

2 Historical Background

Since the mid-19th century, Brazil has been the world's largest coffee producer and the crop one of its main agricultural industries. In Paraná, coffee production took off the early 20th century. By mid-century, the state was the leading coffee producer in Brazil, benefitting from fertile lands, high coffee prices in the years after World War II, and labor flowing from other states. In 1975, Paraná was responsible for 48% of national coffee production and 26% of global production.[2]

Kohlhepp *et al* (2014) and Pozzobon (2006) describe Paraná's coffee economy as a dynamic one, characterized

[1] Hanlon (2017) and Juhász (2018) also examine sectoral shocks. Their focus is on shocks to factor or product prices, while the temporary shock I consider is the destruction of physical capital.

[2] Author's calculations, based on IBGE and FAO data.

by social and spatial mobility, small farmers, and an entrepreneurial spirit typical of frontier regions.[3] Still, coffee was a stable industry, in which investment in new trees kept the capital stock constant.[4]

In the morning of July 18, 1975, after the coffee harvest period for that year, the state of Paraná was affected by the most severe frost in over half-century. Temperatures dropped far below the killing point of coffee trees, reaching -5°C in many regions (Pozzobon 2006). It was a severe type of black frost, the most feared by farmers since the low humidity and strong winds freeze and kill large portions of the plants. The frost had large effects on coffee trees, that had to be cut close to the ground. Surviving trees would take at least 2 years to start producing again and 3 years to resume their production levels, roughly the same time new trees take to start producing (GERCA 1975). Historical accounts indicate that the damage was uniform throughout the state (Oliveira 1978). The day after the frost, Jayme Canet, then governor of Paraná and a coffee farmer himself, was gloomy in his evaluation of the losses: "I already know what next year production will be: zero" (Pozzobon 2006). In the end, his prediction proved accurate: the 1976 coffee production was only 0.4% of the output from the year before.

International coffee markets reacted to the frost, with prices doubling in less than a year. Other events, such as the decline of the Angolan coffee industry after independence, also played a role in the price rise (Pendergast 2010). Throughout the late 1970s and most of the 1980s, coffee prices remained higher than they were in the early 1970s, before the frost. In fact, the decade following 1975 was a profitable period for coffee production, with production increasing in other countries such as Colombia and other Brazilian states such as São Paulo and Minas Gerais. But in Paraná, the world's largest coffee producing region, despite the higher prices, the industry did not recover its pre-frost levels.

Experts in the Paraná coffee industry, however, expected recovery to occur (Mores 2017). In a government-issued booklet, which was distributed to farmers soon after the frost (GERCA 1975), agronomists wrote that the frost should not affect farmers' decisions of whether to grow coffee. An excerpt reads: "the 1975 frost did not affect the evaluation of weather risks in the affected zones. Frosts like these are exceptional and extremely rare." The position was shared by the policy makers, who implemented credit policies favoring the recovery of the coffee sector. As a newspaper correspondent (from the *Estado de São Paulo*) in Paraná wrote in 1978: "good judgment and the acknowledgment of northern Paraná's (coffee growing) potential predominated, and the coffee activity was incentivized in the region" (Oliveira 1978, p.71).

[3] See also Margolis (1973) for an ethnographic study of a Paraná coffee growing community in the pre-frost period, Mores (2017) for an environmental history of coffee production in the state, and Carvalho (1991) for a description of the changes undertook by the Paraná coffee industry from 1960 to 1985.

[4] In 1970, 13.4% of the coffee trees in Paraná were young (less than 3 years old). Coffee trees at the time were usually kept for at least twenty years (Pozzobon 2006, p. 81).

Figure 1 graphs the series of coffee harvested areas (Panel A) and of coffee production (Panel B) in Paraná and in Brazil; the series are scaled by their value in 1975. Due to the damage to the trees, the frost directly depressed Paraná coffee production in 1976 and 1977. Moreover, the recovery of the Paraná coffee industry was only partial, and its production level never went back to what it was before the frost.

In the rest of Brazil, the coffee industry grew during the period that followed the frost. The neighboring state of São Paulo, which was the second-largest coffee producing state in 1975, experienced a much weaker frost. By 1977, production in São Paulo had already surpassed its 1975 level, remaining high during much of the 1980s. Production growth was even more impressive in Minas Gerais, the third-largest coffee producing state in 1975, a state where the frost led to minimal damages. Over time, Minas Gerais would become the largest coffee-producing region of Brazil. In 1975, Paraná, São Paulo and Minas Gerais accounted for 92% of the Brazilian coffee production; the tables and maps presented in this section focus on the three states.

The relative decline of the Paraná coffee industry can be seen in the investment in new trees, which are shown in Table 1. In 1970, 13.4% of the coffee trees in Paraná were young; in São Paulo and Minas Gerais the share was, respectively, 12% and 17.2%. These numbers indicate that the industry was stable before the frost. By 1980, Paraná had a lower share than the other two states: 15.5% against 18.1% and 31%. The difference is more remarkable when we consider only the share of trees that were planted in 1980: only 5.2% of the Paraná trees had been planted in that year, as opposed to 7.8% in São Paulo and 16.9% in Minas Gerais. These numbers illustrate how the late 1970s were a favorable period for investment in coffee. In Paraná, however, investment was low.

Figure 2 maps the spatial distribution of the coffee industry over time, as measured by the coffee tree density from the Brazilian agricultural censuses.[5] In 1970, the coffee industry was concentrated in the state of Paraná. By 1980, however, this was no longer the case: the coffee industry grew in São Paulo and Minas Gerais, which became the leading coffee producing areas of Brazil, while it declined in Paraná. The right map on the bottom panel shows coffee tree density in 2006, which was concentrated in Minas Gerais; clusters of coffee production in Paraná and São Paulo survived up to 2006. The maps reveal two patterns of the Brazilian coffee industry: it is geographically concentrated and the spatial distribution dramatically shifted over time.[6]

[5] Appendix Figure B.2 shows the location of the study states in Brazil.

[6] A recent study by IBGE (2016) details the geographical changes in the production of coffee since the 1975 frost. See also Moreira and Laverdi (2016) for a history of the Paraná coffee sector after the frost. In Appendix B.1, I measure indexes of the spatial concentration of the coffee industry over time. In Paraná, the coffee industry became more concentrated after the frost.

3 Empirical Approach

3.1 Empirical Framework

To examine the effects of the coffee capital destruction of the frost on the spatial distribution of agricultural employment, I estimate a linear model of the association between log changes in agricultural employment between two different periods and the density of coffee trees before the 1975 frost. The unit of observation is a municipality.[7] The effect of capital destruction is identified from the difference between the coefficient on coffee tree density in Paraná and the coefficient in a control state. I estimate the following model by ordinary least squares:

$$\Delta y_{i,t,s} \equiv y_{i,t} - y_{i,s} = \beta_{t,s} Coffee_i \times PR_i + \alpha_{t,s} Coffee_i + \gamma_{t,s} PR_i + \eta_{t,s} + u_{i,t,s} \tag{1}$$

where Δy_{ist} is the change, in municipality i, of the dependent variable of interest (say, the logarithm of employment in agriculture) from year s until year t. $Coffee_i$ is a measure of coffee tree density before the frost; in the baseline model, it is the inverse hyperbolic sine of the ratio of harvested coffee trees to total area in 1975.[8] PR_i is a dummy equal to one if the municipality is in Paraná and zero if it is in the control state. In this way, $\beta_{t,s}$ measures how different is the relation between coffee tree density and $\Delta y_{t,s}$ in Paraná relative to the control state. Note that the coefficients $\beta_{t,s}$ and $\alpha_{t,s}$ depend on the period examined. The equation includes a time-period fixed effect η_{st} and a time-period coefficient $\gamma_{t,s}$ on PR_i that accounts for statewide changes over time.

The baseline model does not contain control variables. Some robustness exercises, however, include municipality characteristics as control variables. Whenever included, the control variables are interacted with the Paraná dummy and the coefficients are specific to the initial and final time periods s and t.

In the baseline specification, I weight observations by municipality area, so the estimated coefficients may be interpreted as the effects per squared kilometer. The weighting makes it easier to interpret the results as the effects on the spatial distribution of farm employment. Spatial correlation of the errors is accounted for by clustering at the micro-region, a classification by IBGE that groups interconnected municipalities. The metropolitan areas around the state capitals were excluded from the sample. Their exclusion is justified by the fact that they are mostly urban regions, with little agriculture. In Section 5, I show the results are robust to different weighting, alternative inference procedures, or the introduction of the metropolitan regions to the sample.

[7]Due to border changes over time, I aggregate municipalities at minimum comparable areas according to 1970 borders; see Reis *et al* (2011). For simplicity, I refer to minimum comparable areas as municipalities.

[8]The main findings are robust to alternative measures of coffee density; see Section 5.

3.2 Identification Strategy

The identification strategy in this paper interprets the coefficient $\beta_{t,s}$ as a measure of the effects of the coffee capital destruction caused by the frost. Remember $\beta_{t,s}$ is the coefficient on the invesre hyperbolic sine of coffee tree density interacted with the Paraná dummy. Three facts motivate a causal interpretation for the coefficient.

First, the frost losses were concentrated in the coffee sector. In part, this is due to the importance of the sector for the Paraná economy. In 1970, 15% of total employment and 24% of total agricultural employment in Paraná where in coffee. Coffee was also the most important crop by revenue: coffee revenues in 1975 were 58% larger than soybeans revenues, the second crop in the revenue ranking. Moreover, coffee was the economic activity mostly affected by the frost. The frost occurred in July of 1975, at a time of the season in which many temporary crops were not in the fields, so its effects were concentrated on perennial plants. Since there was no other relevant perennial crop in Paraná, the crop damages brought by the frost were concentrated on coffee trees.[9]

Second, the frost resulted in damages to the coffee trees that were somewhat uniform across the state of Paraná, essentially affecting all coffee trees in the state (Oliveira 1978). Therefore, the coffee tree density at the time of the frost is a measure of coffee capital losses per hectare of farmland. Note that, since the frost occurred right after the coffee harvest of 1975, the number of harvested trees in that year was not affected by it.

Third, as discussed in Section 2, the effects of the frost were stronger in Paraná than in the other two large coffee producing states in Brazil. In this way, coffee tree density in Paraná was associated with more coffee capital losses than in the control state. For the baseline empirical exercise, I use the neighboring state of São Paulo as the control state. São Paulo was the second-largest coffee producing state in 1975. The frost in São Paulo was much milder than in Paraná: São Paulo coffee production and harvested area in 1976 were, respectively, 25% and 59% of their 1975 levels, while in Paraná they were only 0.37% and 0.49%. By 1977, coffee production in São Paulo was already above its pre-frost level, while it was only 17% in Paraná. São Paulo is a conservative choice for the control state, given it was partially affected by the frost. An alternative would be to use Minas Gerais, the third-largest coffee producing state. Minas Gerais was mostly spared from the frost, and coffee production in the state actually increased from 1975 to 1976. Robustness exercises presented in Section 5 show that the main findings hold when Minas Gerais is used as the control state.

The identification assumption is that, in the absence of the frost, the difference in the trajectories of more coffee dense municipalities relative to less coffee dense municipalities would be the same in Paraná as in the control

[9] The second and third largest perennial crops were citrus and bananas. Their combined revenues were less than 1% of the revenues in coffee.

state. Note that the identification assumption resembles the assumption of equal differences in trends used in a triple-differences identification strategy. As support for the plausibility of this assumption, I show that the relative trajectories were similar in Paraná and São Paulo before the 1975 frost. Most importantly, I show in Section 4 that the coefficient $\beta_{t,s}$ in the period immediately preceding the frost, from 1970 to 1975, is a precisely estimated zero.

The identification strategy assumes that, in the absence of the frost, any different trajectories of municipalities with different coffee tree densities would have been the result of economic shocks and changes that differentially affected coffee dependent municipalities. These shocks include changes in the coffee sector, such as higher coffee prices or the introduction of new technologies in coffee. A useful sanity check for the identification strategy is then to examine whether the coffee industries of Paraná and São Paulo seem to change similarly and be affected by the same events in periods other than the immediately after the frost.

A comparison between the trajectories of the coffee industry in São Paulo and Paraná, shown in Figure 3, suggests this is the case. The 1960s were a bad decade for coffee production, and, in both states, there was a reduction in harvested area.[10] From the mid-1960s until 1975, the harvested area in São Paulo and Paraná followed a similar downward trend. The main difference in the trajectories occur precisely in 1975. Relative to São Paulo, the harvested area in Paraná persistently declined after the frost. But since the late 1970s, the trends of São Paulo and Paraná are again similar. These similar trends suggest that the coffee industries of São Paulo and Paraná were similarly affected by shocks to the sector, such as the collapse of the International Coffee Agreement in 1989.

Finally, it is important to point out that the credibility of the identification assumption depends on the time frame examined. The larger the time interval examined, the less credible is the assumption that relative trends associated with coffee tree density in 1975 woud be the same in both Paraná and São Paulo in the absence of the frost. Such loss of credibility occurs because local economies are affected by a larger number of shocks over a longer time frame. For this reason, most of my analysis focuses on changes that occurred during the 1970s and 1980s. In this way, the findings mostly refer to the period before the large changes that afffected the Brazilian economy since the 1990s, such as trade liberalization, widespread reforms and deregulation, and macroeconomic stability. Still, Table 4 presents estimates associated with long-run changes which suggest the effects on farm employment persist up to 2006.

[10] Part of this decline was due to a government policy that financed the replacement of inefficient plantations for more modern ones; see Kohlhepp *et al* (2014) and Carvalho (1991) for a detailed description of this policy.

3.3 Data

The main data sources for this paper are the *Censos Agropecuários* (agricultural censuses) collected by the *Instituto Brasileiro de Geografia e Estatística* (IBGE), the Brazilian national institute of statistics. The information from the agricultural censuses is released at the municipality level. I use information from the 1960, 1970, 1975, 1980, 1985, 1995, and 2006. For the earlier censuses (up to 1985), the variables of interest were digitized from the tables released by IBGE.

Additional data, specially with respect to non-agricultural activities, are from the *Censos Demográficos* (population censuses) of 1970, 1980, and 1991. From these censuses, I use information on employment, earnings, and migration. This information is from a representative sample of the Brazilian population that is selected to answer an extended questionnaire in the census. I also use the *Produção Agrícola Municipal* (PAM), a survey conducted since 1974 that contains municipality-level yearly series of production, revenue, and harvested area for many different crops.

Further details on the data sources and preparation is available in Online Appendix B.6.

Table 2 displays descriptive statistics on the pre-frost economic conditions in the sample municipalities. The average labor-to-land ratio in 1970 was 0.1 worker per hectare. The economy was predominantly based on agriculture, as the average share of agricultural employment to total employment in 1970 was 62.5%. There were active land markets, as 1.3% of the total farmland value was transacted each year. On average, coffee tree density right before the frost was 33 trees per hectare.

4 Findings

4.1 Main Results

The main question I consider in this paper is how the frost changed the spatial distribution of agricultural employment and whether these changes were persistent. To do so, I estimate equation 1 when the dependent variable is the logarithm change in agricultural employment between two agricultural census. The estimated coefficients are shown on Table 3.

Column (1) reports coefficients for the change from 1970 to 1975. Although there was a relative decline in agricultural employment in coffee dense municipalities, such decline was similar in Paraná and São Paulo. In fact, the coefficient on the interaction between coffee tree density and the Paraná dummy is quite close to zero and statistically indistinguishable from it. That is, there are no differential pre-trends in the five years leading up to the frost. I interpret this result as a placebo test in support of the identification assumption that, absent the frost, differences in agricultural employment growth according to coffee density would be similar in both states.

Differential changes in farm employment appear immediately after 1975: column (2) shows that, from 1975 to 1980, farm employment declined more in Paraná municipalities with a higher coffee tree density. There was no such decline in São Paulo. The coefficient of -0.081 is large. Since the inverse hyperbolic sine approximates a logarithm, the estimates indicate that employment in agriculture fell by 1% for each increase of 12.3% in coffee tree density.

Moreover, the decline of farm employment seems persistent. Column (3) shows that, from 1980 to 1985, there was no faster employment growth in coffee municipalities. In fact, the employment changes seem to endure in these 10 years after the frost. Columns (4) and (5) of Table 3 show that, from 1970 to 1980 or 1985, coffee tree density in Paraná is associated with much larger declines of agricultural employment in Paraná than in São Paulo. The estimates then suggest a large, negative, and long-lasting impact of the frost on agricultural employment in the affected coffee producing areas of Paraná.

The agricultural censuses of 1970, 1975 and 1980 contain information on the share of agricultural employment that consisted of temporary farmer workers. Columns (6) and (7) present the coefficients when the dependent variables are changes in the logarithm of non-temporary agricultural employment. Note the estimates are similar to the results when using all farm labor, suggesting the effects were due to changes in year-long employment.

As further evidence for the persistence of the effects, Table 4 presents estimates of equation 1 when dependent variables are logarithm changes in agricultural employment over longer periods. Column (1) considers 1970-1995 changes and column (2) considers 1970-2005 changes. Note that the coefficient on the interaction between coffee tree density and Paraná are negative and larger than the coefficients for the 1970-1980 changes in Table 3. The results thus indicate that subsequent employment growth in affected municipalities was not faster than in control municipalities, suggesting persistence in the effects of the frost even 30 years after the event. Still, it is important to point out that, with a longer time frame, the identification assumption becomes less credible, so these results should be interpreted with care.

Changes over longer periods may also be used to further examine pre-trends. Columns (3) and (4) of Table 4 incorporates information from the 1960 agricultural census to estimate equation 1 when the dependent variable is the logarithm change in agricultural employment from 1960 to either 1970 or 1975. Note the reduced number of observations, due to the merge of municipalities in order to account for the municipality splits that occurred in the

1960s. Although positive and larger than for the 1970-1975 period, the estimated coefficients are not statistically different than zero. Such absence of pre-trends during the 1960 brings further credibility to the identification assumption.

4.2 Other Agricultural Outcomes

The empirical findings above indicate that the frost was associated with a decline of agricultural employment in the coffee producing municipalities of Paraná. I now examine its impacts on a variety of other economic variables. In this subsection, I focus on variables from the agricultural censuses. In the next subsection, I use information from the population censuses. The findings below are useful to distinguish between the explanations for why the frost had persistent effects on the spatial distribution of employment in agriculture.

Table 5 presents estimates of equation 1 when dependent variables are changes in agricultural outcomes. Column (1) presents the effects on the logarithm of the wage bill per worker. The denominator includes only paid farmer workers. This variable is a proxy for agricultural wages. Later on this section, I use population census data to construct a more precise measure of wages. The results from column (1) suggest that the frost did not affect wages in the short-run. The estimated coefficients on the interaction term for 1975-80 and 1970-80 are not statistically different than zero, while the coefficient for the 1970-75 changes are marginally significant at a level of 10%. The agricultural censuses of 1985 and 1995 did not collected detailed data on farm finances, so the wage bill per worker cannot be computed for those years. Hence, long-run effects are harder to examine. The changes from 1970 to 2005 suggest a negative long-run effect on agricultural wages.

Columns (2) and (3) present the effects of the frost on two variables that aim to measure the relative importance of labor in the agricultural production function. In column (2), I consider the logarithm of total non-labor and non-land expenditures per workers. This variable proxies for the use of capital and inputs such as seeds and fertilizers per unit of labor. In column (3), I consider the labor share in total non-land expenditures. In either case, there are no clear effects on these variables. I interpret these results as suggestive that the frost did not induce labor-saving technical change in agriculture.

Column (4) presents the effects of the frost on farmland prices. The estimates indicate a negative effect on land prices, that declined sharply between 1975 and 1980. The results for the changes from 1970 to 2006 suggest that the effects on land prices were persistent. These findings will be important to distinguish the hypotheses of multiple spatial equilibria and of changes in beliefs about frost risk from other explanations.

Finally, I also examine how the frost affected agricultural land use. To document these effects, I estimate equation 1 when the dependent variable is the change, from 1974 to any subsequent year, in the harvested area (as a share of total farm area) of either coffee, maize, beans, soybeans, or wheat.[11] In 1975, these were the five most common crops in Paraná according to harvest area. I also estimate the effects on the cattle stock density. Results are shown in Figure 4. As expected, there is a large, persistent, and precisely estimated negative effect on the harvested area of coffee. Although noisy, the estimated coefficients for the other crops indicate that there was no differential change in the harvested area of soybeans, a growth of wheat, and a decline of the harvested area of maize.[12] The results also suggest a positive effect on the cattle stock.

4.3 Other Labor Market Outcomes

The results discussed so far show that the frost resulted in a large and persistent displacement of workers from agriculture. Where did these workers go? They could have reallocated, within the same municipality, to sectors other than agriculture. In other settings, sectoral reallocation has been observed as a response to weather shocks (Emerick 2018). It does not seem to be an important margin of adjustment here. Columns (1) and (2) in Table 6 estimate the effects on employment in non-agriculture. Estimates of total municipality employment in non-agriculture are available from the population censuses of 1970, 1980, and 1991. The dependent variables in columns (1) and (2) are, respectively, the logarithm in non-agricultural employment from 1970 to 1980 and from 1970 to 1991. The coefficients on the interaction term are negative and not statistically significant, suggesting that reallocation to the local non-agricultural sector was not the main destination for farm workers displaced by the frost.

Instead, affected individuals reacted by seeking employment in other municipalities. In columns (3) to (5) of Table 6, I estimate the equation 1 with migration rates as the dependent variables. Migration rates refer to the intercensal period of 1970 to 1980 and are calculated for individuals who were between 20 and 59 years old in 1975. See Appendix B.6 for details on the construction of migration rates. It is important to point out that Paraná experienced significant out-migration in that decade: 19% of the 1970 population in this age group migrated from the state, the highest rate among all 25 Brazilian states.[13] In fact, column (3) suggests strong negative effects on net-migration rates. The results are not driven by in-migration rates, as the small coefficients in column (4) show. Instead, they seem due to out-migration rates; see column (5).

The null effects on in-migration are relevant to interpret the negative coefficients on farm employment. As

[11] 1974 is the first year through which the series is available.
[12] Maize was commonly grown by coffee farms.
[13] The state net-migration rate was -12%, the second lowest (most negative) among Brazilian states.

the frost led to large out-migration flows, some sample municipalities could have suffered a labor supply shock if these migrants disproportionately selected into them. The coefficients in column (4) of Table 6, suggest that neither non-coffee municipalities in Paraná nor coffee municipalities in São Paulo disproportionately received migrants in the period. Therefore, the effects on employment in agriculture shown in Table 3 are unlikely to be due to labor supply shocks affecting control municipalities. They are plausibly explained by a relative decline of farm labor demand in affected municipalities.

Both quantities and prices in local labor markets may have been affected by the frost, so now I turn to the effects on wages. There is little information on wages in the agricultural censuses; in Table 5, I used the wage bill per worker as a proxy for agricultural wages. I now complement these results with data from the population census, which contains better information on worker earnings. However, the census information also has its limitations. First, the 1970 census does not include information on hours worked in agriculture, so I do not observe wages. Hence, I use changes in average earnings per worker as the dependent variable. Second, the 1970 census also does not distinguish from labor earnings and earnings from other sources. To make sure the results reflect the effects on wages and not on working hours, and the effects on wages and not on non-labor income, I estimate the effects on the average total earnings of different groups of workers.

Table 7 shows the estimates of the effects on average earnings per worker. In column (1), the dependent variable is the change in the logarithm of average earnings among all workers from 1970 to 1980 (Panel A) or from 1970 to 1991 (Panel B). Column (2) restricts average earnings to men. Under the assumption that the labor supply from men is less elastic than the labor supply from women, the coefficients on column (2) are more likely to represent the changes in wages than in hours. The coefficients suggest that the frost did not affect wages from 1970 to 1980. The coefficients for the 1970-1991 period suggest that average earnings across workers in all sectors increased in the long-run.

Columns (3) and (4) of Table 7 consider average earnings only for workers employed in agriculture. In this case, the estimated coefficients for the interaction term is never statistically different than zero, so the null hypothesis that the frost did not affect earnings in agriculture cannot be rejected.

Some workers in the sample are farm owners, so their earnings will include not only wages, but also the returns to land and other productive assets. In this case, the coefficients from columns (1) to (4) may not isolate the effects on wages. To address this issue, I calculate average earnings only for agricultural workers who declared to be farm employees or daily laborers. The results are reported in columns (5) to (6) of Table 7. As before, the estimated coefficients show no evidence of an effect of the frost on wages from 1970 to 1980. When the dependent variable is the log change in earnings for this subpopulation from 1970 to 1991, the coefficients on the interaction between coffee tree density and Paraná are now negative, although not statistically different than zero.

Summing up the empirical findings, the frost persistently reduced employment in agriculture, displacing workers toward other municipalities instead of toward other activities within the same municipality. Local labor markets seem to have adjusted on quantities, not wages: the effects on average earnings are less clear and smaller than the effects on agricultural employment. Land prices seem to have been negatively affected by the frost too.

5 Robustness

The findings about the effects on the spatial distribution of employment in agriculture are robust to a variety of alternative specifications, samples, and inference methods. This section presents such robustness exercises.

Alternative control state. The baseline empirical exercise uses São Paulo, the second-largest coffee producing state in 1975, as the control state. Column (1) of Panel A of Table 8 shows the estimated coefficients when the control state is Minas Gerais, the third-largest coffee producing state. The coefficients are qualitatively the same, although the differences between the coefficients in Paraná and Minas Gerais are usually larger than the differences between the coefficients in Paraná and São Paulo.[14]

Alternative sample. For the baseline exercise, the metropolitan regions around the state capitals were excluded from the sample. Column (2) of Panel A of Table 8 present the estimates when municipalities in these regions are included into the sample. The coefficients are similar to the baseline ones.

Control for other pre-frost conditions. A concern about identification is that changes in agricultural employment could be due to initial conditions other than coffee tree density. I examine this question by including a variety of initial conditions as control variables; their coefficients are state and time-period specific. First, since Michaels, Rauch and Redding (2012) have shown that the logarithm of population density in 1970 is positively associated with subsequent municipality growth, I control for it; see column (3) of Panel A of Table 8. Second, I account for threats associated with structural transformation by including the 1970 share of agriculture in total employment as a control variable; see column (4). Third, I account for the dynamic adjustment of local labor markets by including the logarithm of average earnings in agriculture in 1970; see column (5). Fourth, I account for possible damages to pastures by including the logarithm of cattle density in 1974 as a control variable; see column (6). The coefficients

[14] There are some reasons to expect Minas Gerais not to be the most appropriate choice for a control. Other factors, such as technological improvements that allowed intensive agriculture in the Cerrado region, led to the future growth of the coffee industry in the state.

shown on column (7) of Panel A of Table 8 are from a specification that includes all the controls above. The results are robust to the inclusion of the control variables.

Control for pre-trends in agricultural employment. The main findings are also robust to the inclusion of pre-frost changes in the logarithm of agricultural employment as a control variable. As above, the coefficient on the control variable is state and time-period specific. The rationale in the inclusion of pre-frost changes is that they may account for persistent differences in the growth potential of the municipalities. Column (1) of Panel B of Table 8 includes changes from 1970 to 1975; column (2) includes changes from 1960 to 1970.[15]

Alternative measures of coffee specialization. Another concern is that the inverse hyperbolic sine of coffee tree density, the measure of the importance of coffee production used so far, is not the most adequate one to capture the capital losses from the frost. In column (2) of Panel B of Table 8, I use a dummy variable equal to one if coffee tree density is above the sample median (6.022 trees/ha). In column (4), I use instead the level of the density of harvested coffee trees in 1975, measured in 100 trees per hectare. In both cases, the coefficients show the same patterns. In column (5), I use the inverse hyperbolic sine of the ratio of harvested coffee trees in total farm workers in 1975 as the measure for dependence on coffee. Notice it is a measure of the capital-labor ratio, while before I used the capital-land ratio. To analyze the effects on labor markets, the capital-labor ratio seems a better measure. However, labor is a variable input that is adjusted in response to temporary productivity shocks, which potentially adds measurement error to the variable. Reassuringly, results point out in the same direction.

Alternative weighting. The main results are estimated by assigning to each municipality a weight equal to its total area. Panel B of Table 8 shows the main results are robust to the use of alternative weightings of the observations. In column (6), I weight all municipalities equally. In column (7), I weight them by total agricultural employment in 1975.

Spatial correlation of the errors. In Table 9, I show that the results are robust to different inference procedures. Instead of clustering standard errors at the micro-region, I consider two alternative estimates for the standard errors. First, I consider robust standard errors that are robust to heteroskedasticity but not to any cross-sectional dependence of the errors. The standard errors are slightly lower in this case. Second, I estimate robust standard errors that allow spatial correlation of the errors of municipalities that are within a cutoff distance of each other as in Conley (1999). I report standard errors when the cutoff is 100 km, 200 km, 300 km, 400 km, or 500 km. The standard errors do not increase as the cutoff increases, which suggest that all relevant spatial correlation of the errors take place within a radius of 100 km. Since it is possible that, as the cutoff increases, overrejection becomes more frequent

[15] In this case, the units of observation are minimal comparable areas since 1960, as in columns (3) and (4) of Table 4.

as the distribution of the test statistic becomes farther than its asymptotic distribution, I also report simulated rejection probabilities as proposed by Ferman (2019). Ferman (2019) suggests a simple procedure to assess such over rejection: assuming the null hypothesis hold, simulate the probability the test is rejected when errors are homoscedastic. I simulate errors drawn from a standard normal distribution. The rejection probabilities for a test of size 5% are shown on column (5) of Table 9. Note that, up to 300 km, there is no substantial over rejection. This is no longer the case when the cutoff is 400 km or higher. The simulations also show no over rejection when robust standard errors are calculated with micro-regions as clusters, which is my preferred inference method.

6 Explanations

In this section, I examine the plausibility of alternative explanations for the persistent effects of the frost. Two explanations seem fully consistent with the empirical findings. The first explanation is that the frost moved the economy to a different spatial equilibrium. The second explanation is that the frost persistently changed farmers' beliefs about frost risk. At the end of the section, I discuss other hypotheses and argue that they are not fully consistent with the data or the historical evidence.

Before presenting the explanations, it is important to note that the two plausible hypotheses explain the persistent decline of farm employment as a consequent of the persistent decline of the farm industry. The claim is that the decline of coffee led to a contraction of labor demand in agriculture. This is a reasonable assumption since, relative to alternative farm activities, such as cattle ranching or temporary crops, coffee production uses a higher labor-land ratio. Before the frost, coffee occupied 9% of the Paraná farmland, but it represented 24% of agricultural employment. With information from the 1975 agricultural census, I show in Online Appendix B.2 that, in a regression of total farm workers on the harvested area of the main crops, coffee has one of the highest coefficients. Only beans used as much labor, while soybeans and wheat were associated with less than half of the labor employed in coffee.

6.1 Multiple Spatial Equilibria

A central conclusion of many spatial economics models is that natural endowments and technology may not uniquely determine the distribution of economic activity across space; see Proost and Thisse (2019) for a recent survey of these models. Such indeterminacy occurs when agglomeration economies, which are positive spillovers from geographic concentration, are sufficiently strong to overcome dispersion forces. If this is the case, a large shock may persistently change the spatial distribution of economic activity by shifting the economy towards a different spatial equilibrium. Could this be the case with respect to the spatial distribution of agriculture after the frost?

Theoretical framework. Appendix A.1 contains a stylized model that applies this explanation to the case of the frost. The model considers a small open economy in which local farmers decide which crop to grow and how many workers to hire. I briefly explain the intuition, which is based on two contrasting forces, in what follows. First, when there are agglomeration economies within the coffee industry, crop choices may exhibit strategic complementarities, as the benefits of producing coffee are larger when neighboring farmers are also coffee producers. If agglomeration economies within the coffee industry are sufficiently strong, then there are two steady-states that differ on how many farmers produce coffee. Second, when a farmer decides to produce coffee, she invests in coffee trees, a persistent type of capital. By doing so, she creates an incentive to keep producing coffee until the coffee trees depreciate. Due to the agglomeration economies, this incentive may indirectly extend to neighboring farmer.

These two forces give rise to a history versus expectations situation, as in the seminal models of Krugman (1991) and Matsuyama (1991). Technological parameters determine whether the long-run prevalence of each crop depends on initial conditions or on self-fulfilling expectations. Note that the crop choice of a farmer depends on her initial stock of coffee trees, as farmers who inherit productive trees from the previous period might prefer to remain producing coffee; they are locked in coffee. If coffee trees depreciate only slowly, an economy that starts specialized in coffee will remain so, as the presence of locked-in farmers makes coffee production optimal (due to agglomeration economies) also for those who had their trees depreciated. History has thus determined the equilibrium. But when a severe frost hits and destroys many of the trees, then initial conditions no longer matter and, depending on farmers' expectations, the economy converges to an equilibrium in which there is no coffee production.

Testable implications on land prices. Despite its simplicity, the theoretical framework above provides an additional testable implication: the frost reduces land prices only if there are agglomeration economies within the coffee industry. Think of land prices as the value of land only, excluding any investments on it, such as coffee trees. This is precisely the information that is available in the agricultural censuses. If land markets are liquid, as they were in Paraná, then land prices are determined by the present value of discounted profits on a land parcel.[16]

[16] Assunção (2008) argues that macroeconomic uncertainty leads to an additional demand for land as a safe asset, which could prevent land to be priced according to the present value of its earnings stream. Macroeconomic uncertainty in Brazil was indeed high from the mid-1980s to the mid-1990s, but lower in the 1970s and in the 2000s, the periods for which I have land price information. Furthermore, it is not clear that (and how) the non-agricultural component of land demand would be affected by the frost.

Land prices are thus forward-looking and depend on the optimal land use. The frost does not directly impact the future earnings potential of the farmland, so it should not affect land prices directly. However, when there are agglomeration economies and the frost shifts the economy towards a lower level of coffee production, then the frost indirectly reduces land prices as the decline in the coffee industry negatively affects profits through agglomeration economies. As the shift away from coffee was persistent, so should be the negative effects on land prices. In fact, this is precisely what is seen in the data; see column (4) of Table 5.

Agglomeration economies. The explanation illustrated by the theoretical framework depends on agglomeration economies within the coffee industry. Agglomeration economies are the benefits of locating near other agents. Hence, there are agglomeration economies within the coffee industry when coffee farmers benefit from proximity to other coffee farmers. That is, they are external economies of scale caused by the spatial concentration of the activity of the industry (Rosenthal and Strange 2004).

The existence of agglomeration economies within any individual industry is an empirical question. It is a notoriously difficult one (Combes and Gobillon 2015). In the rest of this subsection, I first review some qualitative evidence for agglomeration economies within the coffee industry. I then present an empirical exercise that provides *suggestive* evidence of their existence.

Two qualitative studies document agglomeration economies within the coffee industry. Douthat (2017) examines how agglomeration economies mattered for regional resilience during the coffee crisis. Pereira and das Chagas Ribeiro (2015) document the benefits of cooperation in two coffee producing districts in Brazil; they point out the importance of the sharing of equipment and the exchange of labor during the harvest season.[17] Beyond equipment and labor sharing, there are three other reasons for the existence of agglomeration economies within the coffee sector. First, Pozzobon (2006) and Kohlhepp *et al* (2014), main references on the Paraná coffee industry before the frost, documented the dependence of coffee production on trading, milling, and specialized services. Since most coffee farmers were small, these activities were usually performed by other agents who could generate scale by serving many farmers. As an example, consider milling: in 1975, only 5.14% of the Brazilian coffee was milled in the farm it was produced.[18] Some milling equipment was mobile and taken to farms; often, neighboring farmers would coordinate to hire in-farm milling services (Carvalho 1991). Second, when specialized labor is demanded, industry agglomeration can result in thicker labor markets and lower search costs (Duranton and Puga 2004). Pozzobon (2006) and Carvalho (1991) describe an example: new plantations were usually taken care of by specialized

[17] Carvalho (1991) also documents the importance of equipment sharing, as some farmers in her sample would lease agricultural machinery to neighboring farmers.

[18] Author's calculation based on the 1975 agricultural census.

workers who were more experienced in the care of young trees. Third, a concentration of farmers that specialize in the same activity allows the diffusion of new techniques and the discussion of market conditions. That is, it eases the flows of knowledge and information. Zylbersztajn et al (1993) describe this channel, explaining how less information is transmitted when the coffee farmer is distant from firms in the forward-linked sectors. Consistently with these reasons, the low production volume of the Paraná coffee industry today has been indicated as a barrier to higher profits and industry growth (Folha de Londrina 2015). These barriers could operate through low competition in forward-linked sectors, which, according to a diagnosis by the state government, prevents the growth of the Paraná coffee industry (Valor Econômico 2012).

As empirical evidence for agglomeration economies within the coffee industry, I examine whether a farmer is more likely to plant coffee trees when neighboring farmers are coffee producers. A difficulty in this investigation is that similar crop choices by neighboring farmers could be due to some non-observable characteristic of their farms. To address this identification challenge, I consider a salient feature of the terrain that is prejudicial to coffee production in Paraná. It is well-known that coffee production should be avoided on south-facing terrains (Caramori et al. 2007, Barros et al. 2007, Sediyama et al 2001): they are more exposed to frosts caused by cold winds and, since they receive less sunlight, tend to be more humid and thus vulnerable to rust leaf. In fact, there is a negative relationship between whether a terrain faces south and its coffee tree coverage. I show that, conditional on the cardinal orientation of a site, there is also a negative relationship between coffee tree coverage and the share of *neighboring* sites that face south.[19] I interpret this second negative relationship as suggestive of a positive spillover from nearby coffee farmers.

The exercise requires high-resolution information on the spatial distribution of coffee trees across the state of Paraná. I use a mapping of coffee coverage done through remote sensing by the *Companhia Nacional de Abastecimento* (Conab), an institution of the Brazilian Ministry of Agriculture, in 2017. I divide the state of Paraná into a grid of 30 arc-seconds of latitude by 30 arc-seconds of longitude.[20] I select the sample to exclude urban areas and sites that are unsuitable for coffee production; see Appendix B.4 for an explanation of the data and sample selection. For each cell, I calculate the share that is facing south and the share that is covered by coffee trees. A terrain faces south if its altitude declines in a cardinal direction that is to the south. I then estimate the following equation by ordinary least squares:

[19] The identification strategy used to identify agglomeration or density economies in agriculture by Holmes and Lee (2012) and Richards (2017) is similar. Robalino and Pfaff (2012) use this idea to estimate spillovers in deforesting.

[20] The average area of a grid cell is 79 ha. As a comparison, the median coffee farm in Paraná in 2017 had an area of 11.7 ha.

$$Coffee_i = \alpha_0 South_i + \alpha_1 \overline{South}_{N(i)} + X_i'\theta + u_i \quad (2)$$

where $Coffee_i$ is the share of the cell covered by coffee, $South_i$ is the share that faces south, and $\overline{South}_{N(i)}$ is the average south-facing share of neighboring cells to i. I define neighboring cells as those that are at a distance of up to 5 km of i.[21] All observations are weighted by area. To account for spatial correlation of the errors, standard errors are calculated with multi-way clustering (Cameron *et al.* 2011) according to four interlaced 0.5 × 0.5 degree grids; see Online Appendix B.4 for details.

Results are shown in Table 10. Column (1) documents the negative correlation between the share of a grid cell that faces south and coffee coverage: a 1 p.p. increase in the share that faces south reduces coffee coverage by 0.0044 p.p., a large effect given that coffee covers only 0.35% of the sample area. In column (2), I include also the south-facing share of neighboring cells. By doing so, the coefficient on a cell own share facing south falls by a third, while there is a strong and significant relationship with cardinal direction in neighboring cells.

A first concern regarding identification is the presence of an omitted variable bias, such as differences in micro-climates. To partially address it, I add a variety of controls in column (3). I control for a cell own and neighboring averages of altitude, slope, proximity to water (proxied by the share that is within 500m of a river), and the presence of *Terra Roxa*, a fertile red soil that is considered good for coffee production. I also control for the distance to the nearest road and for a quadratic polynomial of latitude. The controls substantially increase the prediction power of the model, but the coefficients of interest are only slightly affected by the control variables. Online Appendix B.4 contains further robustness exercises.

A second concern is that the coefficient on neighboring cells could be identifying within-farm density economies instead of between-farms agglomeration economies; this would be the case when farms span neighboring grid cells. As evidence that the results reflect agglomeration economies, I estimate equation 2 with farms as the units of observation. I use farm limits from the *Cadastro Ambiental Rural*, a land-use farm registry in the Ministry of Environment. The coefficients are shown in column (4). They should be read with caution, as farm sizes and shapes are endogenous to crop choice, but they are in line with the baseline coefficients from column (3).

[21] Note that there are different geographic scales through which agglomeration economies operate. When a group of farmers shares coffee processing equipment, for example, then this benefit of agglomeration decays quickly with distance. On the other hand, some agglomeration economies operate at the scale of large regions. For example, regional institutions of agricultural research and extension provide a public good that benefits a large area. As the information produced and provided by these institutions is endogenous to which crops the farmers grow in the region, they could lead to agglomeration economies that operate at the regional level. The evidence of agglomeration economies I show is at a small geographic scale. Nevertheless, it is important to keep in mind that some agglomeration economies could operate at a large scale, so my results do not account for all agglomeration economies possibly at play.

Finally, in column (5) I use the south-facing share of neighboring cells as an instrumental variable to estimate the effect of neighboring coffee tree coverage on coffee plantation decisions. The two-stage least squares estimate indicates that a 1 p.p. increase in coffee in the neighboring area increases coffee coverage on a cell by roughly that same amount. Since the first stage is weak, as shown by a robust F-statistic of 5.51 on the excluded instrument, I perform inference by inverting the Anderson-Rubin statistic, as suggested by Andrews, Stock and Sun (2019). The 95% confidence interval is narrow, so we can reject the null hypothesis that neighboring coffee production has no effect on coffee production decisions. The OLS estimate is larger than the 2SLS estimate, as expected when an unobservable variable equally affects coffee coverage in neighboring sites, although it is slightly lower than the upper bound of the 95% Anderson-Rubin confidence interval.

6.2 Updating of Frost Risk Perceptions

A potential explanation for the decline of the coffee industry is that the frost led farmers to believe in a higher frost risk. Previous research has already shown that economic agents often underestimate disaster risk and that the occurrence of a disaster increase risk perceptions; see Brown *et al* (2018). Given the lack of detailed information on farmers' beliefs, there is no direct way of assessing this hypothesis. Still, there is some anecdotal evidence of farmers being scared by the frost; see Mores (2017) and TV Coroados (1980). In this subsection, I discuss to which extent this hypothesis is consistent with the empirical findings and the historical evidence.

Note the hypothesis is consistent with the negative effects of the frost on land prices observed in column (4) of Table 5. In fact, if a farmer updates her beliefs to expect frosts to be more frequent, than the expected profis of planting coffee decrease. If, before the frost, coffee was the optimal crop choice for the farmer, then the discounted expected value of the land to her goes down as well.

However, an update of frost risk is at odds with some of the historical evidence. The 1975 frost was a very extreme event: GERCA (1975) suggests that the only two comparable frosts occurred in 1902 and 1918; no frost like that has happened ever since. The opinion of agronomists and other specialists, as Mores (2017) argues in his environmental history of the region, was that such unlikely event should have no effect on the decisions of planting trees that depreciate within 20 years. As written in a booklet the government distributed to farmers: "the 1975 frosts did not affect the evaluation of weather risks in the affected zones. Frosts like these are exceptional and extremely rare" (GERCA 1975). Indeed, as late as 1978 there were observers who expected the coffee industry to recover on this basis (Oliveira 1978).

It is also at odds with some of the empirical findings on how individuals update disaster risk after a shock.

For instance, Gallagher (2014) shows that floods lead to increased demand for flood insurance, despite recent floods providing barely no new information about flood risk. However, the higher risk perceptions often tend to fall over time. Similar findings have been documented with different methodologies and in different settings; see, for instance, Bin and Landry (2013) and McCoy and Walsh (2018). If this was also the case after the frost, then the strong declines in the coffee industry, farm employment, and land prices should be followed by a subsequent recovery. The empirical results indicate no such recovery.

Finally, it is important to point out that the two explanations might not necessarily be seen as alternatives. They two hypotheses are likely to be complementary. In this case, pessimism about climate risk could have negatively affected the recovery of the coffee industry in the short run, effectively coordinating farmers' expectations toward a non-coffee steady-state.

6.3 Alternative Explanations

There are other alternative hypotheses for why the frost persistently affected the spatial distribution of employment in agriculture. In what follows, I discuss whether these hypotheses could explain the effects of the frost *in the absence of* multiple spatial equilibria or the updating of frost risk perceptions. I do not find compelling evidence in support of the following explanations:

Faster transition to a situation with less coffee. The decision between coffee and other land uses depends on market conditions and available technologies that change over time. In this way, it is possible that, in the 1970s, coffee was no longer the most profitable alternative for farmers who had their trees depreciated. In the absence of the frost, convergence would be gradual, as coffee trees depreciate slowly. The frost accelerates the conversion of coffee farms to other uses, reducing the demand for labor.[22] Two pieces of evidence suggest this hypothesis is unlikely to explain the effects of the frosts. First, the investment in new trees shown in Table 1 suggests a stable coffee industry before the frost. Second, it does not account for the negative effects on land prices shown on Table 5. In this case, land prices would be determined by the discounted profits of non-coffee activities, which were not affected by the frost.

Government policy changes. At the time of the frost, there was still substantial government intervention in the coffee sector. Since 1952, the Brazilian coffee market was regulated by the *Instituto Brasileiro do Café* (Brazilian Coffee Institute, IBC), which had many policy instruments to act on coffee markets. In particular, the IBC could

[22]In terms of the model, this situation can be expressed by assuming an initial share of coffee farmers, γ_0, that is higher than the unique steady-state.

regulate prices by setting minimum prices, buying and stocking coffee, and setting the *quota de contribuição*: an export tariff that allowed domestic coffee prices to fluctuate differently than international ones. Furthermore, the IBC also operated warehouses where coffee beans were received and stocked, provided technical assistance to farmers, and defined subsidies for the planting of coffee trees. Therefore, it is important to carefully understand the policies implemented after the frost. If the government responded by reducing the incentives for coffee production in Paraná, then government policy could explain the effects in farm employment and land prices. In what follows, I briefly review the policy response and argue this was not the case.[23] Guaranteed coffee prices were increased by a July 31 resolution of the IBC. From the frost until August 1, the IBC imposed a ban on new export contracts. In August 1975, the *Conselho Monetário Nacional* (CMN), Brazilian monetary and financial council, approved a new credit line with the purpose of recovering ruined coffee farms. This credit line was called the Emergency Plan for the Recovery of Frosted Coffee Plantations (*Plano Emergencial para Recuperação de Cafezais Geados*).[24] A careful analysis of the plan shows that the credit policy for the coffee sector, instead of being channeled towards farmers and regions not affected by the frost, actually benefited impacted farmers by granting them better conditions. For example, if a farmer had to prune his trees as a result of the frost, then he faced less restrictive collateral requirements and was exempted from showing proof of previously high productivity when obtaining a loan to plant coffee.[25] To summarize, there is no indication that the government policy for the coffee sector acted with the purpose of reducing coffee production in Paraná while incentivizing other areas; the government acted instead with the goal of recovering the affected areas. My conclusion is the same as the perception of the *Estado de São Paulo* correspondent in Paraná, who wrote in 1978: "Good judgment and the acknowledgment of northern Paraná's (coffee growing) potential predominated, and the coffee activity was incentivized in the region" (Oliveira 1978; p.71).

Labor-saving technical change. The previous hypotheses assume that labor demand from agriculture declined as a result of the decline in coffee production. However, labor demand shifts down either if land use patterns change, so crops with high labor-land ratios are replaced by crops with low ratios, or if labor-saving farming technologies are adopted. As there were no increases in wages as a consequence of the frost, it is not clear which incentives farmers would have to adopt labor saving technologies (Hayami and Ruttan 1970). Still, the large shock could provide an opportunity for farmers to modernize. The findings from Table 5 allows me to empirically reject this

[23]The conclusions are drawn from a visit to the *Arquivo Nacional* in April 2018, when I carefully analyzed IBC documents. A summary of the post-frost policies is in Conjuntura Econômica (1975). I thank Elis Granado from the *Museu do Café* for the reference.

[24]The plan is described in the Carta-circular 141 of the *Banco Central do Brasil*.

[25]See also Banco Central do Brasil (1976).

hypothesis by examining the effects of the frost on the labor share of total expenditure and on the total expenditure per worker. If wages do not change and there is a labor-saving technology change, the labor share would go down while total expenditure per worker would go up. The estimated coefficients, shown in columns (2) and (3) of Table 5, reject such pattern. There is no evidence that the frost affected these variables. In this way, the most plausible explanation for the labor demand shift is the decline of coffee production.

Spatial misallocation of labor. In all the explanations above, labor demand by the farm sector persistently contracted after the frost. An alternative explanation is that local labor supply persistently contracted instead. This contraction could occur if there was misallocation in the spatial distribution of labor. Suppose that there were many workers and low wages in the coffee producing region of Paraná, but these workers were unwilling to move out due to migration costs. The frost, by depressing short-run work opportunities, could lead many of the workers to migrate, therefore reducing the preexisting wage gap between origin and destination regions. However, this hypothesis is unlikely to be true. First, it predicts a long-run increase in wages, for which the results on earnings in Table 5 and Table 7 show no conclusive evidence. Second, the 1970s were a period of unprecedented labor mobility in Brazil. The national urbanization rate changed from 56% in 1970 to 68% in 1980, the fastest 10-year increase in Brazilian history. The average interstate out-migration rate in the decade was 8.1%, almost double what it was from 2000 to 2010 (4.5%).[26] Given this high mobility, it seems unlikely that labor was spatially misallocated. Third, the effects are robust to the inclusion of 1970 average earnings as a control variable; see column (5) of Panel A of Table 8. Hence, the decline of employment is not explained by initial wage differences.

7 Conclusion

Climate change has been associated with increased weather variability (Bathiany *et al* 2018), so a detailed understanding of how weather shocks affect economic activity is essential to understand the effects of climate change. No other sector is as vulnerable to weather shocks as agriculture, but there is little evidence on how the spatial distribution of agriculture is affected by such shocks. Of particular interest is whether the spatial distribution of farm employment is vulnerable to such shocks, which could then lead to significant weather-induced migration. In this paper, I documented the effects of a large weather shock, the 1975 frost, that damaged coffee trees in the Brazilian

[26] Bell et al. (2015) estimate measures of labor mobility for a variety of countries. Migration intensity in Brazil during the 2000s was less than half of what it was in the United States, less than in Paraguay or Bolivia, but more than Mexico and China, and almost four times higher than in India or Egypt. Since migration rates in the 1970s were 80% larger than in the 2000s, Brazilian labor mobility at the time of the frost would be comparable to what it is observed today in a high internal migration developing country such as Cameroon or Senegal.

state of Paraná. The empirical findings suggest that the frost led to a persistent decline of agricultural employment in affected areas. Such displacement of labor resulted in large out-migration flows from these areas. The decline of farm employment is associated with the decline of the labor-intensive coffee industry, which is plausibly explained by multiple spatial equilibria in agriculture or changes in beliefs about frost risk.

Previous literature has suggested that some, but not all, weather shocks affect the spatial distribution of labor (Bohra-Mishra *et al* 2014, Boustan *et al* 2020). Whether this occurs will depend on characteristics of the local economies affected by the shock. The case of the 1975 frost in Paraná illustrates some of these characteristics. First, the frost damaged coffee trees, a fixed and specific type of capital. Perennial plants like coffee trees account for 21% of global agricultural production and are a particularly weather-vulnerable type of capital.[27] Capital destruction is a particularly devastating way agriculture is affected by weather shocks. Moreover, as shown in the theoretical framework in Appendix A.1, capital may anchor the economy to a given activity, so the destruction of perennial plants could allow the spatial distribution of agriculture to change. Second, the effects on the spatial distribution of agricultural employment may be particularly important when the damages of the shock are relatively stronger on crops that require a larger labor-to-land ratio, such as coffee. Third, the effects on employment may persist over time when farming decisions permanently change after the shock. This is likely to be the case when agricultural activities in a region present significant agglomeration economies or when beliefs about climate risk could be affected by extreme weather events. In this case, policies that promote the coordination of farmers and that align weather risk perceptions may bring resilience to local economies.

[27] Author's calculation, based on the gross production value available from FAOSTAT for 2016.

References

[1] Ager, P., Eirksson, K., Hanse, C.W., & Lonstrup, L. (2020). How the 1906 San Francisco earthquake shaped economic activity in the American West. *Explorations in Economic History*, 77, 101342.

[2] Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: theory and practice. *Annual Review of Economics*, 11, 727-753.

[3] Assunção, J.J. (2008). Rural organization and land reform in Brazil: the role of nonagricultural benefits of landholding. *Economic Development and Cultural Change*, 56(4), 851-870.

[4] Banco Central do Brasil (1975). *Carta-circular 141*. Brasília, September 5 1975.

[5] Banco Central do Brasil (1976). *Carta-circular 188*. Brasília, July 30 1976.

[6] Barros, M.A., Moreira, M.A., & Rudorff, B.R.T. (2007). Dados interferométricos para modelagem topográfica e caracterização ambiental do café em escala municipal. *Sociedade e Natureza*, 19(2), 33-50.

[7] Bathiany, S., Dakos, V., Scheffer, M. & Lenton, T.M. (2018). Climate models predict increasing temperature variability in poor countries. *Science Advances*, 4.5: eaar5809.

[8] Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. *American Economic Journal: Applied Economics*, 9(2), 219-55.

[9] Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M., & Kupiszewska, D. (2015). Internal migration and development: comparing migration intensities around the world. *Population and Development Review*, 41(1), 33-58.

[10] Bin, O., & Landry, C. E. (2013). Changes in implicit flood risk premiums: empirical evidence from the housing market. *Journal of Environmental Economics and Management*, 65(3), 361-376.

[11] Bohra-Mishra, P., Oppenheimer, M., & Hsiang, S. M. (2014). Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proceedings of the National Academy of Sciences*, 111(27), 9780-9785.

[12] Boustan, L. P., Kahn, M. E., Rhode, P. W., & Yanguas, M. L. (2020). The effect of natural disasters on economic activity in US counties: A century of data. *Journal of Urban Economics*, 118, 103257.

[13] Brakman, S., Garretsen, H., & Schramm, M. (2004). The strategic bombing of German cities during World War II and its impact on city growth. *Journal of Economic Geography*, 4(2), 201-218.

[14] Brown, P., Daigneault, A. J., Tjernström, E., & Zou, W. (2018). Natural disasters, social protection, and risk perceptions. *World Development*, 104, 310-325.

[15] Cameron, A. C., Gelbach, J.B., & Miller, D. L. (2011). Robust inference with multi-way clustering. *Journal of Business and Economic Statistics*, 29(2), 238-249.

[16] Caramori, P.H., Androcioli Filho, A., Carneiro Filho, F., de Oliveira, D., Morais, H., Carneiro Leal A., & Galdino, J. (2007). Métodos de proteção contra geadas em cafezais em formação. *Informe de Pesquisa IAPAR.*

[17] Carvalho, M. S. (1991). A pequena produção de café no Paraná. *PhD dissertation, University of São Paulo, Departtment of Geography.*

[18] Combes, P. P., & Gobillon, L. (2015). The empirics of agglomeration economies. *Handbook of regional and urban economics*, vol. 5, Elsevier.

[19] Conjuntura Econômica (1975). Café -efeitos da geada. September 1975.

[20] Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1), 1-45.

[21] Davis, D. R. & Weinstein, D. E. (2002). Bones, bombs, and break points: the geography of economic activity. *American Economic Review*, 92(5), 1269-1289.

[22] Davis, D. R. & Weinstein, D. E. (2008). A search for multiple equilibria in urban industrial structure. *Journal of Regional Science*, 48(1), 29-65.

[23] Dillon, A., Mueller, V., & Salau, S. (2011). Migratory responses to agricultural risk in northern Nigeria. *American Journal of Agricultural Economics*, 93(4), 1048-1061.

[24] Douthat, T. H. (2017). *Adaptive efficiency in coffee clusters: resilience through agglomeration, global value chains, social networks, and institutions*. Doctoral dissertation, Georgia Institute of Technology.

[25] Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics* (Vol. 4, pp. 2063-2117), Elsevier.

[26] Emerick, K. (2017). Agricultural productivity and the sectoral reallocation of labor in rural India. *Journal of Development Economics*, 135, 488-503.

[27] Feigenbaum, J.J., J. Lee & Mezzanotti, F. (2018). Capital destruction and economic growth: the effects of Sherman's march, 1850-1920. *Working paper*.

[28] Ferman, B. (2019). A simple way to assess inference methods. *Working paper*.

[29] Folha de Londrina (2015). Cafeicultura enfrenta desafios.

[30] Gallagher, J. (2014). Learning about an infrequent event: evidence from flood insurance take-up in the United States. *American Economic Journal: Applied Economics*, 206-233.

[31] GERCA (1975). *Recuperação e condução de cafezais geados*. Instituto Brasileiro do Café.

[32] Gray, C., & Mueller, V. (2012a). Drought and population mobility in rural Ethiopia. *World development*, 40(1), 134-145.

[33] Gray, C. L., & Mueller, V. (2012b). Natural disasters and population mobility in Bangladesh. *Proceedings of the National Academy of Sciences*, 201115944.

[34] Gröger, A., & Zylberberg, Y. (2016). Internal labor migration as a shock coping strategy: evidence from a typhoon. *American Economic Journal: Applied Economics*, 8(2), 123-53.

[35] Hanlon, W. W. (2017). Temporary shocks and persistent effects in urban economies: evidence from British cities after the US Civil War. *Review of Economics and Statistics*, 99(1), 67-79.

[36] Hayami, Y., & Ruttan, V. W. (1970). Factor prices and technical change in agricultural development: the United States and Japan, 1880-1960. *Journal of Political Economy*, 78(5), 1115-1141.

[37] Holmes, T. J., & Lee, S. (2012). Economies of density versus natural advantage: crop choice on the back forty. *Review of Economics and Statistics*, 94(1), 1-19.

[38] Hornbeck, R. & Naidu, S. (2014). When the levee breaks: labor mobility and economic development in the American south. *American Economic Review*, 104(3): 963-990.

[39] IBGE (2016). *A geografia do café*. Dinâmica Territorial da Produção Agropecuária, Rio de Janeiro.

[40] Jessoe, K., Manning, D. T., & Taylor, J. E. (2017). Climate change and labour allocation in rural Mexico: evidence from annual fluctuations in weather. *The Economic Journal*, 128(608), 230-261.

[41] Juhász, R. (2018). Temporary protection and technology adoption: evidence from the Napoleonic Blockade. *American Economic Review*, 108(11), 339-3376.

[42] Kleemans, M., & Magruder, J. (2018). Labour market responses to immigration: evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613), 2032-2065.

[43] Kohlhepp, G., Soethe, P.A., Martineschen, D., da Costa Pereira, C.H., Mathias, D., Antoniuk, E., Boechat, F.B., Fullgraf, F., da Silva, N.P., Paulino,S. & Neubauer, S.N. (2014). *Colonização agrária no Norte do Paraná: processos geoeconômicos e sociogeográficos de desenvolvimento de uma zona subtropical do Brasil sob a influência da plantação de café*. Eduem, Maringá, Portuguese translation of the 1975 German original.

[44] Krugman, P. (1991). History versus expectations. *The Quarterly Journal of Economics*, 106(2), 651-667.

[45] Mahajan, P. & Yang, D. (2020). Taken by storm: hurricanes, migrant networks, and US immigration. *American Economic Journal: Applied Economics*, 12(2), 250-77.

[46] Margolis, M. (1973). *The moving frontier: social and economic change in a southern Brazilian community*. University of Florida Press, Gainesville.

[47] Matsuyama, K. (1991). Increasing returns, industrialization, and indeterminacy of equilibrium. *The Quarterly Journal of Economics*, 106(2), 617-650.

[48] McCoy, S. J., & Walsh, R. P. (2018). Wildfire risk, salience and housing demand. *Journal of Environmental Economics and Management*, 91, 203-228.

[49] Michaels, G., Rauch, F., & Redding, S. J. (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2), 535-586.

[50] Miguel, E., & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1), 1-15.

[51] Minale, L. (2018). Agricultural productivity shocks, labour reallocation and rural–urban migration in China. *Journal of Economic Geography*, 18(4), 795-821.

[52] Moreira, J.R.S., & Laverdi, R. (2016). Modernização agrícola e o declínio da cafeicultura pós-geada de 1975 (Norte do Paraná). *Ateliê de História UEPG*, 4(1): 93-106.

[53] Mores, L. (2017). História ambiental do agroecossistema do café (*Coffea arabica*) no norte do Paraná, 1945-1975. *Master thesis*, Universidade Federal de Santa Catarina.

[54] Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, 118(2), 549-599.

[55] Oliveira, F. (1978). O café e o Norte do Paraná. *O Café no Brasil*, Instituto Brasileiro do Café, Rio de Janeiro.

[56] Pendergast, M. (1999). *Uncommon grounds*. Basic Books.

[57] Pereira, L. L., & das Chagas Ribeiro, A. (2015). *O perfil da organização produtiva dos cafeicultores das microrregiões Serrana e Caparaó*. Appris Editora.

[58] Pozzobon, I. (2006). *A Epopéia do Café no Paraná*. Grafmarke, Londrina.

[59] Proost, S., & Thisse, J.F. (2019). What can be learned from spatial economics? *Journal of Economic Literature*, 57(3), 575-643.

[60] Reis, E., Pimentel., M., Alvarenga, A., & Santos, M. (2008). Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. *Rio de Janeiro, Ipea/Dimac.*

[61] Richards, P. (2017). It's not just where you farm; it's whether your neighbor does too: how agglomeration economies are shaping new agricultural landscapes. *Journal of Economic Geography*, 18, 87-110.

[62] Robalino, J. A., & Pfaff, A. (2012). Contagious development: neighbor interactions in deforestation. *Journal of Development Economics*, 97(2), 427-436.

[63] Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of regional and urban economics*, vol. 4, Elsevier.

[64] Sediyama, G.C., Ferreira de Melo Junior, J.C., Rosa dos Santos, A., Ribeiro, A., Costa, M.H., Hamakawa, P.J., Nogueira da Costa, J.M., & Costa, L.C. (2001). Zoneamento agroclimático do cafeeiro (Coffea arabica L.) para o estado de Minas Gerais". *Revista Brasileira de Agrometeorologia*, 9(3), 501-509.

[65] Schumann, A. (2014). Persistence of population shocks: evidence from the occupation of West Germany after World War II. *American Economic Journal: Applied Economics*, 6(3), 189-205.

[66] Siodla, J. (2021). Firms, fires, and firebreaks: the impact of the 1906 San Francisco disaster on business agglomeration. *Regional Science and Urban Economics*, 88, 103659.

[67] TV Coroados (1980). *Filhos do Café*. Londrina, Brazil.

[68] Valor Econômico (2012). A difícil tarefa de resgatar a produção cafeeira no Paraná. 07/03/2012.

[69] Waldinger, F. (2016). Bombs, brains, and science: the role of human and physical capital for the creation of scientific knowledge. *The Review of Economics and Statistics*, 158 (5), 811-831.

[70] Zylbersztajn, D., Farina, E.M.M.Q., & da Costa Santos, R. (1993). *O sistema agroindustrial do café*. Editora Ortiz, Porto Alegre.

A.1 Model

Consider a small open economy inhabited by a continuum of farmers, each holding a 1 hectare plot. Time is discrete and infinite, denoted by $t = 1, 2, 3, ...$ Farmers discount the future by a factor $\delta < 1$. In each period, farmers choose which crop they grow: coffee, denoted by C, and non-coffee, denoted by N. Each farmer hires a quantity n of workers in order to produce, in activity $i \in \{C, N\}$, a revenue of $A^i(\gamma_t)f^i(n)$, where γ_t is the share of farmers growing coffee in period t. I assume $A^C(\gamma)$ is continuous and non-decreasing, while $A^N(\gamma)$ is constant in γ ($A^N(\gamma) = A^N$). The assumption allows for positive productivity spillovers (agglomeration economies) in coffee production. I also assume that workers are spatially mobile and, therefore, supply labor to the municipality at the national wage w.

The profit maximization problem of the farmer implies that the labor-land ratio in a farm will depend on the activity chosen and, if producing coffee, on the share γ_t of farmers growing coffee in the local economy. Denote labor demands by $n^C(\gamma_t)$ and n^N. Notice $n^C(\gamma_t)$ is continuous and non-decreasing in γ_t. I assume that $A^C(0)\frac{df^C(n)}{dn} > A^N\frac{df^N(n)}{dn}$ for every $n > 0$, so coffee always requires more labor than the alternative activity ($n^C(0) > n^N$).[28] A consequence of the assumption is that there is an increasing relationship between the share of coffee producers γ_t and total farm employment in the municipality. I also define revenue minus labor costs in activity i as $\pi^i(\gamma_t) = max_n\{A^i(\gamma_t)f^i(n) - wn\}$. Note π^N is constant, while $\pi^C(\gamma_t)$ is a continuous and non-decreasing function of γ_t.

Coffee production in a plot requires the presence of coffee trees. The planting of coffee trees costs $\kappa > 0$ per hectare. I assume that, at the end of every period t, the trees completely depreciate with probability λ; with probability $1 - \lambda$, trees do not depreciate and the farmer is able to produce coffee in the following period without paying any additional costs and without losing productivity. I also assume that a farmer who chooses the non-coffee activity loses all her coffee tree stock without incurring any extra cost.

I characterize the set of equilibria given an initial condition γ_0, and for two different cases: the **frost case**, in which all trees depreciate from period $t = 0$ to $t = 1$, and the **no-frost case**, in which only a fraction λ depreciates from period $t = 0$ to $t = 1$, as it happens at any other period $t \geq 1$. I consider only symmetric equilibria: in period t, all farmers who inherit (do not inherit) trees choose to grow coffee with probability σ_t^I (σ_t^D). In this way, the law of movement for γ_t is:

$$\gamma_t = \gamma_{t-1}\left[\lambda\sigma_t^D + (1-\lambda)\sigma_t^I\right] + (1-\gamma_{t-1})\sigma_t^D \quad \forall t \geq 1$$

[28] Appendix B.2 uses agricultural census data from 1975 to examine if this assumption holds.

Define the payoffs of farmers who inherited coffee trees at t recursively as:

$$U_t^I = max\left\{\pi^N + \delta U_{t+1}^D, \quad \pi^C(\gamma_t) + \delta(1-\lambda)U_{t+1}^I + \delta\lambda U_{t+1}^D\right\} \quad \forall t \geq 1$$

and for those who did not inherit as:

$$U_t^D = max\left\{\pi^N + \delta U_{t+1}^D, \quad \pi^C(\gamma_t) + \delta(1-\lambda)U_{t+1}^I + \delta\lambda U_{t+1}^D - \kappa\right\} \quad \forall t \geq 1$$

Note that U_t^D may be interpreted as the value of farm land, while $U_t^I - U_t^D$ is the value of coffee trees.

A solution $\left(\gamma_t, \sigma_t^I, \sigma_t^D, U_t^I, U_t^D\right)_{t\geq 1}$ for the three equations above that satisfies the initial condition γ_0 is an equilibrium of the no-frost economy. It is an equilibrium of the frost economy if, for $t = 1$ the law of movement is instead $\gamma_1 = \sigma_1^D$, as the frost destroyed all coffee trees from period $t = 0$.

A straightforward consequence of the farmer crop choice problem is that, if producing coffee is optimal for farmers who do not inherit trees and therefore incur the cost κ, it should also be optimal for farmers who inherit them. That is, $\sigma_t^I = 1$ if $\sigma_t^D > 0$. A similar argument shows that $\sigma_t^D = 0$ if $\sigma_t^I < 1$.

Before discussing further properties of the set of equilibria for an arbitrary initial condition, I characterize the steady-states of this economy. A steady-state is a list $(\gamma, \sigma^I, \sigma^D, U^I, U^D)$ such that $\gamma_t = \gamma$, $U_t^I = U^I$, $U_t^D = U^D$, $\sigma_t^I = \sigma^I$, and $\sigma_t^D = \sigma^D$ for all t defines an equilibrium in a no-frost economy. Since the optimality of farmer's crop choices imposes the restrictions above on σ^I and σ^D, any steady-state belongs to one of three types:

1. a **non-coffee steady-state**, in which all farmers inheriting no coffee trees choose the non-coffee activity ($\sigma^D = 0$ and $\gamma = 0$);

2. a **coffee steady-state**, in which farmers choose coffee production regardless of inheriting trees ($\sigma^D = 1$, $\sigma^I = 1$ and $\gamma = 1$);

3. finally, a **diversified steady-state**, in which some but not all of the farmers decide to plant coffee trees ($\sigma^D \in (0,1)$, $\sigma^I = 1$ and $\gamma = \frac{\sigma^D}{\lambda(1-\sigma^D)}$).

The following proposition describes which steady-states exist as a function of model parameters. It also shows that, when only one steady-state is possible, all equilibria will lead to it. The proof is in Online Appendix B.3.

Proposition 1. *The following characterizes the set of steady-states in the economy:*

(i) *if $\pi^C(0) - \pi^N > \kappa(1 - \delta + \delta\lambda)$, then coffee is the only steady-state.*

(ii) *if* $\kappa(1-\delta+\delta\lambda) > \pi^C(1) - \pi^N$*, then non-coffee is the only steady-state.*

(iii) *if* $\pi^C(0) - \pi^N \leq \kappa(1-\delta+\delta\lambda) \leq \pi^C(1) - \pi^N$*, then both coffee and non-coffee steady states exist. Moreover, if the inequalities are strict or* $\pi^C(\gamma)$ *is constant, then there is also a diversified steady-state.*

Furthermore, when there is a single steady-state, all equilibria converge to it.

Note that the steady-states are determined by how profitable coffee is relative to non-coffee and how this difference compares to the expected discounted cost of capital, $\frac{\kappa(1-\delta+\delta\lambda)}{1-\delta}$. When there are no strategic complementarities, then $\pi^C(0) - \pi^N = \pi^C(1) - \pi^N$ and there is a unique steady-state,[29] where the economy specializes in the most profitable activity. Multiplicity of steady-states occurs when strategic complementarities are sufficiently strong, so coffee is the most profitable activity when all other farmers also grow it, but not when no other farmer grows it. In this case, both the coffee and non-coffee steady-state exist. By continuity, a diversified steady-state also exists.

I now discuss how an economy that starts with a given γ_0 evolves toward a steady-state. In the frost case, the destruction of all capital before period $t = 1$ allows the economy to immediately jump to any steady-state. Therefore, the initial condition γ_0 has no effect on the equilibrium, which will depend entirely on self-fulfilling expectations about $\{\gamma_t\}$. In this way, an economy that was initially specialized in coffee production (high γ_0) can rapidly converge to a non-coffee economy after the frost. Such a transition is possible when strategic complementarities result in the multiplicity of steady-states and farmers expect the equilibrium to converge to a non-coffee steady-state.

On the other hand, a no-frost economy could be anchored in coffee specialization. Suppose that $\pi^C(0) > \pi^N$, so any farmer who inherits coffee trees continues producing coffee. Hence by $t = 1$ there are still many farmers growing coffee: $\gamma_1 \geq (1-\lambda)\gamma_0$. If strategic complementarities are sufficiently strong, then the presence of all these coffee producers makes coffee optimal also for farmers who had their trees depreciated. The economy has thus converged to a coffee steady-state, as summarized by the proposition below:

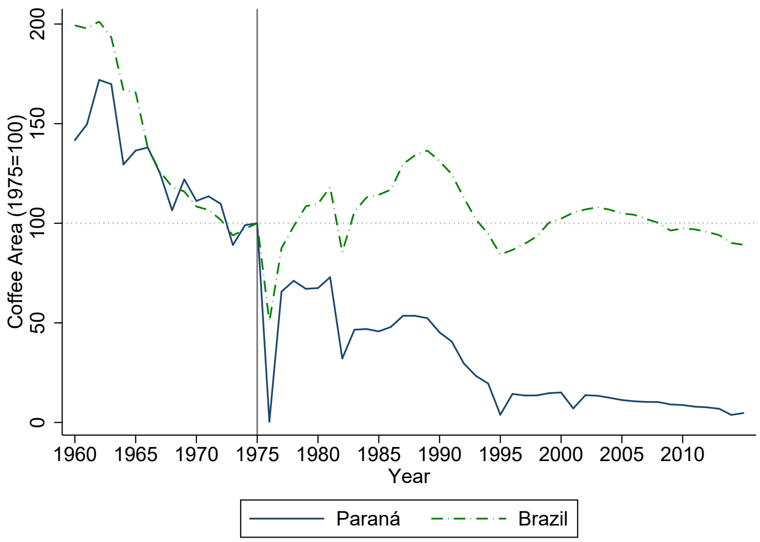
Proposition 2. *If* $\pi^C(0) > \pi^N$ *and* $\pi^C(1-\lambda) - \pi^N > \kappa$*, then, for a sufficiently high* γ_0*, all equilibria of the no-frost economy converge to a coffee steady-state.*

[29] Except in the non-generic case for which $\pi^C(0) - \pi^N = \kappa(1-\delta+\delta\lambda) = \pi^C(1) - \pi^N$.

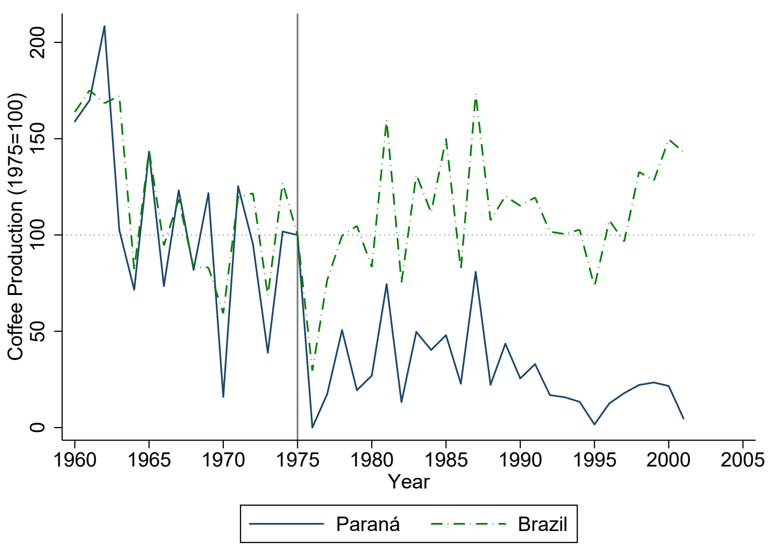
A.2 Figures

Figure 1: Coffee industry in Paraná and Brazil

(a) Harvested area (scaled by 1975 levels):

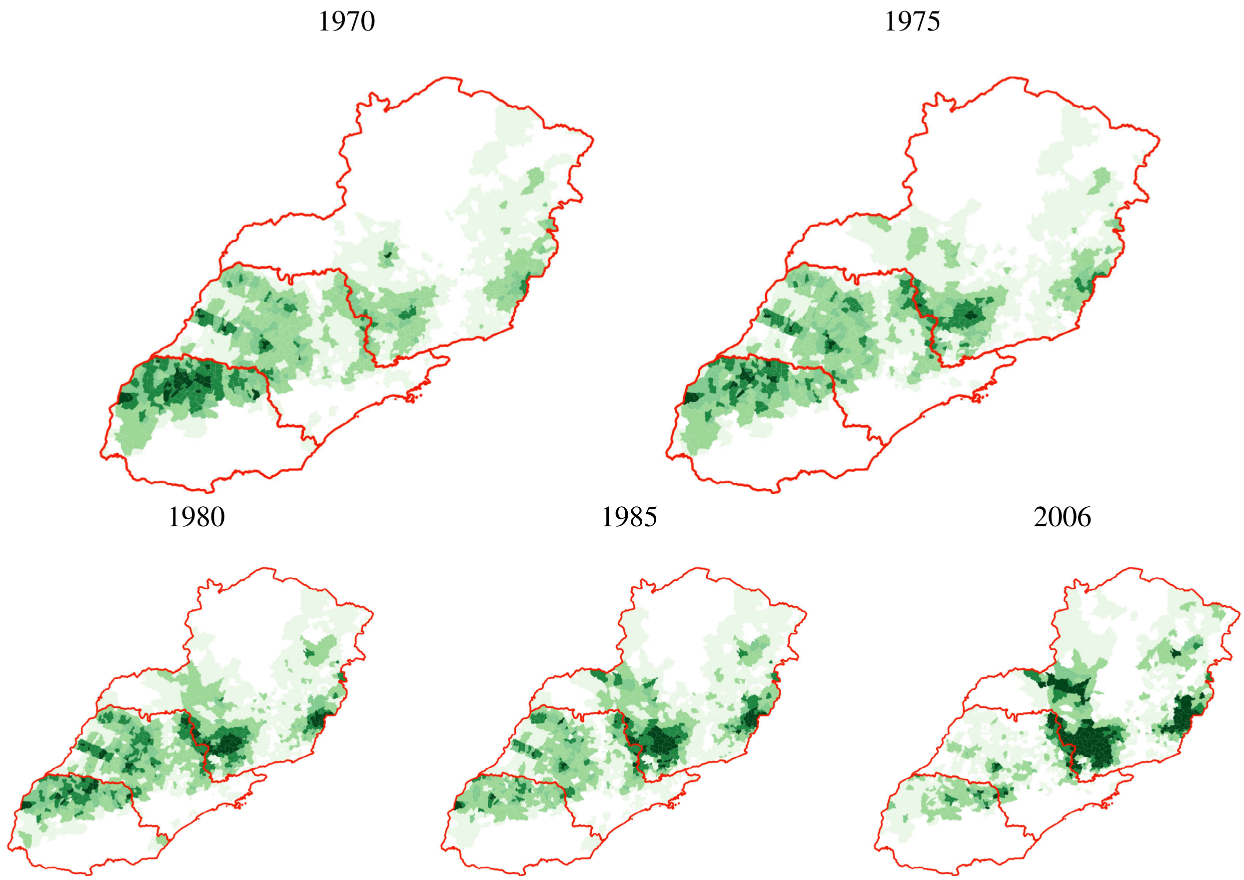


(b) Production (scaled by 1975 levels):



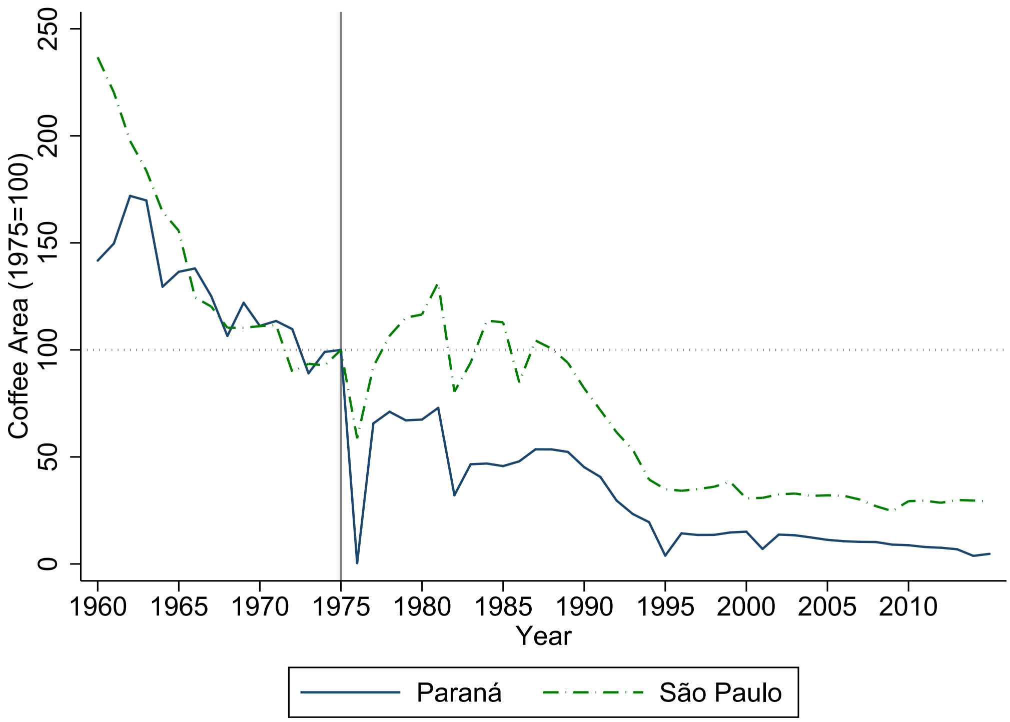
Note: Series are from the *Produção Agrícola Municpal* (since 1974) and from the *Instituto Brasileiro do Café* (from 1960 to 1973).

Figure 2: Spatial distribution of coffee trees



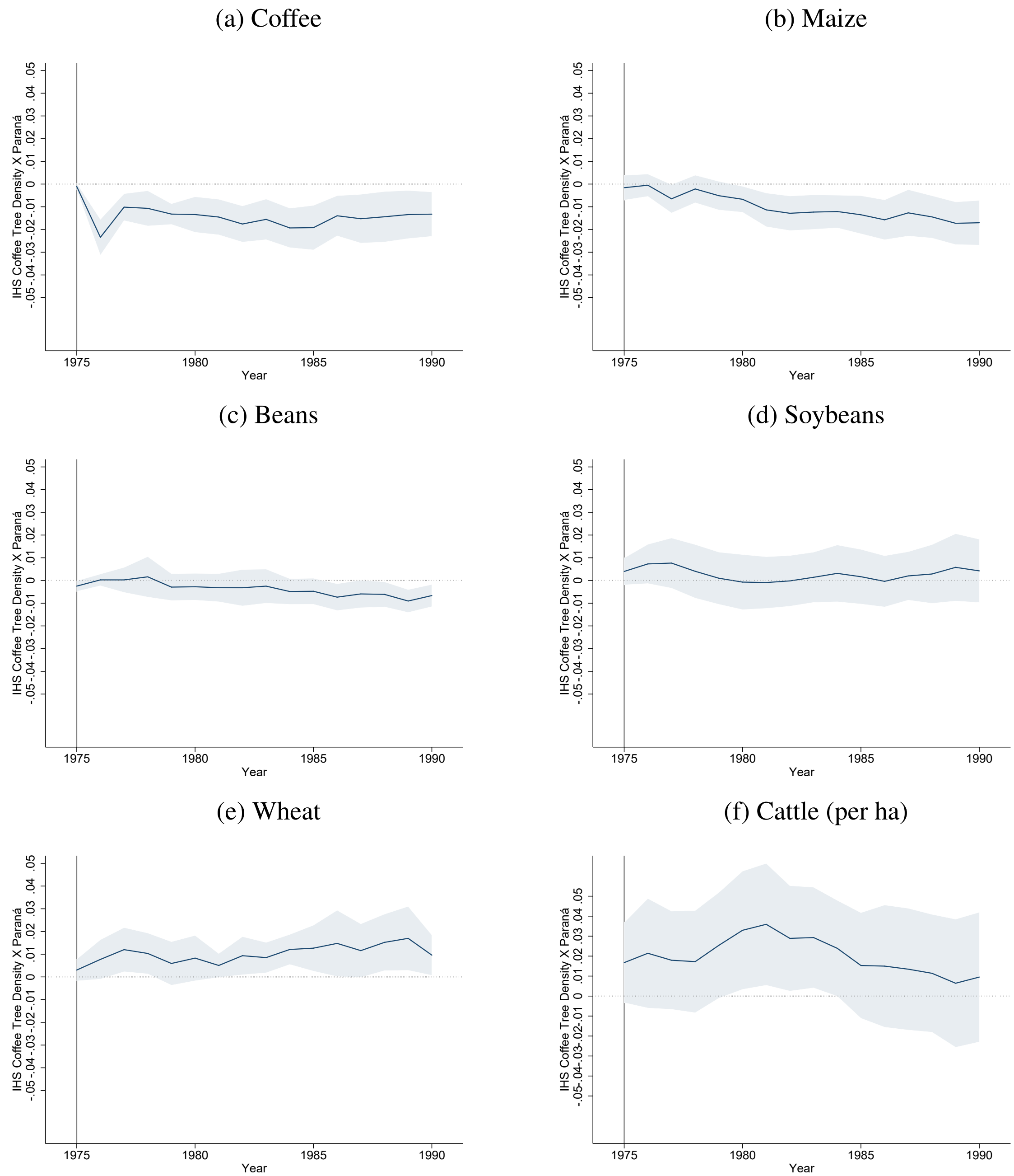
Note: Data from the 1970, 1975, 1980, 1985 and 2006 agricultural censuses. The coffee tree density is measured at December of the census year and refers to the sum of productive-age and young trees. Paraná is the bottom state, São Paulo the middle one, and Minas Gerais the state on the top. Red lines denote state borders. Coffee tree density in white municipalities is lower than the 50^{th} of the 1970 distribution in Paraná (0.778 trees/ha); in light green, between the 50 and 75^{th} percentiles (0.778-14.63); medium-light green denotes density between the 75 and 90^{th} percentiles (14.63-73.34); in medium green, from the 90^{th} to 95^{th} percentiles (73.34-119.44); medium-dark green, from the 95^{th} to the 99^{th} percentile (119.44-259.93); and dark green, in the top percentile (above 259.93 trees/ha).

Figure 3: Coffee harvested area in Paraná and São Paulo (scaled by 1975 levels)



Note: Series are from the *Produção Agrícola Municpal* (since 1974) and from the *Instituto Brasileiro do Café* (from 1960 to 1973).

Figure 4: Effects on the harvested area of alternative crops



Note: Data from PAM and PPM. The dependent variable is the change in the harvested area of a crop (or of cattle) from 1974 to year t, divided by the 1975 farm area. The solid navy line represents the coefficients for the IHS of coffee tree density in 1975 times the Paraná dummy, $\beta_{t,1974}$. The shaded area indicates the 95% confidence interval for β_t^{PR}. The regressions also include the IHS of coffee tree density in 1975 and the Paraná dummy; see equation 1. Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region.

A.3 Tables

Table 1: Share of young trees to total tree stock

	Paraná		São Paulo		Minas Gerais	
	Young	Planted	Young	Planted	Young	Planted
	(1)	(2)	(3)	(4)	(5)	(6)
December 1970	13.4%		12.0%		17.2%	
December 1975	11.9%	5.1%	17.9%	9.3%	29.4%	15.7%
December 1980	15.5%	5.2%	18.1%	7.8%	31.0%	16.9%

Note: Data from the agricultural censuses of 1970, 1975, and 1980. Odd columns show the share of young trees, by state, at the last month of each census year. Even columns show the share of trees that were planted in the census year; this information is available only for 1975 and 1980.

Table 2: Descriptive statistics

	Mean	Stand. Dev.
Farms	2425.04	4617.88
Coffee producing farms	367.96	809.09
Coffee tree density in 1975	33.47	55.95
Agricultural employment	9033.31	17686.57
Agricultural employment per hectare	0.10	0.08
Share of total employment in agriculture (%)	62.51	22.35
Share of agriculture employment in coffee (%)	16.42	23.89
Labor expenditure share in agriculture (%)	23.27	11.40
Land transactions (% of land value)	1.32	1.23
Urbanization rate (%)	38.92	23.41
In-migration rate, 1970-80 (%)	23.13	15.53
Out-migration rate, 1970-80 (%)	34.57	17.02
Observations	778	

Note: Data from the agricultural and population censuses of 1970, except coffee tree density, that is from the 1975 agricultural census, and the migration rates, that are constructed with information from both the 1970 and 1980 population censuses. Statistics weighted by municipality area.

Table 3: Effects on employment in agriculture

	Log total farm employment					Log non-temporary farm employment	
Initial period (s)	1970	1975	1980	1970	1970	1970	1975
Final period (t)	1975	1980	1985	1980	1985	1975	1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IHS Coffee Tree Density in 1975 × Paraná	-0.001	-0.081***	-0.017	-0.082***	-0.099***	-0.006	-0.084***
	(0.015)	(0.017)	(0.011)	(0.022)	(0.023)	(0.018)	(0.017)
IHS Coffee Tree Density in 1975	-0.033***	0.015	-0.018**	-0.018	-0.035*	-0.027*	0.020
	(0.011)	(0.013)	(0.008)	(0.020)	(0.020)	(0.014)	(0.012)
Paraná	0.074	0.118	0.123**	0.192**	0.315***	0.073	0.187**
	(0.059)	(0.073)	(0.049)	(0.087)	(0.088)	(0.064)	(0.075)
R^2	0.064	0.122	0.069	0.194	0.268	0.057	0.115
Observations	778	778	778	778	778	778	778

Note: Dependent variables are from the agricultural censuses of IBGE. The dependent variables are the changes in the variable at the top of the column from the initial period s to the final period t. Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 4: Long differences in agricultural employment

	Log total farm employment			
Initial period (s)	1970	1970	1960	1960
Final period (t)	1995	2006	1970	1975
	(1)	(2)	(3)	(4)
IHS Coffee Tree Density in 1975 × Paraná	-0.144***	-0.163***	0.058	0.062
	(0.030)	(0.033)	(0.047)	(0.051)
IHS Coffee Tree Density in 1975	-0.016	-0.003	-0.040***	-0.073***
	(0.022)	(0.023)	(0.014)	(0.016)
Paraná	0.453***	0.413***	0.592***	0.656***
	(0.120)	(0.129)	(0.127)	(0.141)
R^2	0.236	0.185	0.303	0.310
Observations	778	778	566	566

Note: Dependent variables are from the agricultural censuses of IBGE. The dependent variables are the changes in the variable at the top of the column from the initial period s to the final period t. Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 5: Other agricultural variables

	Log wage bill per worker (1)	Log expend. per worker (2)	Labor share in exp. (3)	Log land price (4)
1970-1975	-0.053*	-0.030	0.007	0.018
	(0.031)	(0.033)	(0.009)	(0.028)
1975-1980	0.020	0.013	0.002	-0.080**
	(0.029)	(0.022)	(0.005)	(0.033)
1970-1980	-0.032	-0.017	0.008	-0.062**
	(0.030)	(0.043)	(0.007)	(0.025)
1970-2006	-0.125***	-0.058	0.004	-0.183***
	(0.045)	(0.056)	(0.007)	(0.033)

Note: Dependent variables are from the agricultural censuses of IBGE. Each cell reports the coefficient on the interaction between the IHS of coffee tree density in 1975 interacted and the Paraná dummy. The regressions also include the IHS of coffee tree density in 1975 and the Paraná dummy as controls; see equation 1. The dependent variable is the change in the variable indicated at the top for the time period indicated at the left. Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 6: Sectoral reallocation and migration of labor

	Log employment in non-agriculture		Migration rates (1970-80)		
	1970-80	1970-91	net-mig.	in-mig.	out-mig.
	(1)	(2)	(3)	(4)	(5)
IHS Coffee Tree Density in 1975 $\times$ Paraná	-0.029	-0.019	-0.045***	0.001	0.046***
	(0.019)	(0.020)	(0.011)	(0.007)	(0.010)
IHS Coffee Tree Density in 1975	-0.010	-0.014	-0.012	-0.005	0.008
	(0.011)	(0.013)	(0.008)	(0.004)	(0.007)
Paraná	0.138	-0.013	0.027	0.009	-0.018
	(0.089)	(0.080)	(0.052)	(0.035)	(0.044)
R^2	0.049	0.028	0.166	0.005	0.320
Observations	778	778	778	778	778

Note: Dependent variables are from the population censuses of IBGE. Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 7: Worker earnings

	All workers		Agriculture		Farm laborers	
	all	men	all	men	all	men
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1970-80						
IHS Coffee Tree Density in 1975 × Paraná	0.010	0.014	0.007	0.008	-0.009	-0.009
	(0.013)	(0.013)	(0.021)	(0.021)	(0.012)	(0.012)
IHS Coffee Tree Density in 1975	-0.001	0.002	-0.006	-0.003	0.012	0.014
	(0.009)	(0.009)	(0.014)	(0.014)	(0.009)	(0.009)
Paraná	-0.018	-0.038	-0.032	-0.051	-0.068	-0.075
	(0.052)	(0.049)	(0.081)	(0.080)	(0.045)	(0.046)
R^2	0.004	0.014	0.001	0.003	0.066	0.079
Panel B: 1970-91						
IHS Coffee Tree Density in 1975 × Paraná	0.021**	0.023**	0.030	0.025	-0.021	-0.017
	(0.010)	(0.010)	(0.022)	(0.020)	(0.022)	(0.022)
IHS Coffee Tree Density in 1975	0.002	0.004	-0.018	-0.017	-0.011	-0.010
	(0.006)	(0.006)	(0.011)	(0.011)	(0.015)	(0.015)
Paraná	-0.150***	-0.165***	-0.305***	-0.279***	-0.263***	-0.271***
	(0.042)	(0.038)	(0.096)	(0.084)	(0.094)	(0.092)
R^2	0.075	0.092	0.097	0.092	0.175	0.173
Observations	778	778	778	778	778	778

Note: Dependent variables are from the population censuses of IBGE. The dependent variable is the change in the logarithm of the average earnings of workers in each group from 1970 to 1980 (Panel A) or from 1970 to 1991 (Panel B). Observations are weighted by municipality area. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 8: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A:	Alternative samples:		Control for initial conditions:				
	Minas Gerais	Met. regions	Pop. density	Ag. share	Earnings	Cattle density	All
1970-1975	-0.031*	-0.003	0.004	-0.005	-0.001	-0.007	-0.013
	(0.016)	(0.015)	(0.017)	(0.016)	(0.015)	(0.016)	(0.017)
1975-1980	-0.050***	-0.084***	-0.072***	-0.074***	-0.081***	-0.064***	-0.051**
	(0.015)	(0.016)	(0.020)	(0.016)	(0.015)	(0.022)	(0.021)
1980-1985	-0.030***	-0.018*	-0.019	-0.010	-0.016	-0.017	-0.015
	(0.010)	(0.010)	(0.015)	(0.011)	(0.011)	(0.013)	(0.011)
1970-1980	-0.081***	-0.087***	-0.068***	-0.078***	-0.082***	-0.071***	-0.065***
	(0.017)	(0.021)	(0.022)	(0.022)	(0.021)	(0.022)	(0.023)
1970-1985	-0.110***	-0.105***	-0.087***	-0.088***	-0.098***	-0.089***	-0.080***
	(0.019)	(0.023)	(0.027)	(0.023)	(0.022)	(0.023)	(0.023)
Observations	951	834	778	778	778	778	778
Panel B:	Control for pre-trends:		Alternative measures:			Alternative weighting:	
	1970-75	1960-70	Dummy	Level	Per worker	Simple	Employment
1970-1975		0.009	-0.058	-0.037	0.003	-0.007	-0.004
		(0.015)	(0.073)	(0.040)	(0.012)	(0.014)	(0.014)
1975-1980	-0.084***	-0.085***	-0.403***	-0.186***	-0.064***	-0.062***	-0.053***
	(0.018)	(0.020)	(0.068)	(0.056)	(0.013)	(0.015)	(0.017)
1980-1985	-0.020	-0.012	-0.101*	-0.013	-0.016**	-0.018	0.002
	(0.012)	(0.010)	(0.051)	(0.043)	(0.008)	(0.013)	(0.011)
1970-1980	-0.084***	-0.075***	-0.461***	-0.223***	-0.060***	-0.069***	-0.057***
	(0.018)	(0.026)	(0.090)	(0.069)	(0.018)	(0.019)	(0.021)
1970-1985	-0.104***	-0.087***	-0.562***	-0.237***	-0.077***	-0.087***	-0.055**
	(0.019)	(0.024)	(0.100)	(0.086)	(0.018)	(0.022)	(0.025)
Observations	778	566	778	778	778	778	778

Note: Dependent variables are from the agricultural censuses of IBGE. The dependent variables are the changes in the logarithm of total employment in agriculture during the time interval indicated at the left. Coefficients for the interaction term. See the main text in Section 5 for details of each specification. Robust standard errors are clustered at the micro-region. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table 9: Robustness to alternative inference methods

	Changes in log employment in agriculture					Simulated rejection probability
	1970-75	1975-80	1980-85	1970-80	1970-85	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	-0.001	-0.081	-0.017	-0.082	-0.099	
Microregion clusters (baseline)	(0.015)	(0.017)	(0.011)	(0.022)	(0.023)	0.040
Robust to heteroskedasticity	(0.012)	(0.014)	(0.010)	(0.014)	(0.015)	0.028
Distance cutoff:						
100km	(0.021)	(0.023)	(0.017)	(0.028)	(0.030)	0.038
200km	(0.022)	(0.022)	(0.018)	(0.029)	(0.032)	0.054
300km	(0.022)	(0.019)	(0.015)	(0.027)	(0.029)	0.060
400km	(0.019)	(0.015)	(0.014)	(0.022)	(0.028)	0.094
500km	(0.017)	(0.012)	(0.015)	(0.019)	(0.027)	0.126

Note: Dependent variables are from the agricultural censuses of IBGE. The dependent variables are the changes in the logarithm of total employment in agriculture during the time interval indicated at the top of the table. Observations are weighted by municipality area. Each cell reports the coefficient on the interaction between the IHS of coffee tree density in 1975 interacted with the Paraná dummy. The regressions also include the IHS of coffee tree density in 1975 and the Paraná dummy as controls; see equation 1. Different estimates for the standard errors in the parentheses. From top to bottom: **(i)** standard errors robust to heteroskedasticity and correlation within microregions; **(i)** robust to heteroskedasticity but not to cross-sectional dependence; **(ii)** robust to heteroskedasticity and to spatial correlation as in Conley (1999), with a triangular kernel and four different cutoffs. Column (6) indicates the simulated rejection probabilities of a 5% t-test assuming the inference method in that line. Simulations assume the null hypothesis and that residuals are homescedastic, according to the method suggested by Ferman (2019).

Table 10: South-facing terrains and coffee tree coverage

	(1)	(2)	(3)	(4)	(5)	(6)
Share of terrain facing south	-0.0044**	-0.0027***	-0.0025***	-0.0034**	-0.0022***	-0.0020***
	(0.0019)	(0.0011)	(0.0010)	(0.0015)	(0.0009)	(0.0009)
Share of neighboring area facing south		-0.0306**	-0.0362**	-0.0540**		
		(0.0163)	(0.0162)	(0.0233)		
Coffee coverage on the neighboring area					1.0327**	1.1090***
					(0.0188)	(0.0092)
Robust F-statistic in the 1st stage					5.51	
Anderson-Rubin 95% CI					[0.97,1.13]	
R-squared	0.001	0.004	0.028	0.028		0.289
Observations	120,118	120,118	120,118	198,056	120,118	120,118
Controls?			Y	Y	Y	Y
Estimation	OLS	OLS	OLS	OLS	2SLS	OLS
Unit of observation	grid cell	grid cell	grid cell	farm	grid cell	grid cell

Note: See Appendix B.4 for a description of the data. The dependent variable is the share of a cell or a farm that is covered by coffee. Control variables include own and neighboring averages of altitude, slope, share within 500m of a river, and share of *Terra Roxa*, as well as latitude, latitude squared, and distance to a road. Observations weighted by area. Robust standard errors in parentheses are multi-way clustered according to four different grids of 0.5° × 0.5 °; see Appendix B.4 for a description of the inference procedure. Statistical significance denoted by: * 10%, ** 5%, *** 1%.

Online Appendix for

Capital as an Anchor of Economic Activity:

Evidence from the 1975 Frost

B.1 Concentration Indexes

The frost was associated with changes in the intra-state distribution of the coffee industry. I calculate Theil indexes to measure the concentration of mature coffee trees relative to farm area, following Brülhart and Traeger (2005). Let $s^C_{i\sigma t}$ and $s^F_{i\sigma t}$ be the share in state σ and year t of, respectively, coffee trees and farm area in municipality i. The Theil Index for state σ and year t is defined as:

$$T_{\sigma t} \equiv \sum_i s^C_{i\sigma t} \cdot log\left(\frac{s^C_{i\sigma t}}{s^F_{i\sigma t}}\right)$$

The larger the index, the higher the concentration of coffee trees relative to farm area. Table B.1 displays the Theil indexes for Paraná, São Paulo, Minas Gerais, and the aggregate index for the three states, for the agricultural census years of 1970, 1975, 1980, 1985, and 2006. Column (1) shows a decline in the concentration of the coffee industry for the aggregated three states from 1970 to 1980. This is mostly due to the reduction of the relative importance of Paraná. In Paraná, the coffee industry became progressively more concentrated over time. On the other hand, the industry in Minas Gerais (column 4) became less concentrated, which is consistent with the expansion of coffee production into new areas.

B.2 Labor-Land Ratios

In the paper, I assume that a movement of farm activity away from coffee leads to a reduction in labor demand. That is, that the labor per hectare used in coffee was higher than in alternative farm activities (at the current wages). In this appendix, I examine this assumption. There is no available data on the allocation of labor to each activity, but it is possible to see how the total labor demand, measured by the number of total farm workers, correlates with the harvested area of each of the main crops. Table B.2 shows these correlations. It reports the coefficients of a regression of farm workers on the harvested area of beans, coffee, maize, soybeans, sugarcane, and wheat, the harvested area of other crops, and the total farm area (and no constant). The sample consists of the Paraná and São Paulo municipalities used in the main empirical analysis. The coefficient on coffee is the second highest, after beans. The coefficient for coffee and beans are not statistically different. Moreover, the coefficient on coffee is statistically different (at a 5% significance level) than the coefficients for maize, soy, sugarcane and wheat. This is evidence that municipalities with more coffee production had higher farm employment per farm hectare.

B.3 Proofs

Proof of Proposition 1

The proof is in four parts. Part 1 shows the conditions for the existence of a coffee steady-state. Part 2 shows conditions for the existence of a non-coffee steady-state. Part 3 shows conditions for the existence of a diversified steady-state. Finally, Part 4 shows that, in the case the steady-state is unique, all equilibria converge to it.

Part 1: coffee steady-state. In a coffee steady-state, $\sigma^D = 1$, so we have that:

$$U^I = \pi^C(1) + \delta\Big[(1-\lambda)U^I + \lambda U^D\Big]$$
$$U^D = \pi^C(1) + \delta\Big[(1-\lambda)U^I + \lambda U^D\Big] - \kappa$$

hence $U^I - U^D = \kappa$. Solving for these two payoffs as a function of parameters results in:

$$U^I = \frac{\pi^C(1)}{1-\delta} - \frac{\delta\lambda}{1-\delta}\kappa$$
$$U^D = \frac{\pi^C(1)}{1-\delta} - \frac{(1-\delta+\delta\lambda)}{1-\delta}\kappa$$

Given these payoffs, $\sigma^D = 1$ optimal, and hence a coffee steady-state exists, when:

$$\pi^N + \delta U^D \leq \pi^C(1) + \delta(1-\lambda)\kappa + \delta U^D - \kappa$$

which is equivalent to:

$$\pi^C(1) - \pi^N \geq (1-\delta+\delta\lambda)\kappa$$

Part 2: non-coffee steady-state. In a non-coffee steady-state, $\sigma^D = 0$, so $U^D = \frac{\pi^N}{1-\delta}$.

If $\sigma^I = 0$, then $U^I = U^D = \frac{\pi^N}{1-\delta}$. In this case, $\sigma^I = 0$ is an optimal response if, and only if, $\pi^N > \pi^C(0)$. This condition also makes $\sigma^D = 0$ optimal.

If $\sigma^I > 0$, then

$$U^I = \pi^C + \delta\lambda U^D + \delta(1-\lambda)U^I$$

which implies that

$$U^I = \frac{1}{1-\delta}\left[\frac{(1-\delta)}{1-\delta+\delta\lambda}\pi^C(0) + \frac{\delta\lambda}{1-\delta+\delta\lambda}\pi^N\right]$$

So, the payoff gain of inheriting coffee trees is $U^I - U^D = \frac{\pi^C(0)-\pi^N}{1-\delta+\delta\lambda}$. Hence, $\sigma^D = 0$ if, and only if,

$$\pi^N + \delta U^D \geq \pi^C(0) + \delta\left[(1-\lambda)U^I + \lambda U^D\right] - \kappa$$

which is equivalent to $(1-\delta+\delta\lambda)\kappa \geq \pi^C(0) - \pi^N$.

Step 3: diversified steady-state. In a diversified steady-state, $\sigma^D \in (0,1)$, so $U^D = \frac{\pi^N}{1-\delta}$. Since $\sigma^I = 1$, we have that:

$$U^I = \pi^C(\gamma) + \delta\lambda U^D + \delta(1-\lambda)U^I$$

which implies the following solution to U^I:

$$U^I = \frac{1}{1-\delta}\left[\frac{(1-\delta)}{1-\delta+\delta\lambda}\pi^C(\gamma) + \frac{\delta\lambda}{1-\delta+\delta\lambda}\pi^N\right]$$

so $U^I - U^D = \frac{\pi^C(\gamma)-\pi^N}{1-\delta+\delta\lambda}$.

This equilibrium occurs only if a farmer without coffee trees is indifferent between growing coffee or non-coffee, which occurs only if:

$$(1-\delta+\delta\lambda)\kappa = \pi^C(\gamma) - \pi^N$$

where $\gamma = \frac{\sigma^D}{\lambda(1-\sigma^D)}$. Due to the monotonicity of $\pi^C(\gamma) - \pi^N$, the condition above implies that $\pi^C(0) - \pi^N \leq \kappa(1-\delta+\delta\lambda) \leq \pi^C(1) - \pi^N(1)$. Also, since $\pi^C(\gamma) - \pi^N(\gamma)$ is continuous, the reverse is also true. This proves all the cases of the proposition.

Step 4: convergence. When there is a single steady-state, it is either a coffee steady-state (**case 1**) or a non-coffee steady-state (**case 2**). First, I analyze **case 1**. We know from Step 1 that $\pi^N < \pi^C(0) - \kappa[1-\delta+\delta\lambda]$. Consider a farmer who inherited coffee threes. Then, her payoff from coffee is higher than her payoff from non-coffee, as:

$$\pi^N + \delta U_{t+1}^D < \pi^C(\gamma_t) + \delta U_{t+1}^D + \delta(1-\lambda)[U_{t+1}^I - U_{t+1}^D]$$

since $U_{t+1}^I > U_{t+1}^D$. Hence, $\sigma_t^I = 1$.

Therefore, we can write U_t^I as:

$$U_t^I = \pi^C(\gamma_t) + +\delta U_{t+1}^D + \delta(1-\lambda)[U_{t+1}^I - U_{t+1}^D]$$

If we subtract the expression for U_t^D from the equation above, we have the following expression for $U_t^I - U_t^D$:

$$U_t^I - U_t^D = min\Big\{\pi^C(\gamma_t) - \pi^N + \delta(1-\lambda)[U_{t+1}^I - U_{t+1}^D],\ k\Big\}$$

In this way, we can define $\underline{\Delta}$ as a lower bound for $U_t^I - U_t^D$ that satisfies:

$$\underline{\Delta} = min\Big\{\pi^C(0) - \pi^N + \delta(1-\lambda)\underline{\Delta},\ k\Big\}$$

The solution to the equation above will be either $\underline{\Delta} = \kappa$ or $\underline{\Delta} = \frac{\pi^C(0)-\pi^N}{1-\delta+\delta\lambda}$.

Consider now the choice of a farmer who does not inherit trees from the previous period. She will choose coffee ($\sigma_t^D = 1$) if:

$$\pi^C(\gamma_t) - \pi^N > \kappa - \delta(1-\lambda)[U_{t+1}^I - U_{t+1}^D]$$

The LHS is greater or equal than $\pi^C(0) - \pi^N$, while the LHS is lower or equal than:

1. $\kappa[1-\delta+\delta\lambda]$ if $\underline{\Delta} = \kappa$;
2. $\pi^C(0) - \pi^N$ if $\underline{\Delta} = \frac{\pi^C(0)-\pi^N}{1-\delta+\delta\lambda}$.

In either case, the LHS is greater than the RHS, so $\sigma_t^D = 1$. The economy thus converges to a coffee steady-state.

Now I analyze **case 2**, when only the non-coffee steady-state exists. In this case, we know from Step 2 that $\kappa(1-\delta+\delta\lambda) > \pi^C(1) - \pi^N$. Consider a farmer who does not inherit coffee trees. She will choose non-coffee ($\sigma_t^D = 0$), since:

$$\pi^C(\gamma_t) + \delta U_t^D + \delta(1-\lambda)[U_{t+1}^I - U_{t+1}^D] - \kappa \leq \pi^C(1) + \delta U_{t+1}^D - \kappa(1-\delta+\delta\lambda) < \pi^N + \delta U_{t+1}^D$$

In this way, we know that $\gamma_t \leq (1-\lambda)^t$. Hence, $\gamma_t \to 0$ and there is convergence to the non-coffee steady-state.

Proof of Proposition 2

Consider an equilibrium of the no-frost economy. Since $\pi^C(0) > \pi^N$, then all farmers who inherit trees grow coffee: $\sigma_t^I = 1$ for all t. A sufficient condition for all equilibria to converge to the coffee steady-state is that $\sigma_t^D = 1$. This will happen if:

$$\pi^C(\gamma_t) - \pi^N > \kappa - \delta(1-\lambda)\big(U_{t+1}^I - U_{t+1}^D\big)$$

If $\pi^C(\gamma_t) - \pi^N > \kappa$, the condition above will necessarily be satisfied.

Now assume that $\pi^C(1-\lambda) - \pi^N > \kappa$. Define:

$$\gamma^* = \inf\{\gamma \in [0,1] | \pi^C(\gamma(1-\lambda)) - \pi^N > \kappa\}$$

Notice $\gamma^* < 1$. Now consider $\gamma_0 > \gamma^*$. This implies that $\sigma_1^D = 1$ is optimal, so $\gamma_1 = 1$. By induction, the same argument shows that $\sigma_t^D = 1$ for all $t > 1$. Hence, the equilibrium converges to the coffee steady-state.

B.4 Agglomeration Economies within the Coffee Industry

Data

The mapping of coffee tree coverage in Paraná was performed by *Conab* based on 10m-resolution satellite images of mission SENTINEL-2. Coffee was identified through a spectral analysis over different periods of the year. The identified coffee polygons were checked against Google Earth images, previous information on the spatial distribution of the coffee industry from IBGE and the state government, and in many cases confirmed on the ground. In order to calculate the share of each cell that was covered with coffee, I convert the polygons into a raster of 1 arc-sec $\times$ 1 arc-sec resolution.

The digital elevation model I use was the ALOS Global Digital Surface Model, a 1 arc-sec $\times$ 1 arc-sec raster released by the Japan Aerospace Exploration Agency (JAXA). For each grid cell, I calculate the median altitude and slope. The south-facing share is the share of non-flat points in the cell that have an aspect between 90° and 270°.

The location of rivers is from version 1.3 of the *Base Hidrográfica Ottocodificada*, a dataset released by the *Agência Nacional das Águas* (ANA). The location of the roads is from the *Base Cartográfica do Brasil ao Milionésimo* of 2016, a detailed map released by IBGE. The soil data I use is a soil map of Paraná from *Embrapa Solos* in 2007 (Embrapa 2007). *Terra Roxa* refers to red latosols, which is any soil with the following codes: LVdf1-11, LVef1-3, LVed1-23, and LVe1-2.

Finally, the *Cadastro Ambiental Rural* (Rural Environmental Registry, CAR) is a mandatory registry for farms in Brazil. Its purpose is to enforce environmental and land use restrictions. I downloaded the data separately for each municipality in Paraná in July 2019. I use the information from the shapefiles *Área do Imóvel*.

Sample Selection

To select the sample, I follow the following procedures:

1. In order to exclude urban areas from the sample, I match each cell with the nearest census block (*setor censitário*) from the 2010 population census by IBGE. I then exclude all the cells for which the closest block was urban and that the distance to the closest block was at most 0.5 km. This procedure is used only for the 30 arc-sec $\times$ 30 arc-sec grid cells, not for farms in the CAR.

2. To select areas that are suitable for coffee production, I exclude cells and farms with a median altitude below 300 m and to the south of latitude 25S. Since coffee is less suitable in these regions, south-facing terrains should not matter for coffee planting decisions. I also exclude observations that were at a distance of more than 40 km to any coffee tree in Paraná; unobservable fundamentals could have made these areas not suitable for coffee production. The results are robust to the latter criterion; see columns (4) and (5) of Table B.3.

3. Since regressions are weighted by area and farm size distribution has a long right tail, I exclude farms in the top 1% of the size distribution.

Inference

Due to the large number of observations, standard errors as in Conley (1999) are computationally impractical. However, the residuals will be spatially correlated by construction. Since the variable of interest is the south-facing share in a 5 km neighborhood, a minimum criterion for inference is to allow spatial correlation between each pair of observations that are at a distance of up to 10 km. Clustering at a grid does not satisfy this condition. The solution I adopt is to use multi-way clustering (Cameron *et al.* 2011) according to four grids. The first grid is a 0.5 $\times$ 0.5 degree grid; each cluster is roughly a square of 53 km side. I then construct three other grids by translating the original grid: the first translation is vertical in 0.25 degrees of latitude, the second is horizontal in 0.25 degree of longitude, and the third is diagonal in 0.25 degrees of latitude and longitude. This approach allows for arbitrary correlation between each pair of observations that are at a same cell of any of the four grids.

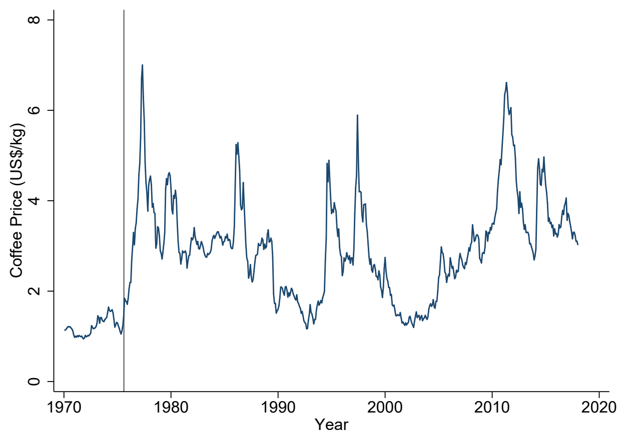
Robustness

I consider a variety of robustness exercises; they are shown on Table B.3. In column (1), instead of the share of a cell that is covered by coffee, the dependent variable is the presence of any coffee tree. The results are qualitatively similar. In column (2), I vary instead the definition of what is the neighboring area; instead of using a 5 km radius, I use a 10 km one. Note the coefficients barely change, which indicates that agglomeration economies do not decline much with distance. In column (3), I estimate a tobit model instead of a linear one by censoring the share at zero and one. Note the coefficient on the neighboring south-facing share is now larger, which makes sense as the identifying variation in the censored model is concentrated in sample regions that are more likely to produce coffee. Finally, in columns (4) and (5) I use a different criterion to select the sample, by including cells that are, respectively, at a distance of 20 km or 60 km to the closest coffee tree. Results are similar.

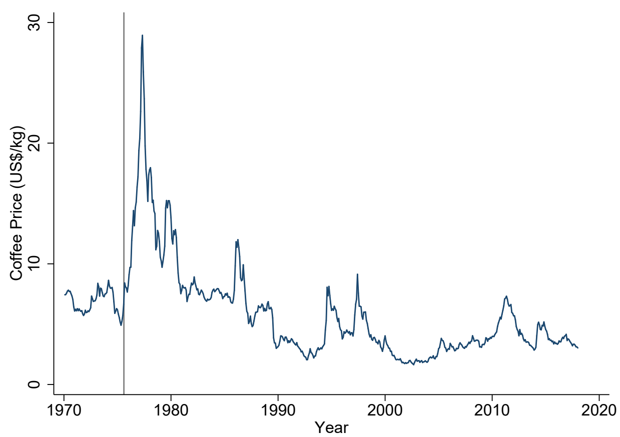
B.5 Appendix Figures and Tables

Figure B.1: Coffee prices

(a) Nominal prices



(b) Inflation adjusted (December 2017 dollars)



Note: Monthly prices from the International Coffee Organization (ICO), represents the average between the markets in New York and Bremen/Hamburg. Inflation-adjusted by the United States CPI in the bottom graph. The vertical line indicates the last observation before the frost (June 1975).

Figure B.2: Main coffee producing states of Brazil in 1975

Note: The map shows the Brazilian states (2010 borders). Paraná is in dark green; São Paulo and Minas Gerais in light green.

Table B.1: Coffee tree concentration - Theil indexes

	All States (total)	Paraná	São Paulo	Minas Gerais	Decomposition: between states	within states
	(1)	(2)	(3)	(4)	(5)	(6)
1970	1.354	0.949	0.805	1.644	0.263	1.091
1975	1.252	0.901	0.809	1.689	0.126	1.116
1980	1.196	0.972	0.777	1.569	0.039	1.157
1985	1.260	1.099	0.868	1.487	0.010	1.250
2006	1.796	1.400	1.879	1.605	0.225	1.571

Note: Data from the agricultural censuses of 1970, 1975, 1980, 1985 and 2006. Column (1) shows the Theil Index computed at the municipality level for all states. Columns (2) to (4) show the index for each of the three states. In column (5), I show the between-state index (using the states as the unit to construct the index); column (6) is the within-state component, defined as (1) minus (5); it is a weighted average of the indexes in (2), (3) and (4).

Table B.2: Labor by crop area

	Coefficient	p-value (= coffee)
	(1)	(2)
Harvested area: coffee	0.535***	
	(0.117)	
Harvested area: beans	0.657***	0.617
	(0.194)	
Harvested area: maize	0.268***	0.022
	(0.044)	
Harvested area: soy	0.263***	0.028
	(0.065)	
Harvested area: sugar	0.135***	0.002
	(0.041)	
Harvested area: wheat	-0.073	0.017
	(0.195)	
Harvested area: other crops	-0.032	0.005
	(0.103)	
Total farm area	0.010	
	(0.006)	
R-squared	0.964	
Observations	778	

Note: Labor and total farm area from the 1975 agricultural census, harvested area in coffee, beans, maize, soy, sugarcane, and wheat from the *Produção Agrícola Mensal*. Observations are weighted by the municipality area. Robust standard errors clustered at the micro-region. Column (1) reports the coefficients, while column (2) reports the p-value of the test that the coefficient for the harvested area of crop is the same as the coefficient for the harvested area of coffee. Statistical significance denoted by: * 10%, ** 5%, *** 1%

Table B.3: Robustness: south-facing terrain and coffee tree coverage

	Any Coffee	10km neighbors	Tobit	20km sample	60km sample
	(1)	(2)	(3)	(4)	(5)
Share of terrain facing south	-0.0133***	-0.0027***	-0.0344***	-0.0029***	-0.0024***
	(0.0037)	(0.0011)	(0.0126)	(0.0012)	(0.0010)
Share of neighboring area facing south	-0.3473***	-0.0580**	-0.8903***	-0.0469***	-0.0361**
	(0.1153)	(0.0267)	(0.3294)	(0.0192)	(0.0161)
R-squared	0.053	0.035		0.035	0.027
Observations	120,118	120,118	120,118	99,399	123,927

Note: See Appendix B.4 for a description of the data. The dependent variable is the share of a cell or a farm that is covered by coffee, except in column (1), when it is an indicator for any coffee tree in the cell. Control variables include own and neighboring averages of altitude, slope, share within 500m of a river, and share of *Terra Roxa*, as well as latitude, latitude squared, and distance to a road. Observations weighted by area. Robust standard errors in parentheses are multi-way clustered according to four different grids of 0.5° × 0.5 °; see Appendix B.4 for a description of the inference procedure. Statistical significance denoted by: * 10%, ** 5%, *** 1%.

B.6 Data Appendix

Agricultural Census

This paper uses the *Censos Agropecuários* by IBGE for the years of 1960, 1970, 1975, 1980, 1985, 1995, and 2006. I collected information for the states of Paraná, São Paulo, and Minas Gerais. Information from the 1995 and 2006 censuses may be obtained online. For the 1960 to 1985 censuses, I digitized the original tables published by IBGE, with information aggregated at the municipality level. The variables digitized include the number of farms, total farm area, number of coffee producing farms, area of coffee producing farms, number of coffee mature and young coffee trees, total land value, total number of workers employed by farms, total number of temporary workers employed by farms, and a variety of expenditures. Stock variables, such as the number of coffee trees, are relative to December of the census year. Flow variables, such as expenditures, are relative to the whole census year. I exclude a few municipalities for which there is no farm activity in some of the census years. These municipalities are Cubatão, Praia Grande, São Sebastião, São Vicente, and Águas de São Pedro.

When calculating the non-labor expenditures per worker, I excluded land rent payments from the expenditures. The labor share is the wage bill (including share-cropping payments to workers) divided by total expenditures minus land rent.

Population Census

I use samples of households who answered an extended questionnaire in the population censuses (*Censos Demográficos*) of 1970, 1980 and 1991. The microdata from these census years may be purchased from IBGE. These samples are representative at the municipality level. I benefited from the STATA applications developed by the Data Zoom team at the *Pontifícia Universidade Católica* of Rio de Janeiro when preparing the 1970 and 1991 census data. In order to calculate employment, I restricted the sample to individuals who declared to have worked in the past 12 months. The definition of the economically active population changed from the 1991 to the 2000 census, so I did not use information from the censuses of 2000 and 2010.

In order to calculate average earnings, I restricted each of the samples to economically active individuals between 20 and 55 years old. This minimizes the effect of non-labor income, which I cannot separate from total income in the 1970 census.

Other data sources

In this paper, I also used the *Produção Agrícola Mensal* (PAM), a series of production, harvested area, and revenue for many crops that is available at the municipality level since 1974. A similar series is the *Pesquisa da Pecuária Municipal* (PPM), from where I know the cattle stock since 1974. Coffee production from 1960 to 1973 is from the *Anuário Estatístico do Brasil*, except for 1971 and 1972, where it is from the *Anuário Estatístico do Café* of 1972. The data sources used for the localization economies exercise are described in Appendix B.4.

Sample Definition

The baseline sample consists of municipalities in the states of Paraná and São Paulo. I aggregate municipalities into *áreas mínimas comparáveis* (AMCs) that satisfy the 1970 borders, following Reis et al. (2011). I exclude municipalities in the micro-regions that contain the three state capitals: Curitiba, São Paulo, and Belo Horizonte. I also exclude Votorantim, for which there is no 1970 census sample data, and Praia Grande, Cubatão, São Vicente, and Águas de São Pedro, for which there is no agricultural census data for some of the years. The baseline sample consists of 778 AMCs.

Migration Rates

In this subsection, I discuss how to construct the migration rates for the intercensal period of 1970-1980. In order to do so, I use the restricted samples of the 1970 and 1980 population censuses. I restrict the population of interest to individuals who were between 20 and 59 years in 1975, the year of the frost. This selection is for two reasons: **(i)** this is the cohort who is most likely to be economic active; **(ii)** the construction of the in-migration and net migration rates require the use in each municipality of survival rates for the whole country, and survival rates for children and for the elderly are more likely to vary spatially than for prime age individuals. Unfortunately, the publicly available data from the 1980 census does not allow to track migrants to their origin municipalities or states, so it is unfeasible to construct a direct measure of out-migration. In this way, I construct out-migration as the difference between in-migration and net migration. All rates use the 1970 cohort population as its base. Furthermore, the survival rates refer to Brazilian born individuals, to avoid measurement errors due to international migration.

In-migration Rates The sample of the 1980 population census contains information on how many years the individual resides in the current municipality. This variable indicates any number between 0 and 5, if she moved in the last five years, or else it indicates if her arrival was between 5 and 10 years. In order to obtain the number of in-migrants, I multiply each migrant by her sample weight and the survival rate for her age between the year of her arrival and 1980. If the migrant arrived between 5 and 10 years, I take the average of the survival rates of 10 and 5 years.

Net Migration Rates The net migration rates are calculated by an average intercensal cohort-component method. See Morrison et al. (2004) for a discussion. For a specific age a, let P_a be the population in 1970, P'_a be the population in 1980, and s_a the (national) survival rate of age a individuals from 1970 to 1980. The number of net migrants of age a, N_a, is calculated as follows:

$$N_a = \frac{(P'_a - s_a \cdot P_a)}{\sqrt{s_a}} = \left(\frac{P'_a}{s_a} - P_a\right) \cdot \sqrt{s_a}$$

Software

The statistical analysis in this paper was performed in Stata. I greatly benefited from the following three user-written packages: **(i)** *geonear*, written by Robert Picard; **(ii)** *reghdfe*, see Correia (2014); **(iii)** *datazoom_censo*, from the Economics Department of the *Pontifícia Universidade Católica* of Rio de Janeiro.

References

[1] Agência Nacional das Águas. 2013. *Base Hidrográfica Ottocodificada 1.3* https://dadosabertos.ana.gov.br/datasets/0e1dd8d2169e49d189ed04184f1a1e3b_0 (last accessed November 3, 2020).

[2] Brülhart, M. & Traeger, R. (2005). An account of geographic concentration patterns in Europe *Regional Science and Urban Economics*, 35(6), 597-624.

[3] Cameron, A. C., Gelbach, J.B., & Miller, D. L. (2011). Robust inference with multi-way clustering. *Journal of Business and Economic Statistics*, 29(2), 238-249.

[4] Companhia Nacional de Abastecimento. 2017. *Mapeamento Agrícola do Café no Paraná*. https://www.conab.gov.br/info-agro/safras/mapeamentos-agricolas/item/11309-cafe-parana-safra-2017 (accessed July 20, 2019).

[5] Conley, T.G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1), 1-45.

[6] Correia, S. (2014). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Boston College Department of Economics*.

[7] Embrapa. 2007 *Mapa de solos do estado do Paraná.*

[8] IBGE. 1960, 1970, 1975, 1980, 1985. *Censo Agropecuário.* https://biblioteca.ibge.gov.br/ (last accessed November 3, 2020).

[9] IBGE. 1995, 2006. *Censo Agropecuário.* https://sidra.ibge.gov.br/pesquisa/censo-agropecuario/censo-agropecuario-2017 (last accessed November 3, 2020).

[10] IBGE. 1970, 1980, 1991. *Censo Demográfico*.

[11] IBGE. 1974-1990. *Produção Agrícola Municipal.* https://sidra.ibge.gov.br/pesquisa/pam/tabelas (last accessed November 3, 2020).

[12] JAXA. 2019. *ALOS Global Digital Surface Model "ALOS World 3D - 30m (AW3D30)"* https://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm (accessed July 20, 2019).

[13] Ministério da Meio Ambiente. 2019. *Cadastro Ambiental Rural* https://www.car.gov.br/publico/imoveis/index (accessed July 20, 2019).

[14] Morrison, P., Bryan, T.M., & Swanson, D.A. (2004). Internal migration and short-distance mobility. In *The Methods and Materials of Demography*, 2nd Edition, ed. Siegel, J.S. & Swanson, D.A..

[15] Reis, E., Pimentel, M., Alvarenga, A.I., & Santos, M.C.H. (2008). Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. *Rio de Janeiro, Ipea/Dimac.*