UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE CIÊNCIAS SOCIAIS APLICADAS

PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

GUILHERME ZAMBALDE PORTELA CUSTÓDIO

ESSAYS ON ECONOMIC GROWTH AND DEVELOPMENT

Tese de Doutorado

Recife

2023

GUILHERME ZAMBALDE PORTELA CUSTÓDIO

ESSAYS ON ECONOMIC GROWTH AND DEVELOPMENT

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Doutor em Economia. Área de concentração: Teoria Econômica.

Orientador: Prof. Dr. Paulo Henrique Pereira de Meneses Vaz

Recife

2023

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

C987e Custódio, Guilherme Zambalde Portela
Essays on economic growth and development / Guilherme Zambalde Portela Custódio. - 2023.
123 folhas: il. 30 cm.

Orientador: Prof. Dr. Paulo Henrique Pereira de Meneses Vaz.
Tese (Doutorado em Economia) – Universidade Federal de Pernambuco, CCSA, 2023.
Inclui referências e apêndices.

1. Crédito financeiro. 2. Rodovias. 3. Agricultura. I. Vaz, Paulo Henrique Pereira de Meneses (Orientador). II. Título.

336 CDD (22. ed.) UFPE (CSA 2023 – 101)

GUILHERME ZAMBALDE PORTELA CUSTÓDIO

ESSAYS ON ECONOMIC GROWTH AND DEVELOPMENT

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Doutor em Economia. Área de concentração: Teoria Econômica.

Aprovado em 31/03/2023

BANCA EXAMINADORA

Prof. Dr. Paulo Henrique Pereira de Meneses Vaz (Orientador)
(Universidade Federal de Pernambuco)

Prof. Dr. Bladimir Carrillo Bermudez (Examinador Interno)
(Universidade Federal de Pernambuco)

Prof. Dr. Rafael da Silva Vasconcelos (Examinador Interno)
(Universidade Federal de Pernambuco)

Prof. Dr. Dieison Lenon Casagrande (Examinador Externo)
(Universidade Federal de Santa Maria)

Prof. Dr. Paulo de Andrade Jacinto (Examinador Externo)
(Universidade Federal do Paraná)

Acknowledgements

Esta tese de doutorado é fruto da minha dedicação, mas é em grande parte fruto da paciência das pessoas ao meu redor, que esperaram pelo meu tempo e pelas minhas dificuldades pessoais. Algumas pessoas que conheci ao longo da vida possuem um traço de obstinação e controle em face de adversidades externas as quais eu sequer consigo imaginar. O meu caso é o exato contrário: o interno é caótico, e o externo é que apruma. Deve ser bom ter os dois à favor...

Agradeço ao apoio da minha família, sobretudo. Meu pai, José Roberto, minha mãe, Josélia, e meu irmão, Fernando. Pela presença, apoio emocional e suporte financeiro. Nos últimos 3 anos voltamos a conviver de perto. Se minha família não fosse tão estável, eu não teria muita chance na vida.

Agradeço à minha noiva, Aline, logo, logo esposa, por aturar essa longa jornada que eu insisto em estender. Por todo carinho, amor, respeito e paciência, obrigado, More.

Agradeço ao meu orientador, Paulo Vaz, antes de tudo, pela paciência (maior impossível), pelos incentivos, mas sobretudo por ser uma referência profissional próxima, junto a todos os bons professores do PIMES, a maioria da safra mais nova. Obter uma formação de ponta no Brasil não está disponível em qualquer canto, e fico feliz de ter recebido algo de excelente nível, e que me deixa em boa posição para construir uma carreira legal daqui para frente. Obrigado à todos pela oportunidade! Foi duro, mas foi legal.

Agradeço aos inúmeros amigos que fiz ao longo da trajetória em Recife. Morei numa pensão com vários alunos de pós e de graduação. Tinha papo para tudo, e é legal ter sempre alguém para jogar um papo fora, até numa madrugada, com insônia (o outro provavelmente também estava...).

Aos amigos do passado, esses eu estou em débito pesado. Fiquei muito distante fisicamente e faltei com atenção. Mas dou um jeito de compensar daqui para frente. São muitos para citar nominalmente...

Aos amigos do PIMES, em especial (por ordem alfabética, hein?) Andrews, Daniel, Gabriel, Leandro (esse não bebe, chatão) e Rubens, que estiveram presentes desde o início, e Caíque, Robson e Yuri, que apareceram depois. Muito obrigado pela amizade. E o pessoal do Gappe também, que convivi menos. Saudades de todos. Por fim, agradeço ao CNPq e ao governo e povo brasileiro por financiar meus estudos. Agora é hora de retribuir um pouquinho.

Obrigado por tudo!

Abstract

This doctoral thesis is comprised of two essays on the matters of economic growth and development economics. In the first essay, I use a reclassification of firms done by the Brazilian Development Bank in 2003 as a natural experiment to causally identify the impact of a shift in credit conditions on capital, labor and scale wedges measuring the misallocation of resources. I find that firms reclassified as small, coming from medium, suffered a shift of - 0.261 in their average log capital wedge 6 years after treatment. Firms reclassified as medium, coming from large, had its largest impact at - 0.123 log average capital wedge 2 years after treatment, but the gap was closed again 6 years after treatment. That change in the average distribution of log capital wedges, particularly when one starts to face favored credit conditions from the small cohort, suggests that the BNDES size classification can potentially have noticeable impacts in resource allocation through this mechanism, particularly capital, either positively or negatively. I then use my difference-in-differences result with the assumption that the log capital wedge around zero is a good approximation of the efficient allocation of funds. I find that new small firms moved symmetrically from the under-invested position to the over-investment one, with potentially no impact on allocational efficiency. New medium firms went from around efficiency into over-investment territory. I also calculate that if not for favored BNDES credit policy, small firms would be far from the efficiency line. Results on the misallocation exercise are heavily dependent on stringent assumptions, and should be taken with caution. The second essay explores the roll-out of an unusually large scale program to pave feeder roads in Minas Gerais, Brazil, from 2002-2014. I run an event study using the new estimators in the literature, robust to heterogeneous and dynamic treatment effects, to identify the causal impact of paving feeder roads on outcomes of labor markets and firm dynamics. I find that there was a large negative demand shock in the agricultural sector, particularly in food crops, which generated large reductions in planted area and yields, combined with an increase in resignations, mostly initiated by the firm, and a decrease in wages for agricultural workers. I also establish statistically significant but not economically meaningful entry and exit of firms. Services also faced some increase in resignations, although without a lowering of wages and with lower exit rates, suggesting a different dynamics than agriculture. My results suggest that better access to remote towns can have noticeable impacts on agricultural workers. They end up unemployed, sent to informality or they are forced to migrate to other town. Short to medium run transition periods should increase caution among policymakers.

Keywords: Misallocation, Subsidized Credit, Feeder Roads, Agricultural Dynamics (**JEL:** O16, O47, O12, O13, O18, R58, R11)

Resumo

Esta tese de doutorado compreende dois ensaios sobre crescimento econômico e desenvolvimento econômico. No primeiro ensaio, eu utilizo uma reclassificação de porte de firmas realizada pelo Banco Nacional de Desenvolvimento Econômico e Social em 2003 como um experimento natural para a identificação causal do impacto de mudanças nas condições de crédito nas distorções de capital, trabalho e escala medindo o *misallocation* de recursos. Eu encontro que firmas reclassificadas como pequenas, vindo da condição de médias, sofreram uma mudança de - 0.261 na sua medida de distorção de capital em log 6 anos após o tratamento. Firmas reclassificadas como médias, vindo de grandes, tiveram no máximo uma mudança de - 0.123 na sua medida de distorção de capital em log 2 anos após o tratamento, mas o *gap* se fechou novamente 6 anos após o tratamento. Essa mudança na distribuição média das distorções de capital em log, particularmente quando se passa a receber crédito do coorte pequenas, sugere que a classificação de porte do BNDES pode gerar resultados notáveis na alocação de recursos, particularmente do capital, tanto positivas ou negativas. Eu então uso meus resultados de diferenças-em-diferenças, juntamente com a hipótese de que a distorção de capital próxima de zero aproxima uma melhor alocação de fundos. Eu encontro que novas firmas pequenas se moveram simetricamente desde a posição de sub-investimento para a posição de sobre-investimento, potencialmente sem impacto para a alocação de recursos. Novas firmas médias se moveram desde próximo à eficiência até o território de sobre-investimento. Eu também calculo que caso as firmas pequenas não fossem favorecidas pela política de crédito do BNDES, elas estariam longe da linha de eficiência. Os resultados do exercício de *misallocation* são fortemente dependentes de hipóteses restritivas, e devem ser tomadas com cuidado.

O segundo ensaio explora o desenrolo de um programa incomumente grande de pavimentação de estradas de acesso em Minas Gerais, Brasil, de 2002-2014. Eu rodo um *event study* usando novos estimadores da literatura, robustos ao tratamento dinâmico e heterogêneo, para identificar o impacto causal da pavimentação de estradas de acesso em variáveis de mercado de trabalho e dinâmica de firma. Eu encontro que há um forte choque negativo de demanda no setor agrícola, particularmente na produção de *food crops*, que gerou grandes reduções na área plantada e nas quantidades produzidas, combinado a um aumento dos desligamentos, sobretudo os iniciados pelo empregador, e uma redução no salário para trabalhadores agrícolas. Eu também estabeleço entradas e saídas de firma que são estatisticamente significantes, porém economicamente desimportantes. Serviços também sofreram um acréscimo nos desligamentos, ainda que sem uma redução nos salários e uma menor taxa de saída de firmas, sugerindo uma dinâmica diferente da agricultura. Meus resultados sugerem que um melhor acesso a cidades remotas pode ter um impacto notório nos trabalhadores agrícolas. Eles podem acabar desempregados, na informalidade ou forçados a migrar para outra cidade. O período de transição no curto e médio prazo deve aumentar o cuidado entre os formuladores de política pública.

Palavras-chave: *Misallocation*, Crédito Subsidiado, Estradas de Acesso, Dinâmica da Agricultura (**JEL:** O16, O47, O12, O13, O18, R58, R11)

List of Figures

List of Tables

Contents

CHAPTER 1

Introduction

This doctoral dissertation (*Tese de doutorado*, in Portuguese) is comprised by two essays focused on the fields of economic growth and of development economics.

The growth literature has identified four main channels by which a nation can grow. Or, at least, what we call the proximate causes: the accumulation of physical capital, the accumulation of human capital, technological progress and the increase in allocational efficiency [1]. The first three have been studied for a longer time. Solow and Neoclassical growth models have long suggested the role of capital accumulation in economic growth. It has also suggested its limitation as the main lever of economic growth. Further research focused on human capital and technological progress. Human capital is probably the area that has benefited the most from microdata. Apart from older research from Gary Becker or James Heckman, for instance, there is a plethora of recent research on the causal impact of better education or nutrition, for instance. As for technological progress, both endogenous growth models initiated by Paul Romer and Schumpeterian growth models initiated by Aghion and Howitt have already 30+ years. Many of the aforementioned researchers have already won their Nobel Prizes. Although allocational efficiency seems to be quite relevant in the eyes of the researcher of today, it has not been always like that. Perhaps its study required a larger availability of data and computer power. Since late 2000's, there has been a larger attention devoted to such matters. We learned that the dispersion in marginal products of factors are responsible for a good chunk of the developing world retardation. We still don't have a clear picture how much, though, mostly because of methodological constraints in data and modelling, and we also don't know which

[1] There is also the study of fundamental causes: institutions, social norms and geography. These would be the more profound causes of the proximate ones.

constraints are the most important for each country, leaving policymakers still somewhat blind about what levers to pull. There are many channels relevant for this new area of research.

In my first essay, I attempt to isolate one of the candidates for misallocation, the financial channel, through an indirect manner. That is, I use a natural experiment in an attempt to identify the impact of a change in credit conditions on generic wedges measuring misallocation. The change in credit conditions are due to a reclassification of firms done by the Brazilian Development Bank (BNDES) in 2003. The Bank was responsible for granting a significant amount of subsidized credit in the Brazilian economy, specially from mid-2000's to mid-2010's.

Understanding how development banks can help or hinder the optimal allocation of resources is important for developing nations. For one thing, they are the most likely to have large amounts of funds being channeled through government agencies. It could be the case that developing nations lack the necessary financial development, thus the necessity of its operations. On the other hand, a heavy handed policy could very well distort the allocation of funds. I attempt to give a modest contribution by capturing how much wedges move when faced with better credit conditions. If firms that received a shock in credit due to being reclassified to a lower size category have seen their wedges move by a significant margin, that would suggest that the credit policy of segregating firms by size and giving them different credit conditions have a potential impact to affect the allocational efficiency, either positively or negatively. Adressing the issue of whether allocational efficiency improved or not due to the credit shock is harder, and depends on invoking some assumptions about the distribution of the other distortions affecting wedges. I perform a misallocation exercise under those strict assumptions as a primary attempt to advance in the matter.

The second essay follows a more development-style route. The causal inference revolution has strongly impacted the development literature. This literature looks at intervention (mostly governmental) in the provision of goods (mainly public goods) to destitute areas. Some study health outcomes from the betterment of water supply, treated sewage, bathrooms, access to medication, treatment, etc. Others will study education outcomes from the betterment of schools stemming from better hirings, class size, equipment, etc. There are also those interested in labor markets and firm dynamics stemming from the impact of some government provision, mainly infrastructure. It is in this spirit that I set to study the impact of paving feeder roads on labor markets and firm dynamics.

I take advantage of a relatively large program by the government of the Brazilian state of Minas Gerais, responsible to expand access to a paved feeder road to the remaining 217 municipalities without them, starting in 1998. The program rolled-out until 2014 and is quite conducive to a staggered differences-in-differences (event study). Using the new developments in the event study econometric literature, I can estimate causal parameters for outcomes related to labor markets and firm dynamics.

Infrastructure projects in general are hard to study under causal inference. They generate a small number of observations, and the roll-out is generally the best option for identification. That is certainly the case of the rural/feeder road literature. Since both under-powered statistics and the lack of a good research design has affected the old literature, this relatively large program by the Minas Gerais state is a good opportunity to make progress in assessing the impact of paving feeder roads to remote towns.

There has been some good research published recently on an Indian experiment that was very large in treatment units and was suitable to both differences-in-differences and a fuzzy RD design. One good experiment is not enough to help policymakers in understanding the dynamics of such matters. In fact, rural/feeder road building was seen before as generating mostly positive and large results, whereas the literature in the Indian program has found more modest and even null results. Apart from that, Minas Gerais state is in a different development level relative to India, which further justify the research, since mechanisms might be quite different in this case.

In addition, paving rural/feeder roads is a noticeable part of funds granted by the World

Bank and the Inter-American Development Bank, for instance. Policymakers need to be better informed when taking the step to execute such types of programs. The fact that I have access to detailed information about the formal labor market already gives my work a relevant differential against others.

Apart from the many interesting development questions generated by an increase in connectivity, it is of notice that the largest economic impacts are mediated through trade. Trade analysis has long been affected by the traditional comparative advantage model, where the transition period effect and income redistribution is abstracted from. Even though the theoretical understanding of such matters come from long ago, the empirical counterpart is more recent. I attempt to further our understanding in those matters too.

The essays follow in the next two chapter.

CHAPTER 2

Misallocation and Access to Subsidized Credit: How Sensible are the Wedges to a Credit Shock?

2.1 Introduction

The misallocation of resources is one of the possible explanations for the income disparity between high and low income countries. Although the notion of resource misallocation has existed since Alfred Marshall, Jules Dupuit and Arnold Harberger (Baqaee and Farhi, 2020), whose work in the field was done in the XIXth and early XXth centuries, only recently, since the works of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), the literature has been revisited with the help of modern mathematical modelling, access to microdata and cheap computational power.

Restuccia and Rogerson (2017) have suggested that there are two avenues of investigation on the matters of misallocation: the indirect approach and the direct approach. The indirect approach is the one related to Hsieh and Klenow (2009)'s seminal paper, where a model allows for the inference of input distortions, or wedges, preventing the optimal allocation of resources. By accounting for the dispersion in wedges, one can perform cross-country comparisons using a benchmark country as allocational reference. This method has yielded several papers identifying relevant misallocation of resources in the developing world, for agricultural, manufacturing and service sectors (Busso et al., 2013; Kalemli-Ozcan and Sørensen, 2016; Hsieh and Klenow, 2009; de Vries, 2014; Dias et al., 2016; Adamopoulos and Restuccia, 2014; Chen et al., forthcoming; Vasconcelos, 2017). Some caveats to this approach exist, though. One shortcoming is the fact that the indirect approach may capture the existence and magnitude of misallocation, but cannot account for the sources of misallocation. Another problem is that, even though the preferred models are fairly general in this literature, the assumptions are still somewhat restrictive. This has come to light with works from Haltiwanger et al. (2018) and Baqaee and Farhi (2020), which identify how demand and supply structures, and input-output networks might affect the implied misallocation measures. The work by Bils et al. (2021) shed light on the mismeasurement problem. As for the direct approach, the researcher establishes a structural model to capture explicitly the mechanism responsible for the possible misallocation. The problem here is that one needs a good measure on the source of misallocation, which is generally not available. In our case at hand, one would need information about the lending interest rates available to specific firms (Restuccia and Rogerson, 2017).

Many papers have tried to connect specifically the financial sector with misallocation. Some papers have established that financial frictions are more relevant for technology adoption and the misallocation of entrepreneurial talent, instead of in the allocation of funds into already established firms. Firms with high productivity but facing financial frictions are able to avoid that constraint by accumulating funds and resorting to self-finance, but entrepreneurial talent in the traditional sector would face a high fixed cost to enter in the modern, technology-intensive sector. Such a shift might not be easily financed in developing countries (Buera et al., 2011; Midrigan and Xu, 2014; Buera and Shin, 2013; Gilchrist et al., 2013).

A recent paper by David and Venkateswaran (2019) attempted to impose more structure in the misallocation exercise and disentangle some of the sources of capital misallocation through the tracking of the moments in firm-level data. That made possible to identify the contribution of transitory factors, like adjustment costs and uncertainty, the contribution of markups and technological heterogeneity, and the size or productivity-dependent component. In the United States, the summation of transitory and technological/markup component is quite relevant (2/3rds of the dispersion in average revenue product of capital), but that is not true for

China, where the unidentified size and productivity-dependent component is what accounts for 2/3rds of the dispersion in the average revenue product of capital. That may suggest that misallocation in advanced economies are largely due to demand and supply characteristics, whereas developing economies potentially face true distortions, or at least policy-relevant distortions, i.e. the ones that could be tackled by removing either government or market frictions. Wu (2018) is a paper that attempts to disentangle financial frictions arising from market imperfections and from other types of policies distorting capital allocation in China. Although the importance of financial frictions are much larger in relation to the developed world, policy distortions are still more relevant an explanation for the dispersion of marginal revenue product of capital.

Brazil is a country with high levels of general misallocation and capital misallocation (Vasconcelos, 2017), specifically. If China is a better reflection of Brazil's financial markets than the developed world, one would expect two things: (i) market failures could be responsible for a significant share of capital misallocation; and (ii) government policy distortions could be responsible for another large share of the dispersion in marginal revenue products of capital. In fact, policy aiming at solving market failures, if badly designed, could worsen misallocation. Just like India and China (Banerjee and Duflo, 2014; Hsieh and Klenow, 2009; Wu, 2018), Brazil also has channelled a significant share of its credit through the public sector. Since mid 2000's, the Brazilian Development Bank (BNDES) started to gain a larger role on credit allocations. General reasons to implement this policy change are given in their documents, but information on specific goals and policy evaluation are scarce. One would like to know, in the best case scenario, how to measure the impact of this policy change on capital allocation. Full data on market interest rates are not easily available at the firm-level, so the indirect approach is the substitute available for measuring wedges preventing optimal capital allocation. Since the indirect approach is unable to disentangle the sources of capital misallocation, we need to couple the generic wedges' evolution with an identification strategy that provides an exogenous shift in credit conditions.

This paper attempts to infer the impact of improved conditions for subsidized credit on capital, labor and scale wedges, estimated through Hsieh and Klenow (2009)'s model using a natural experiment from Brazil, where a reclassification of firms' size by the Brazilian Development Bank [1] (BNDES) in 2003 allowed for a plausibly exogenous shock in credit conditions for firms reclassified to smaller categories. This shift provides an excellent opportunity for an identification by differences-in-differences, since both common trends and the exogeneity of the shock guarantees a clean identification. Under the assumption that mismeasurement arising from modelling and data constraints are not different in the treated and control groups at each time, I can follow the evolution of wedges after the shock to capture the impact of the change in credit conditions on them. By doing so, I can, as a first exercise, quantify the shift in wedges in relation to the overall distribution of wedges, gathering an idea of the impact an improvement in access to credit has in shifting targeted firms across the distribution. As a second exercise, I can observe how the position of targeted firms' wedges relate to the optimal allocation.

This paper is, then, in the spirit of both the direct and the indirect approach. Since the wedges are coming from a general model, I have generic measures of misallocation, like in the indirect approach. Still, the identification strategy allows me to infer changes in the generic measures through a specific channel, the credit condition, which is the objective of the direct approach. In doing so, I follow the lead of Bau and Matray (2023), which used a natural experiment on access to foreign credit to identify changes in wedges, although in their exercise they were able to compute a better measure of allocational change.

I find that the reclassification of firms, which had a quite large impact on the investment rate of firms in treated groups, especially for the new small firms, coming from the medium category (Cavalcanti and Vaz, 2017), had the expected negative impact in capital wedges, and that impact was larger than the impact on both the labor wedges and the scale wedges. New

[1] *Banco Nacional de Desenvolvimento Econômico e Social* in Portuguese

small firms shifted - 0.261 in average log capital wedge in the last year available after treatment. This is not a very large shift to the left, despite the apparent large increase in the investment rate for those firms, but it might still able to generate some fat tail in the overall distribution of capital and scale wedges depending on the original position of the distribution of wedges by cohorts, and that would match, at least partially, with the overall distribution of capital and scale wedge, suggesting the BNDES credit policy would be partially responsible for the larger misallocation present at small firms. A naive approach to compute misallocation against the capital wedge efficiency line, though, suggests the BNDES heavy handed credit policy towards the small group might have actually helped to compensate for financial frictions affecting that cohort. Results for the new medium firms, coming from the large group, are not as impactful. Average log capital wedges shifted - 0.123 2 years after treatment, but that result have not sustained itself through time, mainly because credit conditions became similar for each group after some time.

This chapter is organized in the following manner: section 2.2 explains the policy environment of the reclassification of firms operated by the BNDES; section 2.3 describes the misallocation model used in the analysis, by Oberfield (2013); section 2.4 explains the data used in the exercise; section 2.5 produces some motivational results; section 2.6 explains the identification strategy used to capture the causal effects; section 2.7 presents results and discussion; section 2.8 concludes. In the end of the dissertation is an appendix with tables and graphs.

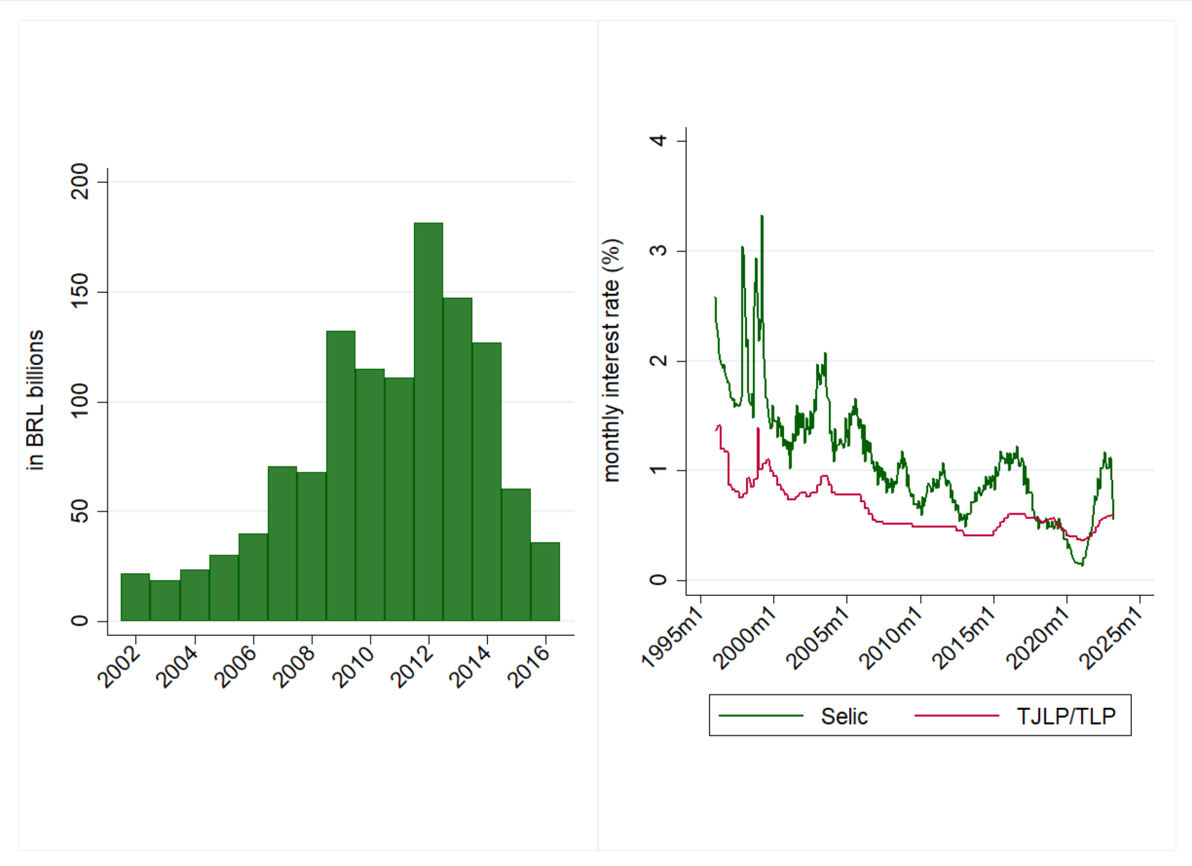
2.2 The Brazilian Development Bank and its Lending Conditions

The Brazilian Development Bank (BNDES) has been present in the discussions of targeted policy for a long time. During the 2000's, the BNDES gained increased importance in credit policy. Figure 2.1 shows the evolution of credit disbursements by BNDES from 2002-2016. It is clear that the government has opted for an increase in its portfolio during the 2000's, but that has been rolled back since 2015. A sample of their today products are described in the footnote.[2]. Many of their operations use the TJLP/TLP ("Long Run Rate"), which is more stable and lower than market interest rate. In Figure 2.1, one can see that the overnight Selic rate (which serves as base rate for monetary policy in Brazil) has a spread against the TJLP/TLP. Treasury operations on bonds follow closely the Selic rate. Risk lending to production activities are presumably done at even higher interest rates than government bonds. Lendings using TJLP/TLP could be seen as subsidies granted by the Brazilian government through the public banking system, mainly the BNDES.

Interest rates and the volume of credit granted by the BNDES through its size classification vary. Interest rates are mostly based on the TJLP/TLP, and are generally higher for smaller firms, followed by medium sized firms and then large firms, which most of the time are able to take credit at the lowest rate. The same is true for market interest rates, which tends to be much higher for smaller firms. Figure 2.2 plots the volume of credit granted for each category at the top left. At the top right, I zoom in for better visualization of small and medium. One can see that volume granted to the large category went through a very large shift. It was bellow *BRL 5 billions* in 2002, reaching almost *BRL 20 billions* in 2007, establishing a very large gap against the others. As for small and medium firms, the gap in volume is very narrow at the beginning, but begins to widen later in favor of medium sized firms. In terms of interest rates available, the years between 2004-2006 pushed higher interest rates to large firms. This reform is brought back to the traditional pattern in 2007.

[2] *Cartão BNDES*, which finances the acquisition of machine, equipment, inputs and services up to BRL 2 millions; *BNDES Automático* targets the acquisition of machine, equipment, construction, installations, training and the acquisition/development of national software, all focused on a more long run necessity than *Cartão BNDES*; *BNDES crédito pequenas empresas* targets the maintenance or generation of employment by micro and small firms; *BNDES Finame* targets the acquisition of machines, equipments, IT and automation products, buses and trucks; *BNDES microcrédito* targets micro firms, both formal and informal, in need of cash flows; *BNDES MPME Inovadora* targets the innovation on products, processes, and improvements in the ability to innovate; *BNDES Exim* targets the production of goods to be exported. *source: https://www.bndes.gov.br/wps/portal/site/home*

Figure 2.1 Credit Disbursements by BNDES from 2002-2016 (left), and Evolution of Key Interest Rates from 1996-2023 (right)



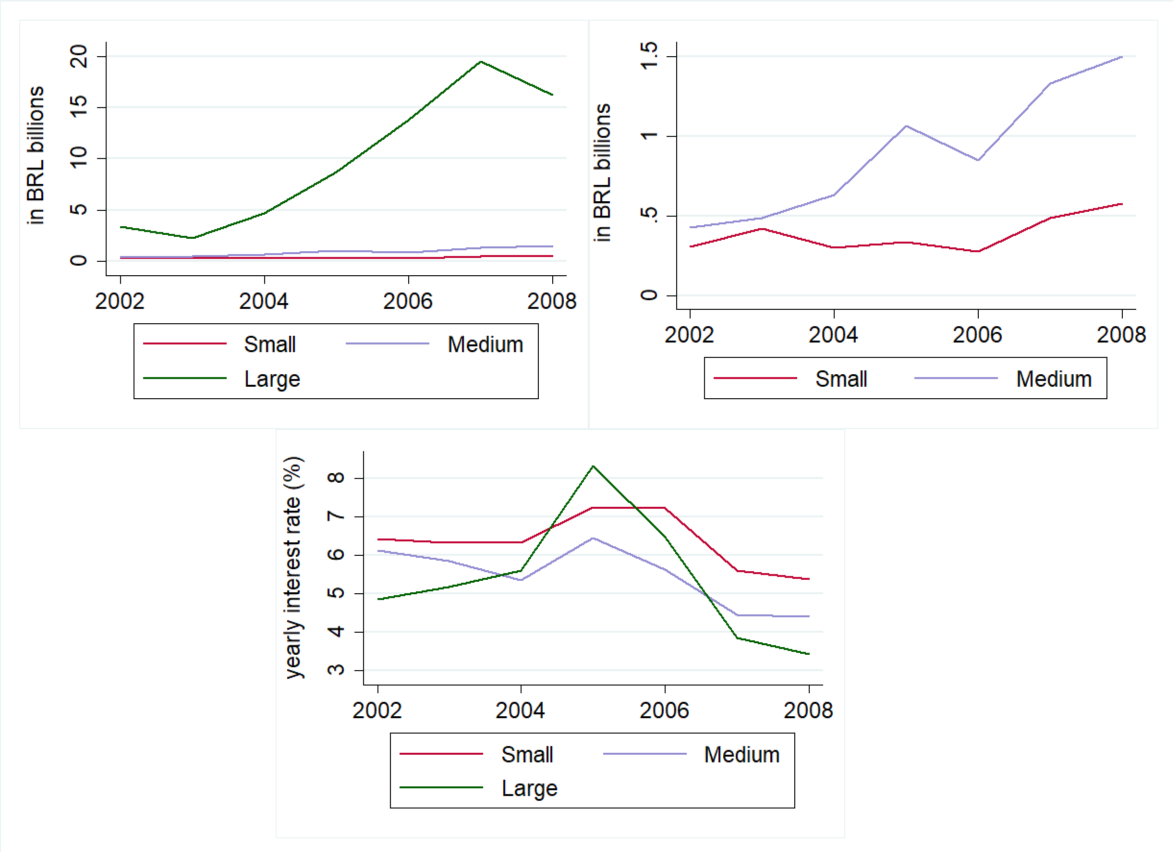
Source: The left graph is sourced from BNDES data for individual loan. The right graph is built with data from the Brazilian Central Bank.

2.3 Misallocation Framework

This section is concerned in explaining how I am going to measure the wedges. I choose to measure it based on Oberfield (2013). It differs from Hsieh and Klenow (2009) by assuming the capital intensity to differ from firm to firm, whereas in Hsieh and Klenow (2009) it is industry specific. In Hsieh and Klenow (2009)'s appendix they do a robustness check using different capital and labor intensities by firms, which ends up being the same model as Oberfield (2013).

I start by assuming a many industries environment. Within those industries, plants produce differentiated products that are combined to an industry aggregate. After that, industries

Figure 2.2 Credit Disbursements by BNDES by Size Category from 2002-2008 (top), and Average Lending Interest Rates by Size Category from 2002-2008 (bottom)



Source: BNDES data for all individual loan: available by size classification.

aggregates are combined into a single aggregate good. Let Y_i be the output of plant i, and $Y_s \equiv (\sum_{i \in I_s} Y_i^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$, the quantity for the industry aggregate for industry s, and $Y \equiv \prod_{s \in S} Y_s^{\theta s}$ is the quantity of manufacturing aggregate with $\sum s \in S\theta_s = 1$. Industries' shares evolve over time and price-taking consumers spend optimally. If P_i is the price of the good produced by plant i, than $P_s \equiv (\sum_{i \in I_s} P_i^{1-\sigma})^{\frac{1}{1-\sigma}}$, and $P \equiv \sum_{s \in S} (\frac{P_s}{\theta_s})^{\theta_s}$ are ideal price for the industry good and aggregate good, respectively.

Oberfield (2013) then constructs a frictionless economy, i.e., an economy where capital and labor are organized as to generate the optimum output.

$$\max_{[K_i, L_i]_{i \in I_s, s \in S}} \prod_{s \in S} [\sum_{i \in I_s} (A_i K_i^{\alpha_i} L_i^{1-\alpha_i})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}\theta_s}$$

subject to $\sum_{s\in S}\sum_{i\in I_s} K_i \leq K$ and $\sum_{s\in S}\sum_{i\in I_s} L_i \leq L$

The maximum attainable output is:

$$Y^{**} = \prod_{s\in S}\left(\sum_{i\in I_s}[A_i(\frac{\alpha_i}{\alpha^{**}}\theta_s K)^{\alpha_i}(\frac{1-\alpha_i}{1-\alpha^{**}}\theta_s L)^{1-\alpha_i}]^{\sigma-1}\right)^{\frac{1}{\sigma-1}\theta_s} \tag{2.1}$$

where α_i is defined to satisfy:

$$\alpha^{**} = \sum_{s\in S}\theta_s \sum_{i\in I_s}\frac{[A_i(\frac{\alpha_i}{\alpha^{**}}\theta_s K)^{\alpha_i}(\frac{1-\alpha_i}{1-\alpha^{**}}\theta_s L)^{1-\alpha_i}]^{\sigma-1}}{\sum_{j\in I_s}[A_j(\frac{\alpha_j}{\alpha^{**}}\theta_s K)^{\alpha_j}(\frac{1-\alpha_j}{1-\alpha^{**}}\theta_s L)^{1-\alpha_j}]^{\sigma-1}}\alpha_i \tag{2.2}$$

Oberfield (2013) notes some features of the frictionless economy. First, $\frac{\partial lnY^{**}}{\partial lnK} = \alpha^{**}$ and $\frac{\partial lnY^{**}}{\partial lnL} = 1-\alpha^{**}$, so that a first order approximation of the frictionless aggregate production function is a Cobb-Douglas production function with capital share α^{**}. This will be relevant when accounting for changes in output.

Second, the capital intensity of the frictionless economy, α^{**} is a weighted average of the capital intensities of individual plants, weighted by optimal size. As plant productivities change, the aggregate capital intensity shifts to better reflect the capital intensity of the lower-cost plants.

Third, α^{**} depends on the capital-labor ratio.

And lastly, with homogeneous capital intensities within industries, as in Hsieh and Klenow (2009), the frictionless factor share depends only on parameters: $\alpha^{**} = \sum_{s\in S}\theta_s\alpha_s$. In that case, α^{**} only changes because of the evolution of industry shares, θ_s.

The previous thought experiment involved imagining a frictionless economy across sectors. I now focus on another thought experiment, which consists in imagining the reallocation of capital and labor within industries.

let $Y^* \equiv \prod_{s\in S}(Y_s^*)^{\theta_s}$, where Y_s^* is:

$$\max_{[K_i,L_i]_{i\in I_s}} [\sum_{i\in I_s}(A_i K_i^{\alpha_i} L_i^{1-\alpha_i})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}$$

subject to $\sum_{i \in I_s} K_i \leq K_s$ and $\sum_{i \in I_s} L_i \leq L_s$. the maximum attainable industry aggregate is:

$$Y_s^* = (\sum_{i \in I_s} [A_i (\frac{\alpha_i}{\alpha_s^*} K_s)^{\alpha_i} (\frac{1-\alpha_i}{1-\alpha_s^*} L_s)^{1-\alpha_i}]^{\sigma-1})^{\frac{1}{\sigma-1}} \tag{2.3}$$

And α_s^* is defined as:

$$\alpha_s^* = \sum_{i \in I_s} \frac{[A_i (\frac{\alpha_i}{\alpha^*} K_s)^{\alpha_i} (\frac{1-\alpha_i}{1-\alpha_s^*} L_s)^{1-\alpha_i}]^{\sigma-1}}{\sum_{j \in I_s} [A_j (\frac{\alpha_j}{\alpha_s^*} K_s)^{\alpha_j} (\frac{1-\alpha_j}{1-\alpha_s^*} L_s)^{1-\alpha_j}]^{\sigma-1}} \alpha_i \tag{2.4}$$

Oberfield (2013) shows that a particular Solow residual can be decomposed into changes in "technology" and changes in the extent of misallocation.

Measuring allocational efficiency between and within industries:

$$M_W \equiv \frac{Y}{Y^*}$$

$$M_B \equiv \frac{Y^*}{Y^{**}}$$

Where M_W measures within-industry misallocation, and M_B measures the additional contribution to output of allocational efficiency between industries.

Changes in the efficient level of output can be decomposed into changes in aggregate capital, changes in aggregate labor, and a residual that reflects change in technology, $dlnA^{**}$.

$$dlnY^{**} = dlnA^{**} + \alpha^{**} dlnK + (1-\alpha^{**}) dlnL$$

changes in actual output can be decomposed as:

$$dlnY = dlnM_B + dlnM_W + dlnA^{**} + \alpha^{**} dlnK + (1-\alpha^{**} dlnL)$$

which can be rewritten as:

$$dlnY - \alpha^{**}dlnK - (1-\alpha^{**}dlnL) = dlnM_B + dlnM_W + dlnA^{**}$$

The equation above makes it possible to disentangle changes in allocational efficiency from changes in technology.

Now I need to focus on how the measures of allocational efficiency, M_W and M_B, and factor shares from efficient production, α^{**} and α_s^*, can be obtained from plant-level data.

Optimal spending by consumer requires $Y_i = Y_s(\frac{P_iY_i}{P_sY_s})^{\frac{\sigma}{\sigma-1}}$. Combining this with $Y_i = A_iK_i^{\alpha_i}L_i^{1-\alpha_i}$ gives:

$$A_i = Y_s\frac{\frac{P_iY_i}{P_sY_s}^{\frac{\sigma}{\sigma-1}}}{K_i^{\alpha_i}L_i^{1-\alpha_i}} \tag{2.5}$$

Plugging this in equations (2) and (4) and rearranging yields:

$$0 = \sum_{s\in S}\theta_s\sum_{i\in I_s}[(\frac{P_iY_i}{P_sY_s})^{\frac{\sigma}{\sigma-1}}(\frac{K/\alpha^{**}}{K_i/\alpha_i})^{\alpha_i}(\frac{L/(1-\alpha^{**})}{L_i/(1-\alpha_i)})^{1-\alpha_i}]^{\sigma-1}(\alpha_i-\alpha^{**}) \tag{2.6}$$

$$0 = \sum_{i\in I_s}[(\frac{P_iY_i}{P_sY_s})^{\frac{\sigma}{\sigma-1}}(\frac{K/\alpha_s^*}{K_i/\alpha_i})^{\alpha_i}(\frac{L/(1-\alpha_s^*)}{L_i/(1-\alpha_i)})^{1-\alpha_i}]^{\sigma-1}(\alpha_i-\alpha_s^*) \tag{2.7}$$

Similarly, plugging (5) into equations (1) and (3) along with $Y = \prod_{s\in S}Y_s^{\theta_s}$ yields:

$$\frac{Y^{**}}{Y} = \frac{1}{M_WM_B} = \prod_{s\in S}(\sum_{i\in I_s}[(\frac{P_iY_i}{P_sY_s})^{\frac{\sigma}{\sigma-1}}(\frac{\theta_sK/\alpha^{**}}{K_i/\alpha_i})^{\alpha_i}(\frac{\theta_sL/(1-\alpha^{**})}{L_i/(1-\alpha_i)})^{1-\alpha_i}]^{\sigma-1})^{\frac{\theta_s}{\sigma-1}} \tag{2.8}$$

$$\frac{Y^{*}}{Y} = \frac{1}{M_W} = \prod_{s\in S}(\sum_{i\in I_s}[(\frac{P_iY_i}{P_sY_s})^{\frac{\sigma}{\sigma-1}}(\frac{K_s/\alpha_s^{*}}{K_i/\alpha_i})^{\alpha_i}(\frac{L_s/(1-\alpha_s^{*})}{L_i/(1-\alpha_i)})^{1-\alpha_i}]^{\sigma-1})^{\frac{\theta_s}{\sigma-1}} \tag{2.9}$$

Note that no assumptions are made regarding plant's choices of capital and labor. The only assumption used are (i): the functional forms of each plants production function; (ii) the functional form of the demand system; and (iii) price-taking consumers make optimal purchasing decisions. The framework takes no stand on how prices are set.

Now I need to define the terms that characterize allocational efficiency. If resources are allocated efficiently within industries, the ratio of value added for plant i satisfies $\frac{\alpha_i P_i^* Y_i^*}{K_i^*} = \frac{\alpha_s^* P_s Y_s}{K_s}$. Define the capital wedge of plant i to be the deviation of this ratio from its efficient (within-industry) level: $T_{K_i} \equiv \frac{P_i Y_i / K_i}{P_i^* Y_i^* / k_i^*}$. Similarly, define the labor wedge to be $T_{L_i} \equiv \frac{P_i Y_i / L_i}{P_i^* Y_i^* / L_i^*}$. Lastly, we define the scale wedge for plant i to be $T_i \equiv T_{K_i}^{\alpha_i} T_{L_i}^{1-\alpha_i}$ A plant scale wedge is related to its within-industry allocational efficiency. A scale wedge larger than one means that the plant is small relative to its size in the efficient allocation.

Those wedges are actually what I am interested in estimating and using it as our main outcome.

2.4 Data

Data comes from the Brazilian Institute of Geography and Statistics (IBGE), specifically, the PIA (Annual Industrial Survey). It reveals firm-level economic information about the manufacturing sector, such as employees, wages and salaries, revenues, costs and expenses, investment, depreciation, output and intermediate consumption. It covers about 40,000 firms that has more than 30 employees.

There are no variables for the capital stock. It must be generated through the available data. The most popular method is the perpetual inventory model. In the exercise of section 5, I compute it through the traditional perpetual inventory method, where missing values on investment could implicate the existence of measurement error.

In the main exercise in Section 7, I use capital stock data available at IBGE's restricted room, where PIA is located. It was built as in Alves and Messa Silva (2008). An initial capital stock is constructed for 1996 using sectoral level data, at the lowest level available. The perpetual inventory method is then applied to capital formation statistics at the aggregate level, and then generated at the firm-level by assuming equal capital-to-labor ratios inside sectors. From then on, it is straightforward to compute the following years' capital stocks by taking account of the investment flows and the depreciation. Apart from that, the data also suffers from missing observations for the investment variables. Imputation is done by propensity score. The first years of the series could distort overall misallocation measures because of the way capital levels are computed for 1996, but this effect is already lowered by 1998. I choose to use this metric in my main estimation because the imputation helps with lowering additive measurement error. Our exercise starts mostly post 1998, and distortions from the methodology are unlikely to affect size cohorts differently.

Labor income is calculated as annual wages paid for blue and white collar professionals, corrected by the operation days of each firm. Wages were deflated by the Brazilian general price index.

The value added is calculated by the industrial gross value of output deducted by the industrial operational cost. Value added is also deflated by the Brazilian general price index.

I also reclassified firms, using the International Standard Industries Classification (ISIC) code system. The disaggregation occurred at the 4-digit level. I did not use sectors comprising of less than 4 firms, and excluded the top 1% and bottom 1% of firms as outliers. In addition, non-manufacturing firms were excluded from the sample.

2.5 Some Motivational Results

I first bring some results on both the evolution of misallocation around the period of the exercise and some indication that the BNDES might be able to target firms with a larger wedge.

Figure 2.3 reports the evolution of misallocation between 1996-2012. Vasconcelos (2017) has already executed this exercise. I replicate it here inside the context of increased directed

Table 2.1 Descriptive Statistics ($\sigma = 3$)

	New Small			Always Small		
Variable	mean	median	Obs	mean	median	Obs
K	5,560,071	2,463,269	12,939	1,997,300	658,801.86	50,030
WL	500,630.91	376,321.66	12,939	260,804	173,573.25	50,030
Revenue	9,097,503.9	9,084,428.5	12,939	2,733,624.5	2,401,756	50,030
Capital Wedge	0.097	0.187	12,939	0.021	0.126	50,030
Labor Wedge	- 0.378	- 0.262	12,939	- 0.419	- 0.279	50,030
Scale Wedge	0.020	0.137	12,939	- 0.081	0.056	50,030
	New Medium			Always Large		
Variable	mean	median	Obs	mean	median	Obs
K	27,022,904	10,866,583	7,002	126,335,786	27,210,427	5,898
WL	1,980,462.8	1,459,063	7,002	7,483,984	3,331,247.5	5,898
Revenue	45,131,330	44,947,968	7,002	279,489,864	108,821,200	5,898
Capital Wedge	0.103	0.178	7,002	0.122	0.145	5,898
Labor Wedge	- 0.440	- 0.3390	7,002	- 0.550	- 0.469	5,898
Scale Wedge	0.03	0.124	7,002	0.042	0.087	5,898

Notes: Descriptive statistics generated with the database presented in the data section (PIA-IBGE) plus estimations of wedges using the model described in section 2.3.

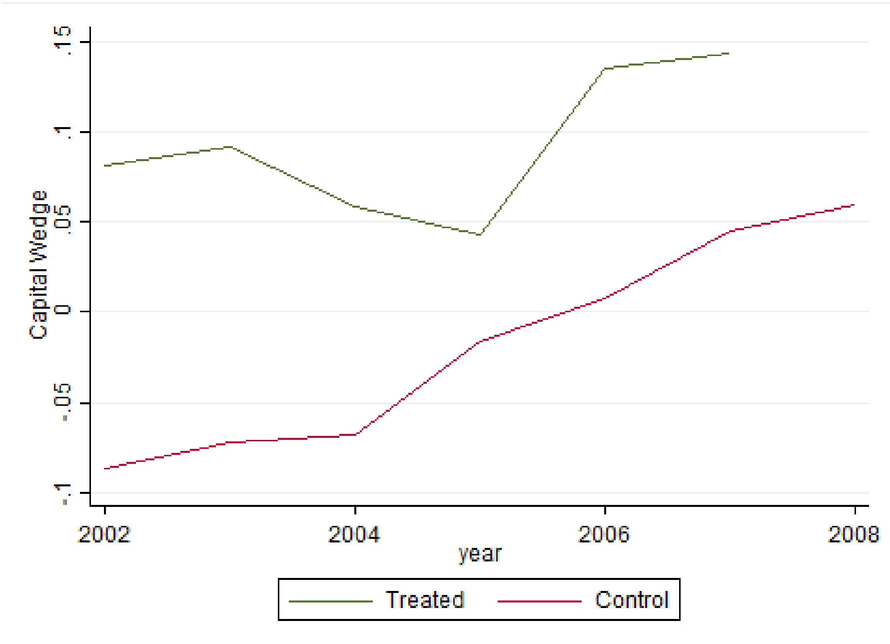
credit. Both within-sector misallocation and between-sector misallocation appear to coincide, on a casual correlational view, with the expansion of public credit. This serves as a motivation for the main exercise, where one could see the impact a change in credit conditions might have on the wedges used to create this measure.

Figure 2.4 plots the evolution of capital wedges from 2002-2008, for both recipients (treated) and non-recipients (control). There is no econometrics here. It is just an overview of treated and control wedges. Treated observations were constructed as firms' capital wedges one year before their first credit taken from BNDES. Control observations were firms that did not take any credit from BNDES during the period. It does appear that BNDES is able to target firms with larger capital wedges on average. What is potentially troubling is the possibility that the difference between treated and control groups are not too large, so that once firms receive the subsidized credit they expand beyond what would be optimal. Another possibility is that private banking is also capable of finding the better investments. Maybe even better than BNDES.

Figure 2.3 Evolution of misallocation between 1996-2012

Source: Graph generated by the model presented in section 2.3. Larger values imply better allocative efficiency.

Figure 2.4 Capital Wedges for BNDES recipients vs. non-recipients, pre concession



Source: Treated observations are the value of the average log capital wedge for 1 year before the firm first received a loan from BNDES. Control observations are firms that never received a loan in this period.

2.6 Identification Strategy

I build my identification strategy as in Cavalcanti and Vaz (2017). As already explained in the introduction, the BNDES operated a reclassification of firms in 2003. Some medium firms were suddenly classified as small, and some large firms were suddenly classified as medium, and then they became eligible to access different credit conditions. The natural experiment is built on this reclassification, as it creates the opportunity to assess the change in credit conditions against a counterfactual.

Table 2.2 BNDES Firm Classification by Revenue Through Time (in BRL)

Year	Small		Medium		Large	
	Min	Max	Min	Max	Min	Max
2000	700,000	6,125,000	6,125,000	35,000,000	35,000,000	-
2001	700,000	6,125,000	6,125,000	35,000,000	35,000,000	-
2002	900,000	7,875,000	7,875,000	45,000,000	45,000,000	-
2003	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-
2004	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-
2005	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-
2006	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-
2007	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-
2008	1,200,000	10,500,000	10,500,000	60,000,000	60,000,000	-

Source: BNDES. This table shows the classification of firms between small, medium and large, from 2000 to 2008.

The quality of the natural experiment lies in the assumption that the reclassification was exogenous to firms, so they could not anticipate it and position themselves in a desired spot. Checking for common trends ensure the quality of the identification. A possible concern would be the contamination of the control group by the treatment. If treated firms respond by further taking more or less credit from private credit markets, or if the possible expansion of public credit is done at the expense of private credit, the control group stops being a good counterfactual. I defend my identification on the basis that I am talking about a small fraction of the economy, i.e. the manufacturing sector, thus the BNDES policy targeting manufacturing firms might not severely change the economy's overall credit condition. In addition, treated groups are small in size in comparison to the rest of manufacturing. Therefore, it is unlikely that the control group becomes seriously contaminated by the treatment.

I build two new categories given the reclassification: *new small* and *new medium*. The *new small* group lies in between $BRL\,7,875,000$ and $BRL\,10,500,000$ in revenue. As one can see in Table 2.2, every firm falling in this bracket from 2000 onwards was classified as medium until 2002, and then they became small in 2003. The same happens with *new medium*, where firms in between $BRL\,35,000,000$ and $BRL\,60,000,000$ in revenue shift from large to medium. I construct the control groups as *always small*, with revenues from $BRL1,200,000$ and $BRL\,6,125,000$ and *always large*, with revenues upwards of $BRL\,60,000,000$.

I employ the following Differences-in-Differences model, then:

$$Y_{it} = \beta_1 New(category)_{it} + \beta_2 Post_t + \beta_3 Post_t \times New(category)_{it} + X_{it}\gamma + \varepsilon_{it} \tag{2.10}$$

where Y_{it} are the capital and scale wedges, or respectively, T_{K_i} and T_i. X_{it} refers to the possible controls added to the model.

What interests me is the parameter β_3, which captures the conditional expected value of wedges before and after the shift.

$$\beta_3 = E[Y_{it}|new\ category = 1, Post = 1] - E[Y_{it}|new\ category = 1, Post = 0]$$

$$-E[Y_{it}|new\ category = 0, Post = 1] - E[Y_{it}|new\ category = 0, Post = 0]$$

For the case of the new small firms versus always small firms, I shift the treatment to 2002, instead of 2003. This anticipation is done because there is a small change in credit conditions in 2002 favoring medium firms, so new small firms started to have better conditions in 2002, and then got the treatment via the reclassification. Common trend would not be violated, though. Effects mount mostly after the reclassification, so one could take 2003 as treatment.

The two exercises in this paper are the comparison of new small firms against always small firms and new medium firms against always large firms after a reclassification took place in

2003. I do not have access to complete data on the periods prior to 2002, but strong common trends are established in the exercise, suggesting the pattern of changing volume and interest rates present post 2004 is not present in the pre-treatment years.

The first exercise interpretation is straightforward: if new small firms were in a trajectory with common trends with the always small firms, their trajectory post-treatment reflect the change in credit conditions. A first inspection of the data presented would suggest that the credit shock is negative, since interest rates are higher for the small firms and volume is lower, in comparison with medium firms. The correct interpretation would have to adjust credit volume as a proportion of the size of the capital stock in each category, though. Cavalcanti and Vaz (2017) conducted a differences-in-differences in the same new small versus always small category, but with the investment rate in mind (investment/capital stock). It grew by 33.9% after the reclassification on a full period differences-in-differences. Although there was no leads and lags analysis, a graph showing the unconditional evolution of the investment rate reaches around 75% larger investment rate in the last years in comparison with the pre-treatment investment rate, suggesting a large positive adjustment in the capital stock over the years. That would suggest falling capital wedges in this exercise. This finding would be just a mechanical one, but I am interested in the magnitude of wedge changes relative to the distribution, and also in the position of wedges relative to the efficiency line.

The second exercise is done by comparing new medium against always large firms. Since they were both large pre-treatment, they necessarily faced the same conditions up until treatment, even if it changed through time. But it then diverged for both groups since they were classified differently. The treated is compared against a group that suffers major shifts in their policy mix post-treatment, the large group. It still serves as an exercise on the impact of being medium against a counterfactual of a changing large.

Capturing common trends before treatment is important to establish that groups moved in tandem when facing similar trends in their credit conditions, suggesting post-treatment results are due to different credit conditions.

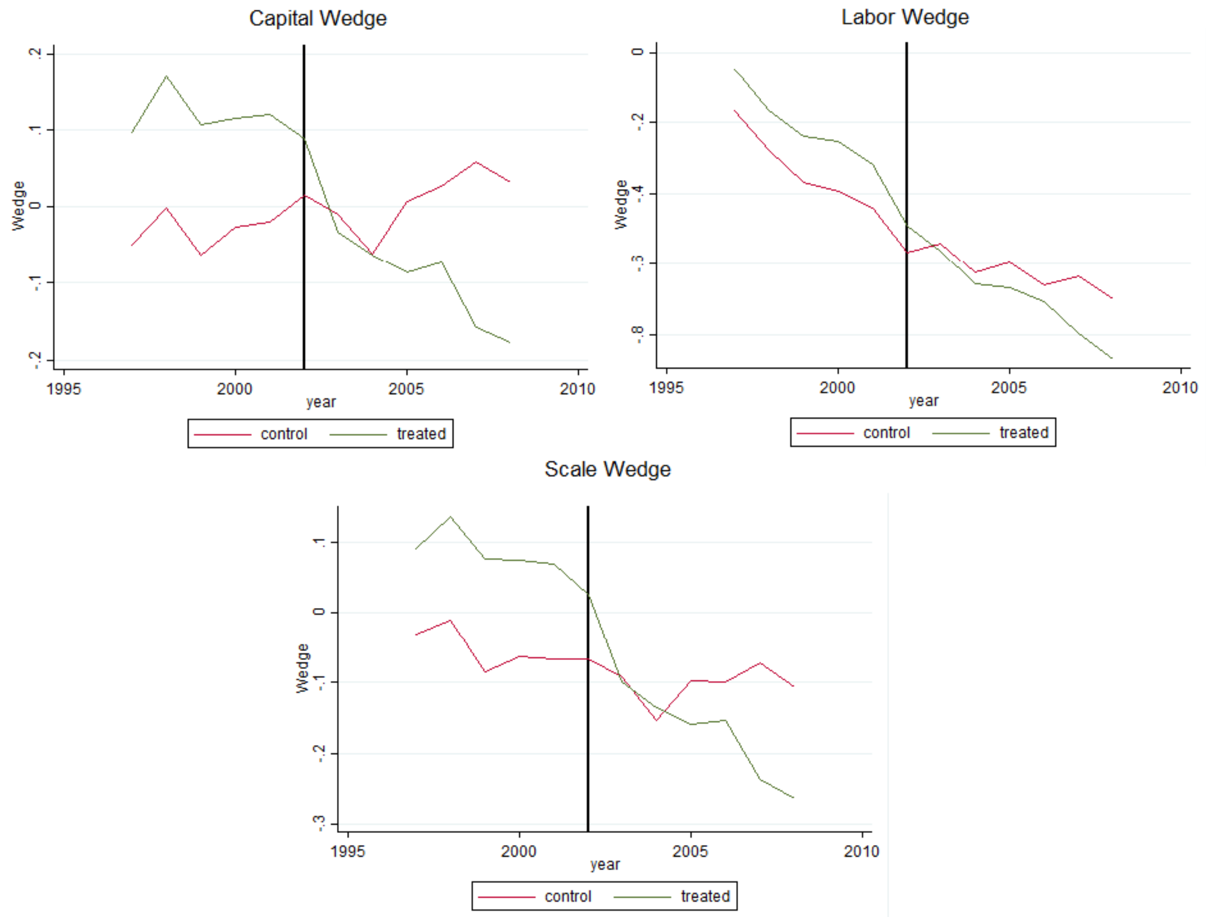
2.7 Results

I start by reporting the unconditional average for wedges during the study period in Figure 2.5. Results do not vary wildly by using different elasticities, so I report $\sigma = 3$ in the main text, following Hsieh and Klenow (2009), while reporting results for $\sigma = 1.5$ and $\sigma = 5$ in the appendix tables 4.1 to 4.6, where one could also find the tables for these two sigmas that were produced for $\sigma = 3$ in the main text. It is already visible that the credit shock did affect the wedge dynamics. The response is more visible in capital and scale wedges, and not so much in labor wedges. The stronger response in capital wedges are as expected. An expansion of the credit available for investment should mainly move the capital wedge, as firms take the opportunity to expand their capital stock and equalize marginal revenue with the new lower marginal cost. The less important dynamics on labor wedges are also expected, since the channel by which firms could use funds to adjust to their optimal labor hiring would be if they had serious cash flow constraints. That appears to be a smaller consideration. The scale wedge mimicked the evolution of the capital wedge.

Taking a first look at the wedge levels visible in the graphs, it appears that the shift in credit conditions sent both new small and new medium firms across the efficiency line at 0. In the case of new small firms, they start at a higher capital wedge level than both the counterfactual and the efficiency line, suggesting potential improvements with more access to capital. But the improvement in capital conditions actually sent new small firms below the efficiency line, from an under-investment position to an over-investment one, suggesting the improvement was actually more than needed from the economy perspective, if one takes the efficiency line as the correct metric for the evaluation of the misallocation of funds. Not only that, equal credit access as always small firms did not put new small firms in the same trajectory as always small firms afterwards. New small firms crossed the counterfactual line (the always small wedges), potentially because they are now the largest members of the small category, with better market interest rates available, making their mix of public and private credit interest rate lower than the smaller members of the small group.

As for the new medium comparison with the always large, their levels were already close to the efficiency line at pre-treatment. The shift in credit conditions to the new medium puts them in a lower level than the counterfactual (always large), even if the always large group also experienced a decrease in capital efficiency due to the large volume received after 2004.

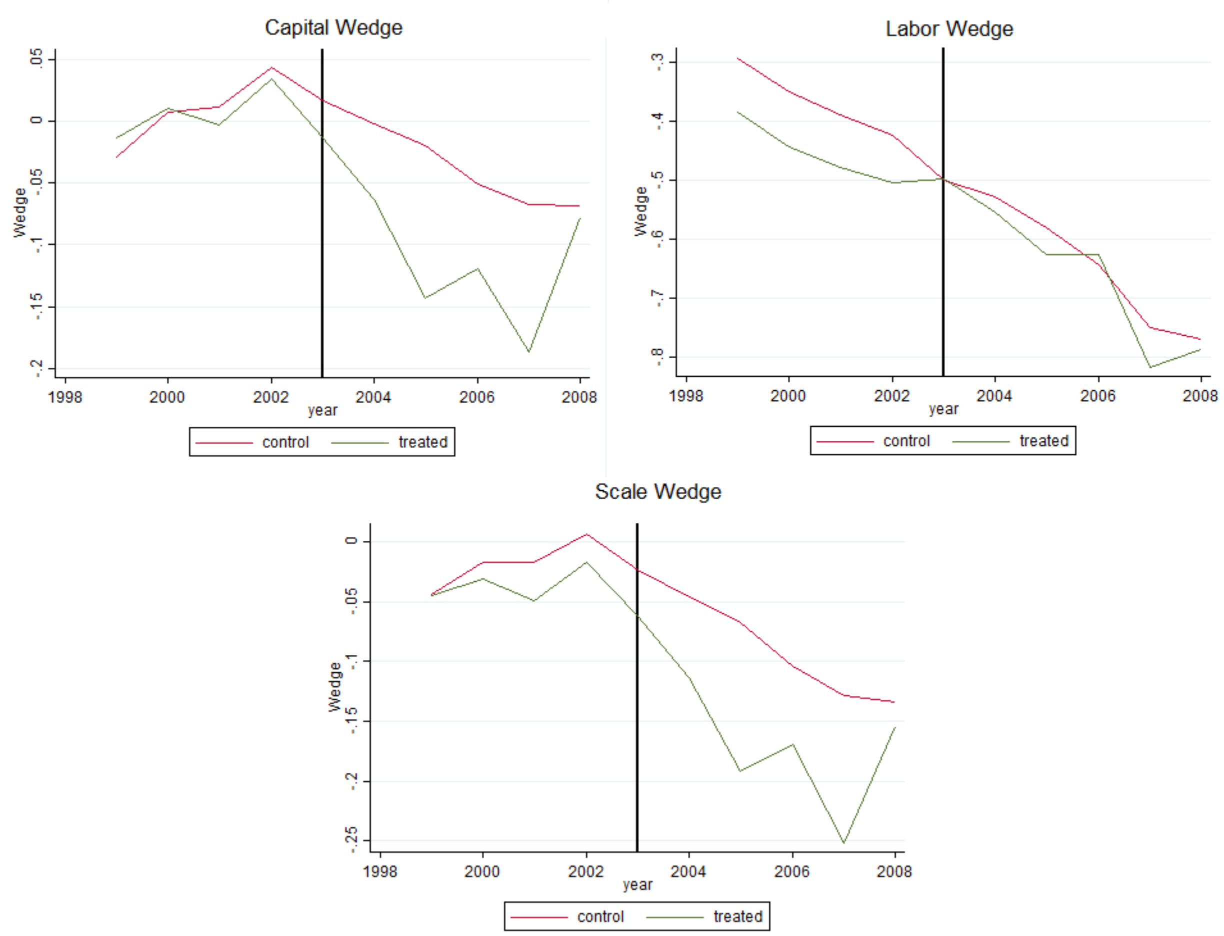
Figure 2.5 Unconditional Averages for Capital, Labor and Scale Wedges (New Small x Always Small)



Source: Graphs generated by plotting the unconditional means of the log capital wedge, log labor wedge and log scale wedge, for the size classifications of new small and always small

Table 2.3 reports differences-in-differences results for capital, labor and scale wedges, in

Figure 2.6 Unconditional Averages for Capital, Labor and Scale Wedges (New Medium x Always Large)



Source: Graphs generated by plotting the unconditional means of the log capital wedge, log labor wedge and log scale wedge, for the size classifications of new medium and always medium

both new small versus always small and new medium versus always large categories. As expected, results for the capital wedge are the most noticeable. Taking the specification with industry fixed effects and controls as the main result, a shift from medium to small generate a negative shift in average log capital wedges of - 0.157, significant at 99% confidence levels. Using results from Cavalcanti and Vaz (2017), the shift in the investment rate of 33.9% for new small firms during the treatment period caused the reduction of - 0.157 in average log capital wedges. I report the 2008 wedge distribution in the Appendix Figure 4.2. I do not hold the standard deviation to calculate the shift relative to the standard deviation. Although it does appear to be a somewhat small move around the distribution, it is enough to make new small firms to cross the efficiency line, locating them at the over-invested side of the distribution.

Results for the labor wedge were significant, but smaller in magnitude, suggesting that access to better credit conditions facilitated hiring. The shift from medium to small reduced the average log labor wedge in - 0.087, significant at 90% confidence levels. Labor wedges were already trending downwards on the over-invested side of the distribution. Increased access to cheap credit pushed it down even further.

As for the scale wedges, one must remember that they are given by $Ti = Tki^{\alpha_{si}} Tli^{1-\alpha_{si}}$. Since the shift in labor wedges were smaller than capital wedges, and manufacturing plants are quite capital-intensive in comparison to the typical 2/3rds labor share and 1/3rd capital share for the whole of the economy, the average log scale wedge effect is in between, of - 0.112, significant at 95% confidence levels. The scale wedge also crosses the efficiency line into over-scale territory.

Moving to the comparison between new medium versus always large, effects are much less pronounced. Only the average log capital wedge returns a significant result. It is of - 0.070 in our preferred estimation. Starting from around efficiency levels, the capital wedge falls faster for new medium firms than for always large firms. Perhaps more interestingly, even the shift in volume favoring the large group with close to 4x more capital available during the treatment period, versus 3x more capital for medium firms, was not enough to compensate against being in a lower category, with presumably more credit per capital stock available for lending. One must also bear in mind that interest rate grew for the large group during mid treatment, which could also have had some effect. If one could use the proportional shift in capital wedges of new medium relative to the shift from capital wedges of new small, of 0.446, one could use the 33.9% change in the investment rate for the new small in Cavalcanti and Vaz (2017), and return a 15.11% change in the investment rate of the new medium category.

I shift to the more interesting dynamics reflected in the event studies. Figure 7 plots event studies for new small versus always small. The shift in the investment rate in Cavalcanti and Vaz (2017) also happens during a period of adjustment, with the investment rate for new small firms adjusting to always small levels close to 2008. That is why the capital wedge keeps falling during the treatment period. The investment rate for new small firms increase around 75% by 2008 (Appendix Figure 4.1). The event study reported in Figure 2.7 is also reported in the second column of Table 2.4. After 6 years, the shock in estimated average log capital wedge is of - 0.261, significant at 95% confidence levels, which is noticeably stronger than the results for the overall differences-in-differences (- 0.157). One should take the estimates for the last period seriously, since it takes time for the investment rate and the capital stock to adjust to the new steady state. These estimates reflect better the endpoints for wedges after the shock in credit conditions. The same is true for labor wedges, which 6 years after the shock in credit conditions reach - 0.166, significant at 90% confidence levels. Results are not significant for scale wedges in any given year, despite being significant for the overall differences-in-differences. I still report here the effects for the last year, which is of - 0.195.

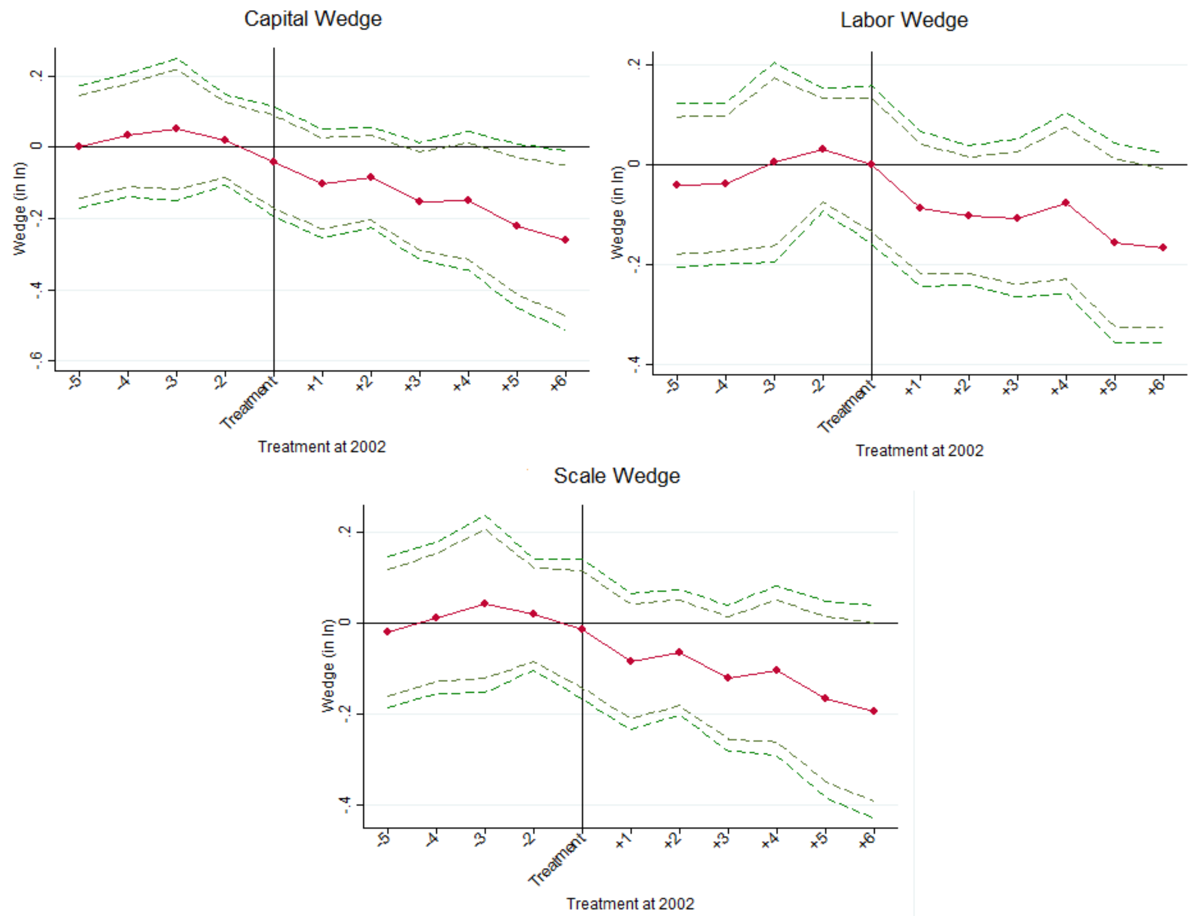
The new medium dynamics are different, and are reported in Figure 2.8 and Table 2.5 (the Figure 2.6 dynamics is shown in column 2). It actually reaches its largest results after 2 years. Credit conditions for medium and large groups did not change dramatically in the first periods, so this effect reflects better the difference in credit conditions in favor of medium firms against a stable counterfactual. The shift in volume and interest rates for the large group starts to appear more after that, and treatment reflect improvements for new medium firms in relation to an improving condition in volume and a temporary increase in interest rates for the large group.

Table 2.3 Treatment Effects on Capital, Labor and Scale Wedges

	New Small x Always Small								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.174***	- 0.163***	- 0.157***	- 0.086*	- 0.083*	- 0.087*	- 0.124**	- 0.115**	- 0.112**
	(0.062)	(0.061)	(0.059)	(0.051)	(0.049)	(0.048)	(0.057)	(0.056)	(0.0548)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	45,464	45,460	45,460	45,464	45,460	45,460	45,464	45,460	45,460
R^2	0.430	0.443	0.447	0.523	0.532	0.533	0.428	0.440	0.444
	New Medium x Always Large								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.072*	- 0.075*	- 0.070*	0.059	0.055	0.057	- 0.061	- 0.064*	- 0.059
	(0.040)	(0.040)	(0.040)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.039)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	17,133	17,128	17,128	17,133	17,128	17,128	17,133	17,128	17,128
R^2	0.511	0.532	0.533	0.589	0.605	0.605	0.508	0.529	0.31

Notes: Treatment effects for the traditional diff-in-diff estimation are given by Eligible*Post. Unit FE and Year FE are unit and year fixed effects. Industry FE are the fixed effects related to the ISIC classification. Controls capture firm characteristics like its location and its workers characteristics. I also report the number of observations and R^2.
*p*0.10 p**0.05 p***0.01*

Figure 2.7 Event Study for New Small vs. Always Small

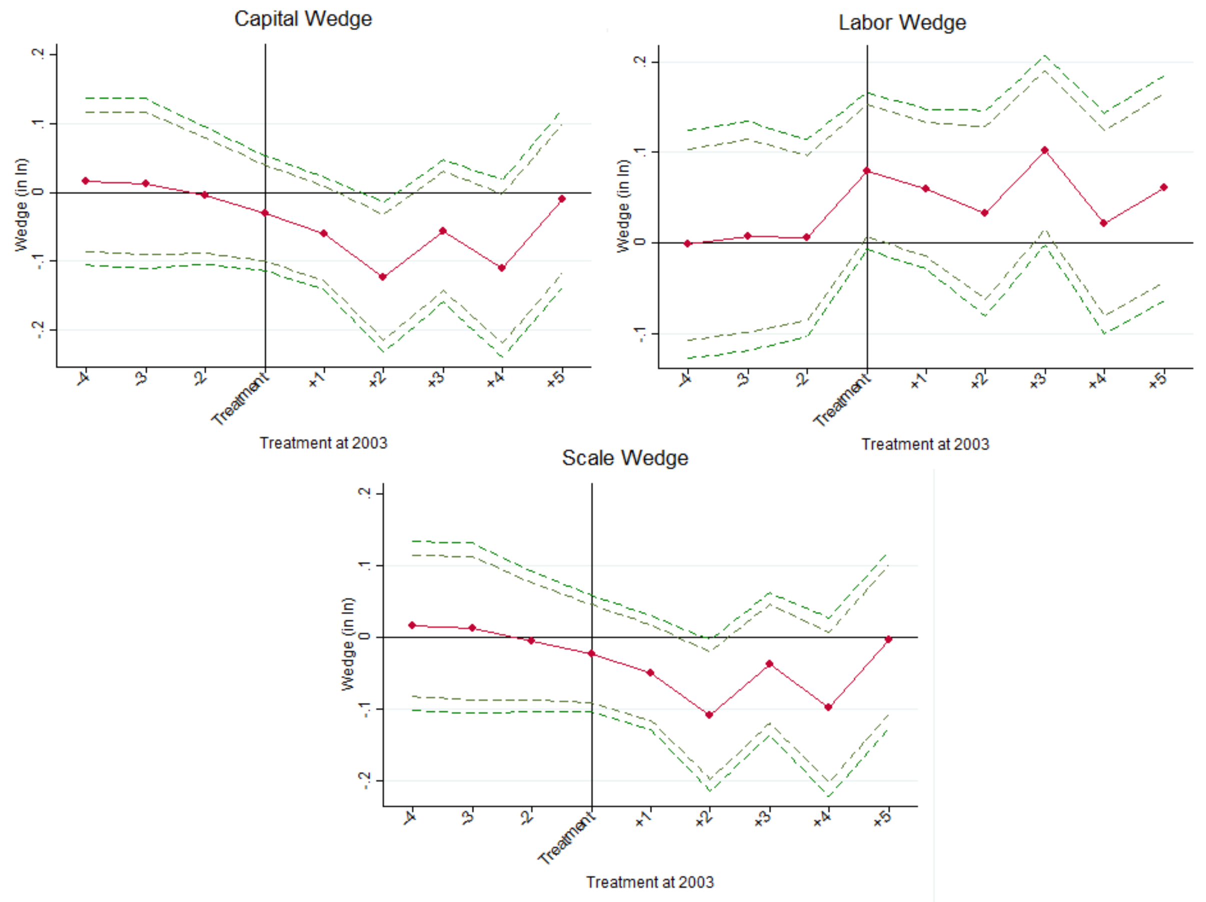


Source: Graphs generated by plotting the results of Table 4, column 2, for the treatment effects on the log capital wedge, log labor wedge and log scale wedge, for the new small vs. always small exercise. Confidence intervals of 90% and 95% are also reported.

The impact in average log capital wedges were of - 0.123, significant at 95% confidence levels, a much lower impact than the top impact of moving from medium to small. Using Cavalcanti and Vaz (2017) results for the investment rate, we estimate that the investment rate peaked at 26.55% increase for the new medium category after two years of treatment.

The average log labor wedge has an unexpected trajectory, which is felt mostly in the treatment year and 3 years after treatment, with 0.102, significant at 90% confidence levels. Results were not significant for the overall differences-in-differences, though. It does not appear that new medium firms had any cash flow problems affecting hiring.

Figure 2.8 Event Study for New Medium vs. Always Large



Source: Graphs generated by plotting the results of Table 5, column 2, for the treatment effects on the log capital wedge, log labor wedge and log scale wedge, for the new medium vs. always large exercise. Confidence intervals of 90% and 95% are also reported.

As for scale wedge, it reaches maximum results 2 years after treatment. The average log scale wedge falls - 0.110, significant at 95% confidence levels.

Our results suggest that shifting to better credit conditions has an impact on the measured wedges. This result is expected and mechanical. The more interesting result comes from estimating the magnitude of the shift in wedges given a certain change in credit conditions. Shifting to the small category appears to bring the most notable shift. As showed above, in the most extreme case, when one waits for the capital stock to adjust the most, the average capital wedge shifted - 0.261. By looking at Appendix Figure 4.2, the shift in the average might look relatively small, but a shift of - 0.261 in the average distribution of a cohort could potentially

Table 2.4 Treatment Effects on Capital, Labor and Scale Wedges (Event Study)

	New Small x Always Small					
	Capital Wedge		Labor Wedge		Scale Wedge	
5 years before	- 0.018	0.001	- 0.047	- 0.042	- 0.039	- 0.0198
	(0.088)	(0.088)	(0.082)	(0.083)	(0.084)	(0.085)
4 years before	0.019	0.034	- 0.031	- 0.038	0.000	0.012
	(0.087)	(0.088)	(0.081)	(0.082)	(0.084)	(0.085)
3 years before	0.036	0.051	- 0.000	0.005	0.029	0.043
	(0.101)	(0.102)	(0.101)	(0.102)	(0.098)	(0.099)
2 years before	0.015	0.021	0.022	0.029	0.013	0.019
	(0.065)	(0.064)	(0.062)	(0.062)	(0.063)	(0.063)
1 year before	0	0	0	0		
Year of treat.	- 0.040	- 0.040	- 0.000	- 0.001	- 0.012	- 0.013
	(0.079)	(0.079)	(0.081)	(0.081)	(0.078)	(0.078)
1 year after	- 0.136*	- 0.101	- 0.095	- 0.089	- 0.112	- 0.084
	(0.079)	(0.078)	(0.080)	(0.079)	(0.077)	(0.076)
2 years after	- 0.114	- 0.085	- 0.105	- 0.102	- 0.087	- 0.065
	(0.073)	(0.072)	(0.073)	(0.071)	(0.071)	(0.070)
3 years after	- 0.180**	- 0.151*	- 0.101	- 0.107	- 0.140	- 0.120
	(0.089)	(0.084)	(0.087)	(0.080)	(0.087)	(0.082)
4 years after	- 0.190*	- 0.150	- 0.086	- 0.077	- 0.136	- 0.105
	(0.104)	(0.099)	(0.098)	(0.092)	(0.099)	(0.095)
5 years after	- 0.270**	- 0.221*	- 0.160	- 0.157	- 0.208*	- 0.167
	(0.123)	(0.117)	(0.107)	(0.101)	(0.115)	(0.110)
6 years after	- 0.258*	- 0.261**	- 0.147	- 0.166*	- 0.194	- 0.195
	(0.136)	(0.128)	(0.104)	(0.097)	(0.126)	(0.119)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	45,464	45,460	45,464	45,460	45,464	45,460
R^2	0.430	0.448	0.523	0.533	0.428	0.444

Notes: Treatment effects for the event study diff-in-diff estimation are given by each year relative to one year before. Unit FE and Year FE are unit and year fixed effects. Industry FE are the fixed effects related to the ISIC classification. Controls capture firm characteristics like its location and its workers characteristics. I also report the number of observations and R^2.
*p*0.10 p**0.05 p***0.01*

produce a fatter tail in the negative side of Appendix Figure 4.2. To better visualize it with the distribution of wedges by revenues, Figure 2.9 below reveals that most of the concentration of negative capital wedges are present in the small category.

Table 2.5 Treatment Effects on Capital, Labor and Scale Wedges (Event Study)

	New Medium x Always Large					
	Capital Wedge		Labor Wedge		Scale Wedge	
4 years before	0.020	0.015	- 0.007	- 0.002	0.019	0.015
	(0.060)	(0.062)	(0.063)	(0.064)	(0.059)	(0.060)
3 years before	0.0204	0.013	0.005	0.007	0.018	0.012
	(0.061)	(0.063)	(0.063)	(0.064)	(0.059)	(0.060)
2 years before	0.000	- 0.004	- 0.002	0.005	- 0.002	- 0.006
	(0.050)	(0.051)	(0.055)	(0.055)	(0.049)	(0.049)
1 year before	0	0	0	0	0	0
Year of treat.	- 0.021	- 0.030	0.083*	0.079*	- 0.015	- 0.023
	(0.041)	(0.042)	(0.044)	(0.044)	(0.040)	(0.041)
1 year after	- 0.052	- 0.059	0.053	0.059	- 0.044	- 0.049
	(0.040)	(0.042)	(0.043)	(0.045)	(0.039)	(0.040)
2 years after	- 0.110**	- 0.123**	0.041	0.033	- 0.099*	- 0.110**
	(0.055)	(0.055)	(0.057)	(0.058)	(0.054)	(0.054)
3 years after	- 0.055	- 0.056	0.099*	0.102*	- 0.038	- 0.037
	(0.052)	(0.0525)	(0.052)	(0.053)	(0.050)	(0.050)
4 years after	- 0.113*	- 0.110*	0.017	0.021	- 0.103	- 0.098
	(0.067)	(0.066)	(0.062)	(0.062)	(0.065)	(0.064)
5 years after	- 0.024	- 0.010	0.045	0.060	- 0.020	- 0.004
	(0.065)	(0.066)	(0.061)	(0.063)	(0.063)	(0.063)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	17,133	17,128	17,133	17,128	17,133	17,128
R^2	0.511	0.534	0.589	0.605	0.508	0.531

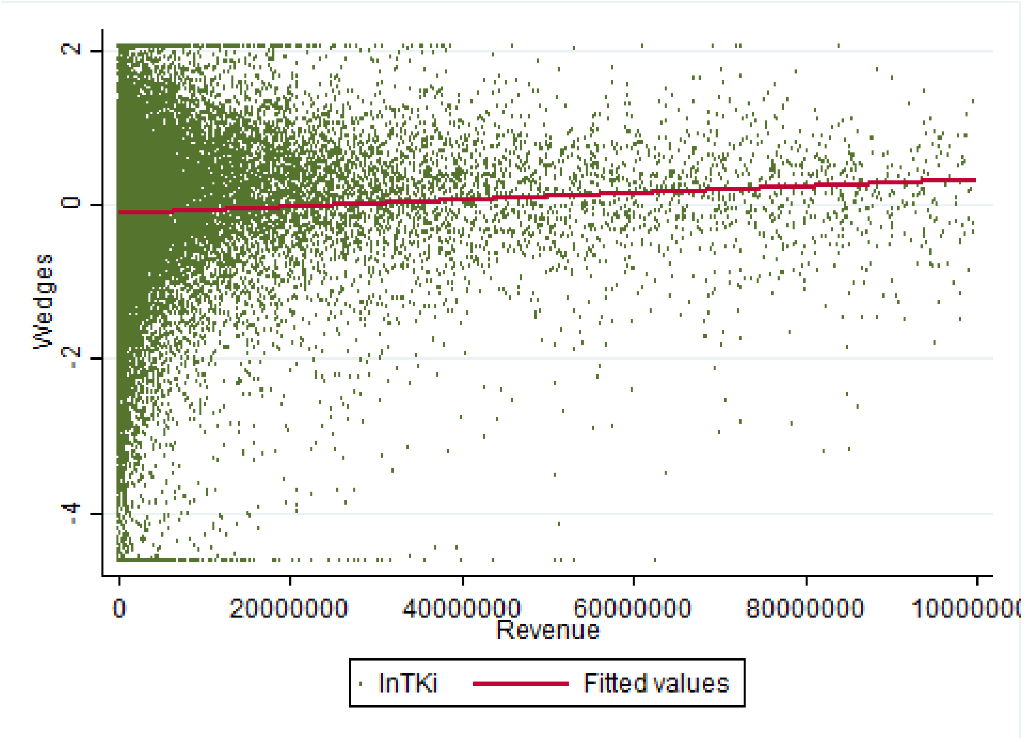
Notes: Treatment effects for the event study diff-in-diff estimation are given by each year relative to one year before. Unit FE and Year FE are unit and year fixed effects. Industry FE are the fixed effects related to the ISIC classification. Controls capture firm characteristics like its location and its workers characteristics. I also report the number of observations and R^2.
*p*0.10 p**0.05 p***0.01*

Since the estimate is the causal effect of access to better credit conditions from the BNDES, one can be sure that better credit condition was making small firms to operate at least 0.261 average log capital wedge below what they would operate if their conditions were similar to medium firms around treatment time. And medium firms operated at most 0.123 average log capital wedge below large firms during one treatment year. So I have established that the BNDES definitely affected wedges differently depending on the size cohort defined by the institution. The average capital wedge gap for the small group versus the medium group appears to have been 2.12 times larger than the gap between the medium group vs. the large group around treatment time. This is our main and most trustworthy result. It suggests that the BNDES should be sure if their targets, specially in the small group, are indeed credit constrained, otherwise it could give an important contribution to capital misallocation.

The second interesting exercise would be to determine if these shifts in wedges generated by the shift in credit conditions are actually improving or reducing allocational efficiency. If one takes the efficiency line at 0, this implies the assumption that all types of wedges affecting capital allocation act symmetrical on the distribution, so shifts in credit conditions improving or reducing optimal allocation could be approximated by the position of wedges relative to the overall efficiency line. Under this assumption, the new small group experienced a run into over-investment, after starting in under-investment position. Using estimates from the event study, I find that the reduction of 0.261 average log wedges 6 years after treatment against the counterfactual implies an average capital wedge of - 0.106, starting from 0.115 pre-treatment, which is basically a symmetric shift that would not have changed overall misallocation. In the case of the new medium group it shifted to - 0.138 because of changing credit conditions, coming from 0.03, which is an increase in misallocation due to the shock. If one could use estimates of the new small comparison against the always small in the reverse way, since the always small group is always close to the efficiency line, giving them the same condition of new small firms prior to treatment would send them to under-investment territory. Adding 0.261 to the first year pre-treatment average log capital wedge of - 0.01 would give us 0.260. That would suggest that the BNDES was actually giving the right amount of subsidies to the small group, and that handing them the same condition given to medium firms would increase misallocation.

The best way to use the indirect method coupled with an exogenous shock to capture misallocation is as in Bau and Matray (2023), a very recent paper pioneering this combination. Their shock is at the industry-level, through an event study. So any shrinkage of the distance between wedges above the median and the wedges below the median is enough to capture improved misallocation, which is what they report for their exercise on access to improved foreign credit. A next step would be to use similar intuition to this exercise to achieve better measures of allocational efficiency inside a given cohort.

Figure 2.9 Capital wedges by revenue in 2004



Source: Distribution of log capital wedges by revenue in 2004, showing the existence of a fat tail on the negative side of the smaller revenues

2.8 Conclusion

Summarizing my results, first and foremost, an improvement in credit conditions for some medium firms that were reclassified as small made their average log capital wedge and average log scale wedge shift significantly to the left of the distribution of their respective wedges. In terms of magnitude, it would be enough to generate a visible fat tail if I start by assuming that different cohorts starts with an equal distribution. That makes the BNDES credit policy of granting noticeable better conditions to the small firms a good candidate to explain part of the fat tail concentrated at negative wedges of small firms. A first approximation using the efficiency line as a guide to interpreting the misallocation of funds actually suggest that the BNDES credit policy is helping the achievement of allocational efficiency, not hindering it, although this interpretation relies on a stringent assumption of the efficiency line being a useful guide to interpret the misallocation of funds.

The improvement in credit conditions for the new medium firms relative to the always large firms is much smaller and temporary. The contribution of the BNDES policy to the misallocation of funds in the large and medium category could be related to the dynamics of wedges inside the cohort, and not across those cohorts, which is perhaps an even more interesting question, which also apply to the inside of the small cohort, but which I cannot answer without applying the recent developments by Bau and Matray (2023).

CHAPTER 3

Paving Feeder Roads to Remote Towns: Labor Markets and Firm Dynamics from Minas Gerais, Brazil

3.1 Introduction

Transportation infrastructure is probably understudied given the size of funds directed to such matters by governments and multilateral organizations such as development banks. The World Bank Group has funded $\sim$ 11.5% of its disbursements in transportation projects in 2022 (WB, 2022), through both IBRD and IDA, which is only smaller than health and public administration disbursements of all 11 sectors the bank categorizes. Smaller regional development banks, like the Inter-American Development Bank (IDB) (IDB, 2021), also pays a lot of attention to such matters. In 2021, the IDB directed $\sim$ 6.5% of its disbursements to transportation infrastructure. Feeder and rural roads are a relevant part of these funds. A recent interesting experience that has generated a lot of research comes from India. The government of India funded, with the help of the World Bank and the Asian Development Bank, a very large program to expand the network of roads connecting small villages to regional markets by an all-weather road, the PMGSY program, starting in early 2000s. This program cost USD 40 billion and was able to pave 400,000 kilometers of roads, benefiting 185,000 villages(Aggarwal, 2018; Asher and Novosad, 2020) .

Most rural and feeder road building programs use strong adjectives to describe its intended impacts. India's PMGSY suggested that "poor road connectivity is the biggest hurdle in faster rural development" (Asher and Novosad, 2020); "The best ways to promote rural development is to ensure good accessibility to growing and competitive urban markets" (WB, 2008); or, to cite the program this study aims: "(...) the adequate supply of transport infrastructure consists in a decisive stimulus to economic growth and the attraction of production investments".

Despite the apparent salience of the matter in the policymaker eyes, there is not the same amount of research done on the topic. One can understand why: the very nature of large infrastructure projects are of a small number of observations and a lack of randomization, which makes causal inference hard to obtain. An important exception is the aforementioned Indian PMGSY. This program is specially conducive to identification given its massive size and potential for good research design through diff-in-diff and fuzzy RDs (Adukia et al., 2020; Aggarwal, 2018; Asher and Novosad, 2020; Shamdasani, 2021). Apart from that major program, others were of much lower size or more relevant for urban environments, and not rural, backward areas of developing nations. This paper attempts to contribute to the literature on the impacts of transportation infrastructure, particularly the paving of rural, feeder roads, which provide access to a main, already paved road, and the regional markets nearby. I do that by exploring the roll-out of a program named *Pro-Acesso*, from the state government of Minas Gerais, Brazil, which provided the paving of feeder roads to 217 of the remaining small municipalities without an all-weather road from 1998-2014. Although not comparable to the PMGSY program of India in scale, it is relatively big for a feeder/rural roads program and quite amenable to a staggered differences-in-differences design (event study), provided we use the newest methods of the differences-in-differences literature that are able to correct for the problem of negative weights given to ATTs in the traditional two-way differences-in-differences regression Goodman-Bacon, 2021; Callaway and Sant'Anna, 2020.

Although Brazil is an upper middle-income country, and the state of Minas Gerais, specifically, is quite average and representative of Brazil in its development statistics, some of its rural areas are much poorer than the average, and connecting those municipalities to the more ample regional or Brazilian markets (or perhaps world markets) could have noticeable development implications. It is of notice that Brazil's backlands might be a different kind of feeder road target than India. The *Pro-Acesso* program offered an all-weather road to the remaining quarter of Minas Gerais municipalities without them. It is, then, a project to extend the benefit to all unconnected towns in the state, but starting at an already large share of paved connections within the state, which is not the same case of PMGSY in India[1], for instance. The average distance of unconnected towns to the nearest paved road in Minas Gerais in 1998 is 26 kilometers, and the median is of 22.4 kilometers, with the most autarkic town being exposed up to 122 kilometers of a dirt road. The level of autarky in Minas Gerais appears to be much smaller than that of the villages targeted by the PMGSY. We would like to know how the effects compare to the literature, and particularly to the Indian PMGSY, given its relevance in the recent literature.

This paper will focus on labor markets and firm dynamics outcomes, with attention to heterogeneity by sector, but specifically to the situation of the agricultural sector given its importance on tradables for the target municipalities. The effects of better connecting municipalities to foreign regions should be understood through the lens of trade theory, as is explicitly done in Asher and Novosad (2020). The treatment here is analogous to a reduction in trade tariffs, and that may shift production and consumption in treated municipalities. Recent research from the Indian PMGSY finds moderate to null results on their main outcomes, with more relevance to the reallocation of labor out of agriculture (Adukia et al., 2020; Aggarwal, 2018; Asher and Novosad, 2020; Shamdasani, 2021). Older research papers tend to be more optimistic about rural/feeder road infrastructure Hine et al. (2016). Since trade has distributional effects on top of welfare effects, it is interesting to pay attention to sectors or factors of production that might end up worse off, or at least pay some price through the transition period.

After running my regressions, I find that the agricultural sector suffers with higher resigna-

[1] As reports Aggarwal (2018), between the years 2001 and 2010, PMGSY provided roads to more than 110 million people, about 14.5 percent of the entire rural population, or 47% of the unconnected rural population as of 2001 in India. And as reports Asher and Novosad (2020), half of Indian villages were unconnected before the program, as opposed to 25% in the case of Minas Gerais state.

tions and lower wages in the sector. To test for the mechanism responsible for these effects, I show that agricultural output in the food crops sector falls, together with a larger contract termination initiated by the employer in the agricultural sector. The combination of these results suggest that the improved connection in treated municipalities actually caused a large demand hit, sending formal workers to unemployment, informality or migration, since the manufacturing and services sectors appear to not be highly affected. Just as in the international trade literature, there appear to be some noticeable transition effects in the medium run, which should be salient to policymakers in order to avoid overoptimistic expectations.

The paper develops as follow: section 3.2 builds some theoretical predictions arising from simple trade models; section 3.3 reviews the related literature, with more attention to the recent experience of India; section 3.4 gives the program background; section 3.5 explains the data used; section 3.6 explains the identification strategy; section 3.7 presents the results and discussion; section 3.8 provide some evidence on mechanisms; and finally, section 3.9 concludes.

3.2 Theoretical Predictions

Starting from a simple comparative advantage model, transport costs are analogous to trade tariffs. Once they fall, domestic and foreign markets interact by exploiting their comparative advantage.

Tradables from small local markets either gain foreign markets where they can sell at a higher price, possibly expanding production by absorbing factors of production from other industries, or they lose local markets from foreign competition, either shifting production to the new comparative advantage or lowering prices.

Trade in comparative advantage models increase welfare for both domestic and foreign markets. Still, there are interesting distributional shifts that might hurt different agents. Let's assume prices for agricultural goods are higher in foreign markets. That would increase agricultural prices for domestic producers and domestic consumers. The reverse would be true for lower foreign agricultural goods. These distributional shifts might be important from a policy perspective, since vulnerable parts of society might be hurt the most, and there might be adjustment costs present, although backward areas are characterized by low education, and therefore lesser technologically intensive production. Sector-specific skills might not be an important constraint in the reallocation of the labor force.

Trade also increase access to foreign inputs. It might be important for domestic farmers if they end up having access to better crops or fertilizers, for instance. Productivity might grow from this new access.

As for manufacturing, the channels are similar to agriculture. This natural experiment concern small towns with an agriculturally based economy. Manufacturing is present, but not large prior to the paved feeder road. Less technologically intensive firms might consider entering a town with paved access. Still, manufacturing tends to be more related to agglomeration economies (Krugman, 1991). Without a large demand nearby, a cluster of manufacturing firms would make sense if there is an incentive to a whole network of suppliers to establish around. That could be the case for some towns, but it does not appear realistic for most. In addition to that, I am only measuring the effects on the short to medium run. The establishment of a manufacturing base might take more time. Therefore, results pointing to industrialization are not expected.

Apart from the shifts in the domestic market structure of tradables, services might be affected by the arrival of a cheaper and larger supply of goods.

The paper by Casaburi et al. (2013) develops several theoretical predictions, assuming different market structures and intermediated trade, where city dwellers are responsible for transporting the goods from local markets to the city, while farmers transport them to local markets. Their theoretical models show multiple possible outcomes of a reduction in transport costs. Notwithstanding, their models assume from the start that local market prices (the town that received the paving in this exercise) are lower than in urban areas (the regional markets nearby in this exercise), which suggests that comparative advantage is set to work in favor of tradables in newly connected areas. When they run an RD on paved feeder roads in Sierra Leone, they find a lowering of prices in local markets. An interpretation from the simple comparative advantage model would be that prices in local markets were actually higher to begin with, and the paved feeder road made the distribution of foreign tradable goods feasible. To generate the same results with their assumption on lower domestic prices to begin with, the reduction in transport costs must manifest more in the cheaper transportation from farms to local markets than in the cheaper transportation from local markets to the city.

Although I take the results of Casaburi et al. (2013) as possible, it seems not realistic in this exercise, since feeder roads are linking a main road to a small town, which presumably is still linked to farms nearby through dirt roads. The shift in transport costs have to affect both exports and imports from and to the town mainly, and not from farms to the town. A reduction of domestic prices in this exercise would almost surely mean that foreign goods have become cheaper than domestic goods.

3.3 Related Literature

Most of the literature on the impact of rural and feeder roads rely on either underpowered statistics (less than 100 treated observations) or non-ideal designs, involving sub-optimal differences-in-differences[2], propensity score matching and case studies. Their findings show a range of results for various outcomes, including GDP per capita, sectoral employment, agricultural, manufacturing, services, schooling, consumption and health outcomes, mostly on the positive and optimistic side, although a few papers do find null results for their outcome of interest. I

[2]They either lack tests for common trends or they ignore the problem of negative weights (only recently formalized)

leave their main results to the footnotes, with peer-review papers. Other papers following the same spirit but only published as working papers or by development institutions are found in the literature review by Hine et al. (2016)[3].

The more immediate connection of this paper is with the literature on the Indian PMGSY program, which is suitable to identification on both sample size and research design. Four papers have studied its effects recently, and they serve as an interesting comparison with the *Pro-Acesso* experience. Three of the papers use a differences-in-differences strategy (Aggarwal, 2018; Shamdasani, 2021; Adukia et al., 2020), while Asher and Novosad (2020) explores a discontinuity on the probability of villages receiving treatment above or below a population threshold determined by the program plan to generate a fuzzy RD design.

Their findings suggest that the arrival of paved roads generated a large reallocation of labor out of agriculture. Asher and Novosad (2020) finds a 9 p.p. decrease in the share of workers in agriculture. Shamdasani (2021) finds a similar decrease, but only for villages that were close to towns, and no significant change in the composition of sectoral jobs for more distant villages. Shamdasani (2021) extended the analysis to movements of labor inside the agricultural sector between villages, showing that remote villages are able to hire more agricultural workers.

As for agricultural outcomes and technology adoption, Aggarwal (2018) finds an increase of 9% in usage of fertilizer and 7% in hybrid seeds for food crops, while no effect was found for cash crops. Asher and Novosad (2020) finds no effect for agricultural productivity using

[3] An increase in GDP per capita and working hours is found in Cuong (2011). Increased availability of goods and services is found in (Mu and Van de Walle, 2011; Cervero, 1990). Increased consumption is found in (Dercon et al., 2009). A Decrease in poverty rates is found in (Dercon et al., 2009; Fan et al., 2008; Owuor et al., 2007), although Fan et al. (2008) finds a small impact in comparison with other government spending. Positive impacts on agricultural outcomes, such as land value, fertilizer usage, can be found in (Jacoby, 2000; Mazlumolhosseini, 1990; Cervero, 1990; Dalton et al., 1997), while no significant impact on agricultural productivity is found in Fan et al. (2008). As for non-farm activities, road quality is important in Gibson and Olivia (2010). Evidence of sectoral reallocation to manufacturing and services can be found in (Mu and Van de Walle, 2011). Increased entry of small enterprises is found in Lokshin and Yemtsov (2005). Improved schooling metrics are found in (Mu and Van de Walle, 2011). Skepticism on the impact of improved health outcomes is found in Airey (1991), but optimistic findings are found in Al-Taiar et al. (2010); Buor (2003); Lokshin and Yemtsov (2005). Chomitz and Gray (1996) studies road building and deforestation, finding low economic impacts and high environmental impacts in Southern Belize, where soil is of poor quality. Porter (1995) finds negative results for women in the labor markets.

remote sensing data. They also do not find an increase in mechanized farming or irrigation equipment, and no effects on the composition of farming (subsistence or not) and land extensification. Shamdasani (2021) confirms increased access to fertilizers and seeds and also finds 15% increase in selling of agricultural products, mostly concentrated in small-scale landowners. When discriminating between cereal and non-cereal crops in the category that Asher and Novosad (2020) treated as subsistence farming, she shows that there is a 17 p.p. increase in the share of non-cereal crops, which she shows are mostly sold in regional markets. As for cash crops, it was confirmed to have a null effect.

There is no increased consumption in Asher and Novosad (2020), but Aggarwal (2018) finds a large reduction in import prices for some items imported from urban centers, while also finding an increase in product availability. Although Aggarwal (2018) had a more specific database on prices, only a small pool of items could be used to test for import price fall, so it might not be representative for all items. Another potential problem for this disparity in findings could be that for this specific test, Aggarwal (2018) could not test for common trends.

Another interest of those papers was in education and the employment of school-year individuals. Aggarwal (2018) found an increase of 5 p.p. in school enrollment for 5-14 year olds, while finding a reduction of 11 p.p. enrollment in 14-20 year olds. Adukia et al. (2020) focused on primary and middle school, finding an increase in 7% to 10% enrollment in middle school, but no increase in enrollment in primary school.

Another typical strand of the literature uses large highway programs as a source of exogenous shock to transport costs in rural counties and municipalities on the path of the highway. Michaels (2008) uses the American Interstate Highway System, whose construction spanned the 1950-70s. Treated rural counties showed an increase in trucking activity and retail sales at about 7-10 p.p. in per capita terms. He also found an increase in the wage of high-skilled workers relative to low-skilled workers where skill was abundant, a feature of Heckscher-Ohlin models of trade. Morten and Oliveira (forthcoming) use the same strategy for the highways built after the construction of Brasília. They find that welfare increased by 15.9%, the majority of which (88% of the increase) stemming from increased trade flows and the reduction of prices, while another 12% stems from increased migration. In the same spirit, Faber (2014) explores the expansion of China's highway system between the 1990s and the 2000s. He finds a reduction in GDP per capita in peripheral counties, mostly led by a reduction in industrial output growth.

One must bear in mind that the construction of highways linking major centers could generate different effects to rural areas in its vicinity than the expansion of feeder and rural roads into towns relatively far away from a major center. The immediate connection of this exercise is with nearby towns and regional centers, mostly through the connection of the paved feeder road to a secondary state route, not necessarily a major highway connecting metropolitan areas. Therefore, the papers investigating the PMGSY and the older literature on rural/feeder roads falls closely to this one.

3.4 Program Background

Minas Gerais is a large state in both land area and population. It is slightly larger than France on land area and the second largest state in Brazil by population, only below *São Paulo* state. It is also the state with the largest number of municipalities, most of which are very small in population.[4].

Although part of the relatively developed Southeast, the state of Minas Gerais is marked by regional inequalities. The Southern and Western sides of the state are more developed and more connected with the dynamics of prosperous agricultural and manufacturing hubs of the state of *São Paulo*. Its Northern and Eastern sides lag behind, appearing to follow the dynamics

[4]*Minas Gerais* has 586,528 km², while France has 551,695 km². As for population, *Minas Gerais* had 20.85 million inhabitants as of 2015. The state had 853 municipalities as of 2022, and the median population is only 8,323 in 2021 (IBGE)

of the Brazilian Northeast, or the neighboring state of Bahia, which is significantly poorer. The majority of paved feeder roads were executed in the poorest areas of the state, as it can be seen in Figure 3.1 [5], although there is a significant portion of them in the more dynamic areas too. As already mentioned in the introduction, the program paved a dirt feeder road for 217 municipalities, almost all of the remaining towns without a paved road in the state as of 1998. The DEER-MG, a body of Minas Gerais' state government responsible for the implementation of the *Pro-Acesso*, produced 2 documents on the program, with the following content: in the first document one finds the intentions of the program, as well as the original planning for the paving of roads, with data on the planned kilometers of length as well. On total there were 225 planned stretches of roads; and in the second document there is data on the actual amount of paved roads, 217 of the 225 planned, as of December 2014, the document's date. In addition, there is information on the kilometers paved in a given stretch of road, the end date of complete pavement, the end date of the contract, as well as the cost of each paved section.

The first document outlines the main objectives as providing paved roads for 225 municipalities which at that moment could not count with paved access to the main state highway system. They were expecting to allow for permanent traffic movement in and out of the municipalities [6], with smaller costs and more comfort and safety. In addition, they were expecting more access to job opportunities and basic services (education and health, mostly), while "contributing to the growth of economic activities, strengthening local capabilities, and facilitating the integration to outside markets and the attraction of new enterprises" (our translation). It is

[5]Brazil's IBGE has a sub-regional classification of states. FJP (2002) provides GDP per capita data for each macro-region inside *Minas Gerais* state. The richer ones are closer to the border of the states of *São Paulo* and *Goiás* and are to the South and West of the state. They are the following regions, with their 1999 GDP per capita in brackets: *Central* (R$ 6,408.49), *Sul de Minas* (R$ 4744.63), *Triângulo* (R$ 5,494.50), *Alto Paranaíba* (R$ 5,268.26), *Centro-Oeste de Minas* (R$ 4,392.68) and *Noroeste de Minas* (R$ 4,928.27). The *Rio Doce* (R$ 4,526.42) region has received many new pavings and is the richest region to the East of the state, mostly because of mining and heavy industry activities. The remaining regions border the states of *Bahia*, *Espírito Santo* and *Rio de Janeiro*, and are to the North and East of state, and the poorer ones: *Norte de Minas* (R$ 2,773.15), *Jequitinhonha/Mucuri* (R$ 1,735.73) and *Zona da Mata* (R$ 3,646.10). One can see that the GDP per capita is generally between half or a third of the more developed regions. And they are the ones that received most of the new paved roads provided by ProAcesso.

[6]Presumably, the dirt road must affect normal traffic during the rainy season, which is in the summer.

explicitly stated that better access to agricultural inputs was one of the aims of the program.

The document stated that those are the last 225 municipalities without access to the main highways (26% of the state municipalities). The total amount of roads to pave was estimated to be around 5,600 kilometers. The document also states that there was a large growth in the amount of roads paved in the state between 1975 and 1985 (10% growth), while before the program the growth rate was closer to 2% in the previous decade. It is also stated that older efforts were concentrated in connecting the main cities and economic hubs of the state, while the remaining towns are the smaller and less dynamic ones.

The second document allows us to generate figure 3.2, which is the roll-out of *Pro-Acesso* between 2002 and 2014, where one can see that most road pavings were completed between 2005 and 2012.

3.5 Data

3.5.1 RAIS

This database comes from *Relação Anual de Informações Sociais - RAIS* ("Social Information Annual List", on a direct translation). It is kept by the Brazilian Labor Ministry and every formal employer is required by law to report information from workers in an annual basis. I have access to the database for the years 2002-2014. I exclude data from 4 municipalities that received a paved road before 2002, since I do not have access to the data.

Among the information available, one could track the job of each worker, as he/she is identifiable. This is also true for firms' plants. The information has specific geographical quality. It can be accessed on the municipality level. They are also readily accessed by larger region classifications inside the state. As for the plants, one can have detailed information about

Figure 3.1 Minas Gerais Roadmap with ProAcesso's New Pavings

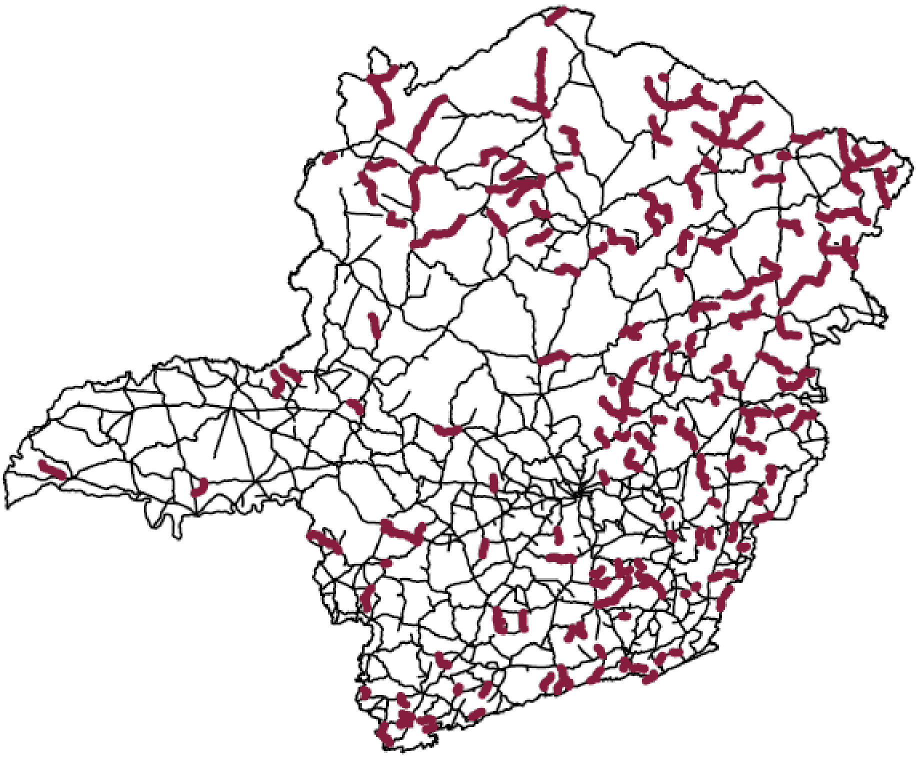


Figure 3.1 ***Notes:*** The above map shows the already paved roads before the ProAcesso program in black, with the new pavings from the program stressed in "cranberry" (color). It was created using data about the program provided by the DEER-MG

their industry classification in the Brazilian CNAE, which is relatable to the ISIC [7]. One can calculate the total number of workers at a specific plant on a monthly basis, since the database has detailed information on admissions and resignations. There is also information on the cause of the resignation - if the contract was terminated by the employer or the employee, for instance - which will be useful in my analysis. It is also possible to identify their profession through data on occupation. There are several variables available for wage. One could choose from using the december wage or the average monthly wage during the year. They are in current nominal R$ (BRL - Brazilian *Reais*). There is also the possibility of using wage information expressed as a

[7]CNAE stands for *Classificação Nacional de Atividades Econômicas*, or "National Classification of Economic Activities" in a direct translation, and is cataloged by the IBGE. ISIC stands for International Standard Industrial Classification, and it is provided by the United Nations Statistics Division. It helps to guide a standardized national classification.

Figure 3.2 The Program Rollout Throughout the Years (Date of Road Completion)

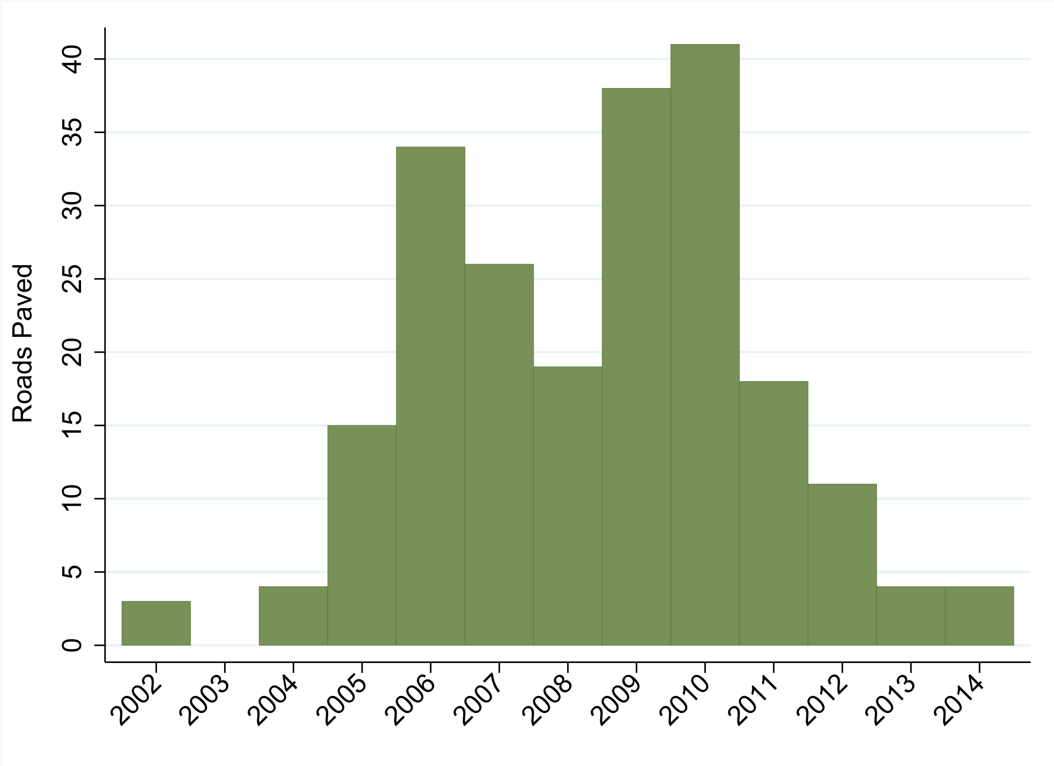


Figure 3.2 ***Notes:*** The above bar graph was created using data on ProAcesso provided by the DEER-MG

minimum wage multiple. Working hours are also available in the database. Demographic data on the employee is available on a number of dimensions. It provides information on education, gender, race, nationality and age.

I use this database to build the outcome variables related to labor market and firm dynamics. My first outcome of interest is admissions. This data is available on a monthly basis, but I consider it by year. The same is true for resignations (contract termination), my second outcome of interest.

The third outcome is wages. I choose to work with the average monthly wage in the year, since some workers are not present during the whole year. Values are at current prices, so I adjust it to 2022 R$ [8].

My fourth outcome is firm size. This one is not readily available, so it needs to be created

[8]BRL - Brazilian *Reais*

from other variables. I sum all workers working in a given year at a given plant. Since some workers are not present for the whole of the year, instead of treating them as a unit, it is possible for them to assume a fraction of a unit, depending on the amount of months they spent at that particular firm. With these steps I achieve a variable for firm size, identified at the plant level.

My fifth and sixth outcome of interest are the entry and exit of firms. These are gathered each time a firm appears or disappears in the database. There are strange appearances and disappearances happening during the years. I choose to take each one of them as an entry and exit. They will cancel out in the comparison between treated and control groups in the main exercise, so there is no need to further investigate why firms appear and reappear in the database.

To test for mechanisms, I use available information on the cause of contract termination, which could be initiated by employers or employees.

I make use of the following variables for controls in many of the regressions to be employed: work experience (at the current firm), working hours, education, gender, race and age.

3.5.2 PAM

The second database is *Produção Agrícola Municipal - PAM*, or "Municipality Agricultural Production", in a direct translation, which is a database collected by the *Instituto Brasileiro de Geografia e Estatística - IBGE* at an annual basis. It supplies the data on agricultural outcomes by municipality. I use the following outcomes on regressions: planted area (in hectares) and yields (in tons).

Data are available on a crop basis, so I choose to separate them in cash crops and food crops, as in Aggarwal (2018). The former are commodities that are important in the Brazilian export flows, while the latter are mostly sold inside Brazil, presumably in regional markets.[9]

[9] cash crops were defined as: coffee, soybeans, corn, sorghum, sugar cane, orange and wheat. food crops were the remaining crops: avocado, pineapple, cotton, garlic, peanuts, rice, oat, olive, banana, sweet potato, yellow potato, latex, cocoa, khaki, cashew nut, onion, barley, tea plant, coconut, palm oil, mate herb, pea, broad bean,

3.5.3 Descriptive Statistics

I provide the summary statistics for the 213 municipalities used for the study. Since the identification strategy uses the not yet treated municipalities as controls, most municipalities will act as treatment and controls, therefore I present the descriptive statistics for all municipalities in a single group. For data at the individual level, I take the mean of all characteristics inside the municipality first, and then compute the mean, standard deviation and median for all characteristics across municipalities.

Table 3.1 Descriptive Statistics

Municipality Characteristics			
Variable	Mean	Std. Deviation	Median
GDP per capita	3,001.264	1,962.8816	2,538.113
Population	6,103.426	4,282.125	4,997.5
Kilometers Paved	26	16.16	22.4
Workers Characteristics			
Variable	Mean	Std. Deviation	Median
Education	4.580	0.656	4.588
Gender	1.423	0.125	1.429
Age	35.562	2.345	35.485
Wage	390.692	112.592	367.449
Working Experience	53.885	25.492	51.784
Working Hours	38.57	4.642	39.434
Agricultural Characteristics			
Variable	Mean	Std. Deviation	Median
Planted Area (Cash Crops)	1,684.069	2,660.73	841
Planted Area (Food Crops)	643.497	780.575	380
Yield (Cash Crops)	10,909.36	30,473.32	4,763.5
Yield (Food Crops)	2,687.335	5,358.658	1163

Notes: This table provides the descriptive statistics for the data. They are all in the year 2002, the first year we have data, apart from the information on the kilometers of paved roads, which is related to the treatment. GDP per Capita and population were extracted from IBGE. All Workers Characteristics comes from RAIS. Education is on a scale from 1 (illiterate) to 11 (PhD); Gender is 1 for female and 2 for male; Wage is the average monthly wage in 2002 BRL (Brazilian Reais); Working Experience is computed for active months of work for current employer; Working Hours are the weekly hours of work; Planted Area are in hectares and; yields are in tons. All variables are described in their means, standard deviations and medians.

beans, fig, tobacco, guava, guarana, jute fiber, lemon, linen, apple, mauve, papaya, melon, nuts, palm heart, pear, passion fruit, black pepper, ramie fiber, sisal fiber, tangerine, tomato, triticale, tung tree, anatto, grape, manioc, mango, peach, quince and watermelon.

Table 3.1 provides the details. One can see by the descriptive statistics, specially the low standard deviation, that the distribution for most variables are not very different by municipality, mainly the ones related to workers characteristics. Still, it is clear from the standard deviations that there are substantial heterogeneities in GDP per capita and population, as well as for the variables related to working experience in the current firm and planted area and yields in the municipality. Heterogeneity in planted area and yields are probably related to the area size of each municipality, as well as GDP per capita. Big differences in the working experience for the current firm probably suggests that turnover can be very different for each municipality. I control for observables in one specification, as described below.

3.6 Identification Strategy

To estimate causal parameters, I explore the roll-out of the *Pro-Acesso* program to perform an event study. I take advantage of the recent literature on the matter. Several papers identified the problem arising from negative weights in naive event study fixed effects regressions (Athey and Imbens, 2022; Borusyak and Spiess, n.d.; Chaisemartin and D'Hautefoeille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2022).

The Two-Way fixed effects regression follows the known form, with leads and lags here:

$$Y_{it} = \gamma_i + \theta_t + \sum_{\tau=0}^{m} \delta_{-\tau} D_{s,t-\tau} + \sum_{\tau=1}^{q} \delta_{+\tau} D_{s,t+\tau} + \mathbf{X}_{\mathbf{it}}^{'}\beta + \varepsilon_{it} \tag{3.1}$$

Where γ_i is the municipality fixed-effects, θ_t is the year fixed effect, the various $D's$ are dummies for the leads and lags, with the $\delta's$ being the fixed-effect estimator associated with each lead and lag, and X' is the control vector, with β the coefficient associated with that vector.

I follow Chaisemartin and D'Hautefoeille (2020) in a brief explanation for the potential problems with the Two-Way estimator above, in the aggregate form. They show that under the following assumptions: (i) balanced panel of groups; (ii) sharp design; (iii) independent groups; (iv) strong exogeneity; and (v) common trends, the expectation of the average treatment effect (ATE) is given by:

$$\delta^{TR} = \mathbb{E}\left[\sum_{g,t:D_{g,t}=1} \frac{N_{g,t}}{N_1}\Delta_{g,t}\right]$$

where $\Delta_{g,t}$ is the average treatment effect on the treated (ATT) for each (g,t) cell, and $N_{g,t}/N_1$ is the share of treated observations in cell (g,t) over all treated observations. That acts like a weight given to that particular ATT. But applying the same assumptions for the fixed effects estimator in the version of equation (1) without the leads and lags, will generate the following expectation:

$$\beta_{fe} = \mathbb{E}\left[\sum_{(g,t):D_{g,t}=1} \frac{N_{g,t}}{N_1} w_{g,t}\Delta_{g,t}\right] \; where \; w_{g,t} = \frac{\varepsilon_{g,t}}{\sum_{(g,t):D_{g,t}=1} \frac{N_{g,t}}{N_1}\varepsilon_{g,t}}$$

Both estimators differ if $w_{g,t} \neq 1$. So heterogeneous effects across groups and time are given different weight than their proportion in the sample. Another potential problem comes from using already treated units as controls if treatment effects build over time. In that situation, those later effects are going to assume the position of counterfactuals in comparisons against treated units, canceling out legitimate treatment effects. In the worst case scenario, the signal can invert.

To deal with such problems, one needs to use event study estimators that are robust to heterogeneous and dynamic treatment effects. Chaisemartin and D'Hautefoeille (2020) provide their own estimator in their paper, which I use to report estimates in this paper. I now review their estimators.

Under the assumptions of: (i) sharp design; (ii) no anticipation; (iii) non-pathological design; (iv) independent groups and strong exogeneity; and (v) common trends [10], the average

[10](i) Sharp design implies that every unit inside a treatment group gets treated; (ii) no anticipation implies that reaction to treatment happens after treatment; (iii) non-pathological design requires that at least one group goes

treatment effect is given by:

$$\delta_+^{tru} = E\left[\frac{\sum_{g:2\leq F_{g,1}\leq NT}\sum_{t=F_{g,1}}^{NT} N_{g,t}(Y_{g,t}(\mathbf{D_g}) - Y_{g,t}(\mathbf{0}))}{\sum_{g:2\leq F_{g,1}\leq NT}\sum_{t=F_{g,1}}^{NT} N_{g,t}D_{g,t}}\right]$$

where δ_+^{tru} is a truncated at NT. It includes all treatment effects that can be estimated under the assumptions (i)-(v). $F_{g,1}$ denotes the first year where group g was treated, and NT is the last date where a group remains untreated. Essentially, I am interested in the comparison between the potential outcome of treated groups at some point in time where they were treated against their potential outcome if untreated. Chaisemartin and D'Hautefoeille (2020) develop the following estimator, which is an unbiased estimator of δ^{tru}:

$$DID_{+,t,l} = \sum_{g:F_{g,1}=t-l} \frac{N_{g,t}}{N_{t,l}^1}(Y_{g,t} - Y_{g,t-l-1}) - \sum_{g:F_{g,1}>t} \frac{N_{g,t}}{N_t^{nt}}(Y_{g,t} - Y_{g,t-l-1})$$

$DID_{+,t.l}$ is the DID estimator comparing the outcome evolution from period $t-l-1$ to t in groups treated for the first time in $t-l$ and in groups untreated from period 1 to t. Then, they define:

$$DID_{+,t,l}^D = \sum_{g:F_{g,1}=t-l} \frac{N_{g,t}}{N_{t,l}^1}(D_{g,t} - D_{g,t-l-1}) - \sum_{g:F_{g,1}>t} \frac{N_{g,t}}{N_t^{nt}}(D_{g,t} - D_{g,t-l-1})$$

This equation simplifies to $DID_{+,t,l}^D = \sum_{g:F_{g,1}=t-l} \frac{N_{g,t}}{N_{t,l}^1} D_{g,t}$

Let $L_{nt} = NT - min_{g:F_{g,1}\geq 2} F_{g,1}$ denote the number of time periods between the earliest date at which a group goes from untreated to treated and the last period at which a group has been untreated all along.

from untreated to treated at a date where some group has been untreated all along; (iv) independent groups and strong exogeneity imply that the treatment status of a group does not affect other group's potential outcomes, while strong exogeneity requires that a group treatment is not related to its outcome evolution, avoiding that a group gets treated because it suffered a shock, for instance; and (v) common trends imply that the expectation of the never-treated outcome follow the same evolution through time.

$$DID_{+,l} = \frac{\sum_{t=l+2}^{NT} N_{t,l}^1 DID_{+,t,l}}{\sum_{t=l+2}^{NT} N_{t,l}^1}$$

$DID_{+,l}$ is a weighted average of the $(DID_{+,t,l})_{t \in l+2,...,NT}$. Chaisemartin and D'Hautefoeille (2020) establish that $DID_{+,l}$ is an unbiased estimator of the average effect of having switched from untreated to treated for the first time l periods ago.

Similarly, let:

$$DID_{+,l}^D = \frac{\sum_{t=l+2}^{NT} N_{t,l}^1 DID_{+,t,l}^D}{\sum_{t=l+2}^{NT} N_{t,l}^1}$$

be a weighted average of the $(DID_{+,t,l}^D)_{t \in l+2,...,NT}$. The estimator of σ_+^{tru} is the ratio of weighted averages of the $DID_{+,l}$ and $DID_{+,l}^D$. Let:

$$w_{+,l} = \frac{\sum_{t=l+2}^{NT} N_{t,l}^1}{\sum_{l=0}^{L_{nt}} \sum_{t=l+2}^{NT} N_{t,l}^1}$$

Then, let:

$$\hat{\delta}_+^{tru} = \frac{\sum_{l=0}^{L_{nt}} w_{+,l} DID_{+,l}}{\sum_{l=0}^{L_{nt}} w_{+,l} DID_{+,l}^D}$$

which is the unbiased estimator of δ_+^{tru}. Then, Chaisemartin and D'Hautefoeille (2020) proposes placebo estimators of the assumptions underlying $\hat{\delta}_+^{tru}$. For any $l \in \{0, ..., \left[\frac{T-3}{2}\right]\}$ and $t \in \{2l+3, ..., T\}$, let

$$DID_{+,t,l}^{pl} = \sum_{g:F_{g,1=t-l}} \frac{N_{g,t}}{N_{t,l}^1}(Y_{g,t-2l-2} - Y_{g,t-l-1}) - \sum_{g:F_{g,1}>t} \frac{N_{g,t}}{N_t^{nt}}(Y_{g,t-2l-2} - Y_{g,t-l-1})$$

Finally, let:

$$DID^{pl}_{+,l} = \frac{\sum_{t=2l+3}^{NT} N^1_{t,l} DID^{pl}_{+,t,l}}{\sum_{t=2l+3}^{NT} N^1_{t,l}}$$

The $DID_{+,t,l}$ and the $DID^{p}l_{+,l}$ forms the event study together. The common trends assumption relies on the placebo estimations. A modification of the estimator can be made by allowing covariates. In that case, the common trends assumption could be achieved even if the unconditional DID does not hold, provided the covariates are able to explain deviations from common trend.

I take the Chaisemartin and D'Hautefoeille (2020) version with controls as my main specification.

As suggested above, common trends could be checked by the placebos. That is also true for strong exogeneity, since there would be some sort of Ashenfelter dip before treatment. The assumption of independent groups is defended on the grounds that the vast majority of towns are already neighbors with a connected town, so the treatment of some of its neighbor could generate only a small contamination.

3.7 Results and discussion

3.7.1 Labor markets and firm dynamics

I begin by testing the impact of the arrival of paved feeder roads on outcomes of labor markets. Both estimators reported on the Figures will not differ much in most exercises. For reasons reported in Section 3.6, I choose to focus mainly on the estimator by Chaisemartin and D'Hautefoeille (2020) with controls as the main specification.

I first look at the dynamics of admissions, resignations and wages, in Figure 3.3. There is no interesting dynamics emerging on admissions, with results remaining insignificant for the whole period. Resignations grow starting at treatment year, with 18.5% growth in resignations

(0.185, ci=95%). Point estimates keep growing during the period, but significant estimates are found one year after treatment, with 16.9% growth (0.169, ci=90%) and 3 years after the treatment, with 35.3% growth (0.353, ci=95%) relative to the first pre-treatment year. Since effects remain for a long period, one would assume the results could be generating unemployment, informality, migration or the opening of new businesses.

The increase in resignations without an increase in admissions suggests, at first instinct, that wages would suffer a negative shock. This is not what is seen in the wage graph, although one can see a trend present in the estimates of C&D'H, which could make the DID interpretation invalid. Further examination of the heterogeneity between sectors below will reveal different wage dynamics.

(Asher and Novosad, 2020) found that the arrival of a paved road has generated a small shift towards wage labor. Presumably, people would either abandon their subsistence activity to join the workforce because of higher wages or because it would be profitable to sell the land to more productive buyers aiming at increasing production. Another possibility, as revealed by (Aggarwal, 2018), would be that individuals leave school at teenage years to participate in the labor force, although (Adukia et al., 2020) report positive schooling effects for all educational cohorts. Our results suggest the opposite: there is a decrease in wage labor activity, at least inside the municipality and for formal labor markets. That alone asks for better understanding of the mechanisms at work. Migration to foreign labor markets could be a positive impact of the arrival of paved roads if domestic labor markets ends up exporting labor, as in (Faber, 2014). If unemployment or informality are growing instead, there could be substantial frictions preventing job transition, even for low skilled workers, as there is evidence for the Brazilian trade opening of 1990 (Dix-Carneiro, 2014; Dix-Carneiro and Kovak, 2019).

Moving to Figure 3.4, and looking at the firm side, established firms do not appear to exhibit interesting growth dynamics. They remain on the same path as before treatment. Entry and exit of firms are more interesting. I do not use controls for entry and exit, since controls are mainly firm level data one can track through time. A firm counts as an entrant or exiter at a given point in time, so my main focus is in the unconditional estimator of C&D'H. Entry grows 17.3% in the treatment year (0.173, ci=95%), 20.2% in the first year after treatment (0.202, ci=90%) and 20.8% (0.208, ci=90%), and exit falls - 31.2% 2 years after treatment (-0.312, ci=95%), - 52.4% 3 years after treatment (-0.524, ci=95%) and - 59.1% 4 years after treatment (0.591, ci=95%). Although established firms did not grow or shrink, it does appear that their chances of exiting has fallen. The entry dynamics would suggest improving conditions at a first glance, but one cannot rule out entry is related to the increase in resignations reported above. Another important caveat on the entry and exit results is that I am talking about low levels. Translating these results to levels, one would find an increase in entry of only 1.5 firms 1 year after treatment (1.50, ci=95%), and a decrease in exit of only - 2.65 firms 4 years after treatment (-2.65, ci=95%), results that are barely meaningful in economic terms.

As discussed above, a possibility for the resigned worker would be to produce as an entrepreneur. Even though I cannot check that mechanism with our data, the amount of extra workers resigning from treated areas are too large for the very small number of extra entrants. Unemployment, informality and migration are better candidates to explain the fate of resigned workers.

Figure 3.3 Treatment Effects on Labor Markets and Firm Dynamics (percent change)

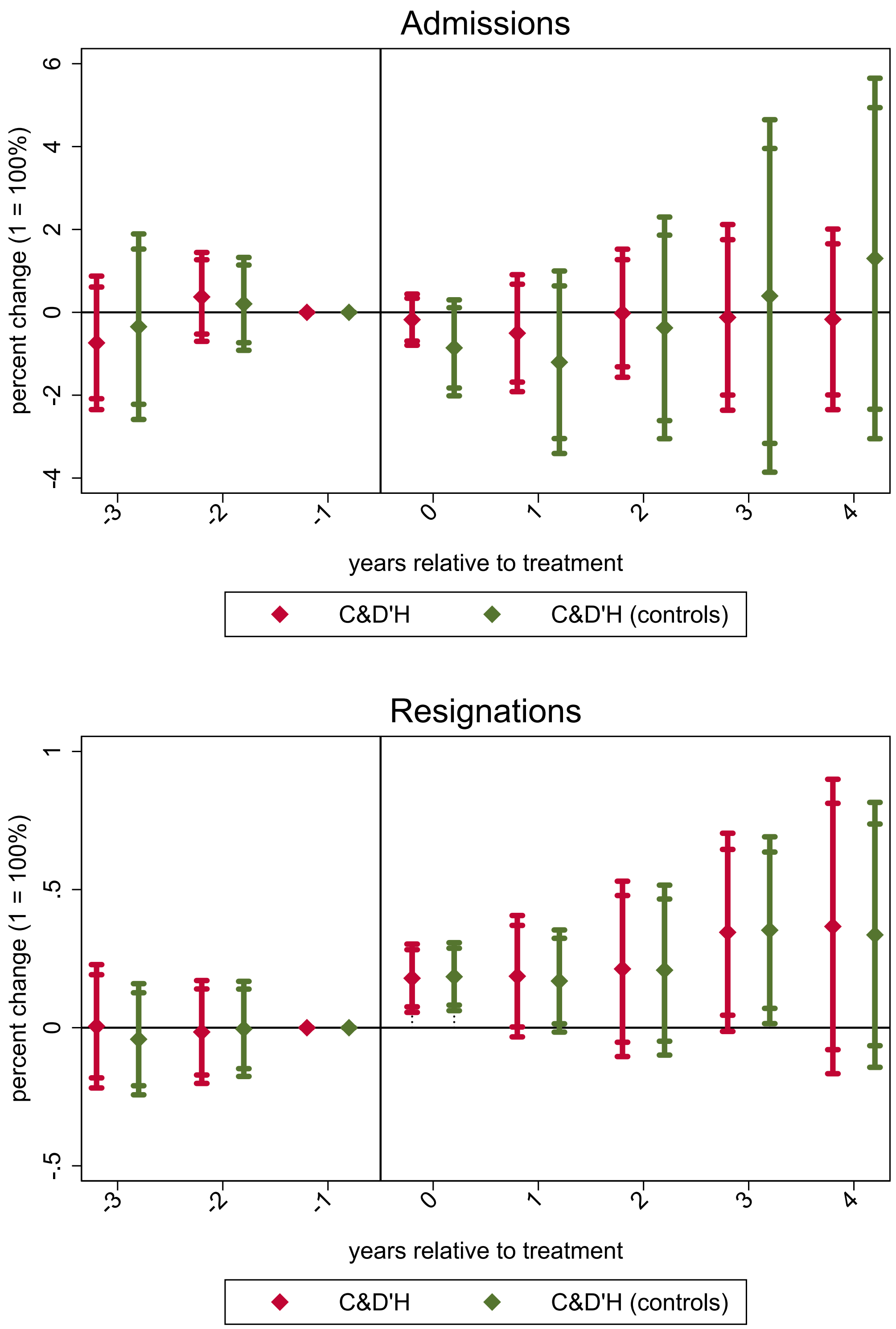


Figure 3.3 Treatment Effects on Labor Markets and Firm Dynamics (percent change)

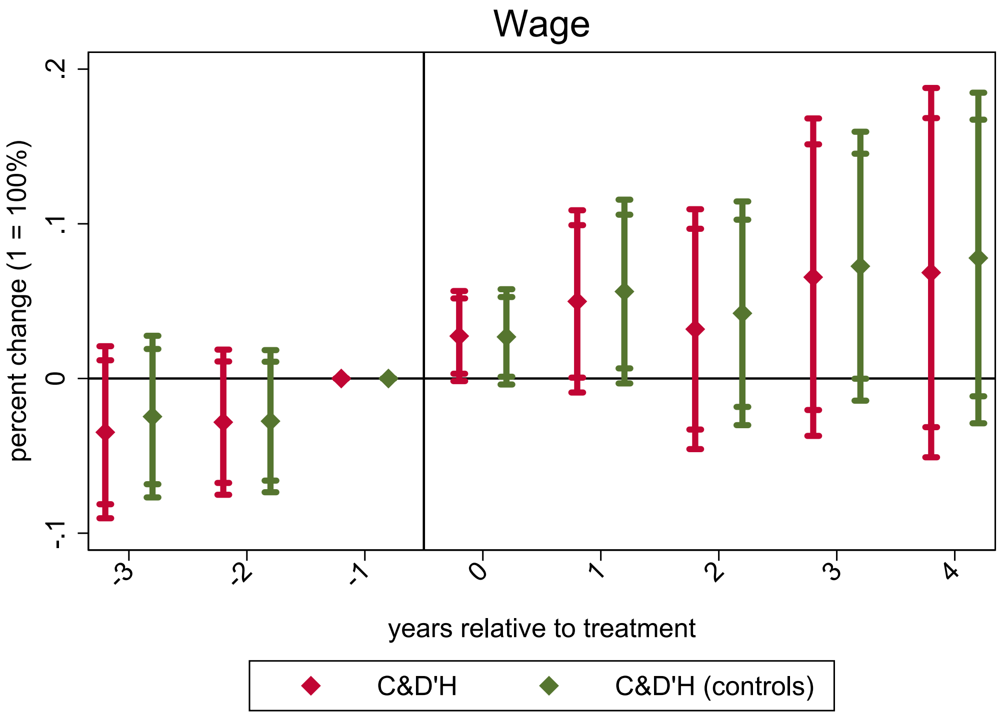
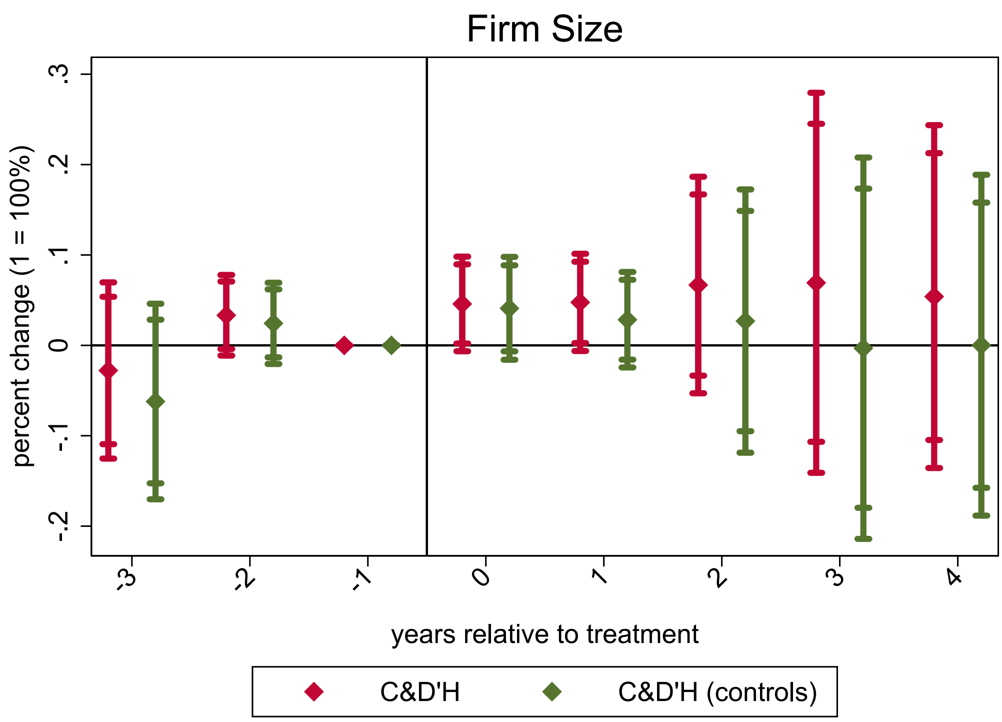


Figure 3.3 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Admissions, Resignations and Wage. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.7.

Figure 3.4 Treatment Effects on Labor Markets and Firm Dynamics (percent change)



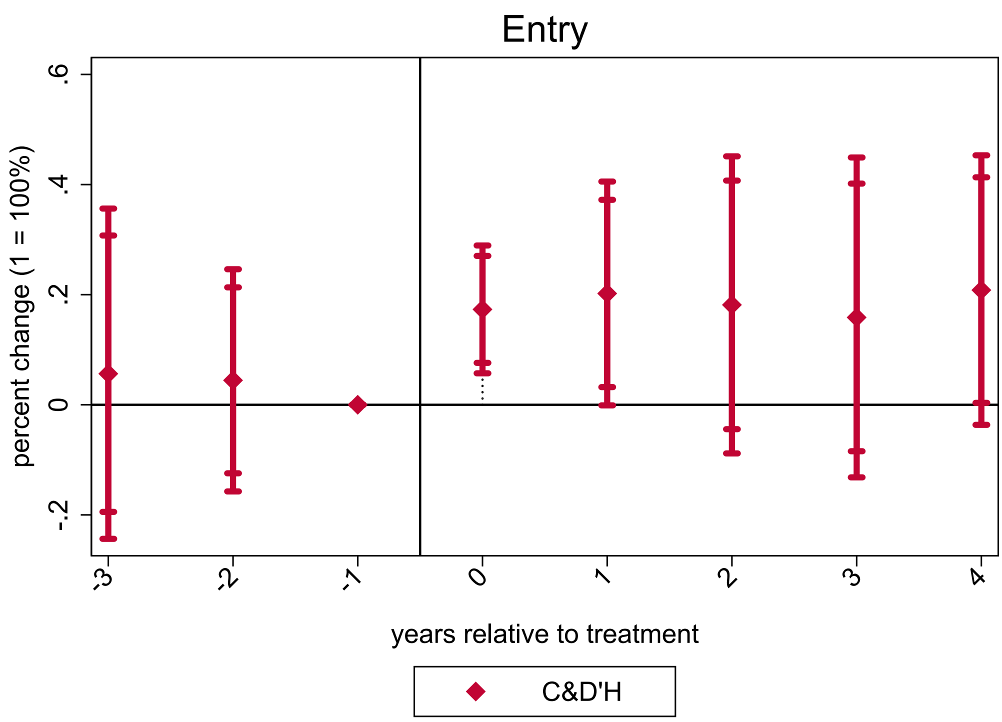


Figure 3.4 Treatment Effects on Labor Markets and Firm Dynamics (percent change)

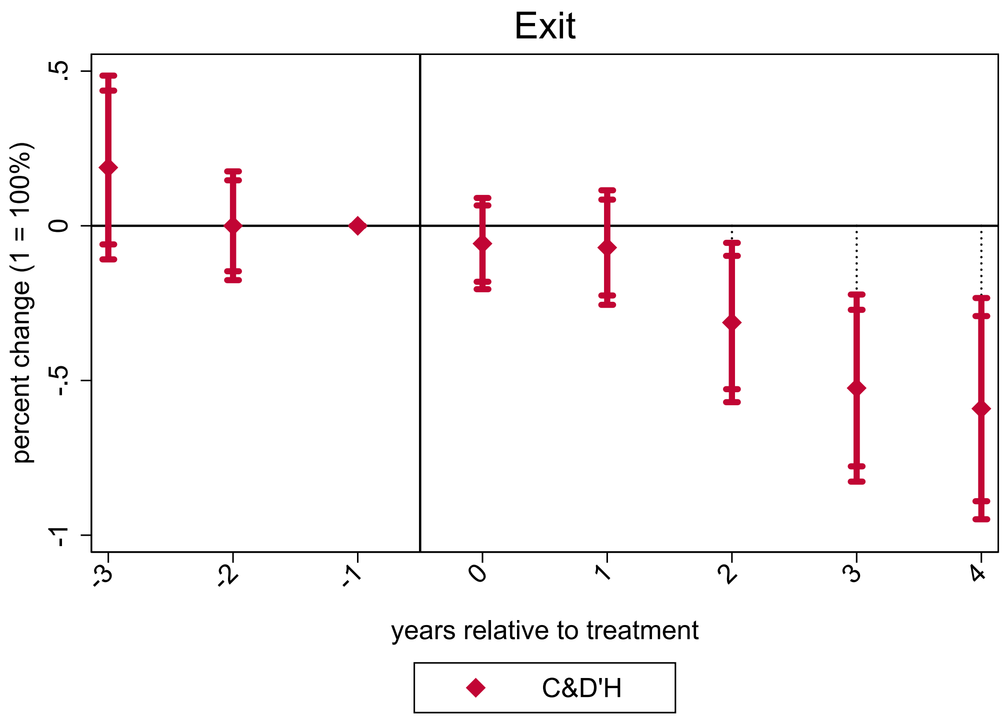


Figure 3.4 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Established Firm Size, Entry and Exit. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.8.

3.7.2 Heterogeneity in labor markets and firm dynamics

I move to heterogeneity by sector in order to spot if changes are mainly in tradable markets or in the service sector. I report only the interesting results helping to unpack the aggregate dynamics.

I start with resignations by sector, in Figure 3.5. An increase in resignations clearly show up in agriculture right after treatment. It grew by 26.3% (0.263, ci=90%) in the treatment year, and reached 55.2% 3 years after treatment (0.552, ci=90%). Resignations seems to have increased notably in manufacturing in the first year after treatment (0.860%, c.i=90%), but it returned to pre-treatment dynamics by the end of the time period, while agriculture seems to have kept the new levels of resignation, although there is a high uncertainty about that. As for services, there is an increase in resignations right at the treatment year, of 16.8% (0.168, ci=95%), while 3 years after treatment, resignations increase by 29.5% (0.295, ci=90%). All sectors contributed to the increase in resignations, but mostly agriculture and services, which also happen to be the largest ones.

Moving on to wages by sector, in Figure 4, it is in agriculture that the shock is most severe. Two years after treatment, wages fell by 8.36% (-0.836, ci=95%), and then it was still 6.40% lower than the pre-treatment year in the next year (0.640, ci=90%). Manufacturing wages were not significant, but it seems like there was an increase in wages in the treatment year. Results for services are also not significant, although it mimics the developments of aggregate data, with a possible trend present.

(Shamdasani, 2021)'s research points to a reallocation of workers out of agriculture on aggregate, but also captures an increase in admissions in certain treated villages, attracting agricultural workers from other treated villages. The lack of any dynamics in admissions in agriculture suggest that this labor market connection for agricultural workers is not relevant for our domestic municipalities. They do not attract workers in agriculture, and also not in manufacturing and services.

The most interesting dynamics comes from resignation by sector and wages by sector. The most clear correlation is between an increase in resignations and a lower wage in agriculture. That is most clear two and three years after treatment. A first intuition suggests a negative demand shock. Still, there is the possibility of a change in the composition of the agricultural work force. If higher skilled workers earning better wages are attracted to jobs elsewhere, average wages would also fall. A first glance at the mechanism at work could be seen on the overall and sector admission dynamics: they follow the same trajectory pre-treatment, suggesting that if the higher skilled workers are leaving domestic agriculture to a better job, they must be migrating, since it is unlikely they would choose unemployment or informality. This information narrows potential channels to a demand shock displacing workers or workers migrating to foreign markets.

As for the apparent positive correlation between resignation and wages in the service sector, that could be explained by the consolidation of more productive services and the termination of low wage workers' contracts due to the lesser demand coming from the fall in agricultural activity or from the migration of low wage, presumably lower skilled workers.

Figure 3.7 reveals that entry is mostly happening in agriculture, with a growth of 36.7% in the treatment year (0.367, ci=95%), 26.3% in the first year after treatment (0.263, ci=90%) and 48.3% 4 years after treatment (0.483, ci=90%). Figure 3.8 reveals that exit dynamics comes from both agriculture, with a lower exit of 57.2% 4 years after treatment, and services, with a lower exit rate of 30% in the treatment year (-0.300, ci=95%), 60% 2 years after the treatment (-0.600, ci=95%), 107.1% 3 years after treatment (-1.071, ci=95%), and 88.8% 4 years after treatment (0.888, ci=95%). It is important to bear in mind that this result is not as impressive in levels, as shown for the aggregate.

Figure 3.5 Treatment Heterogeneous Effects on Resignations by Sector (percent change)

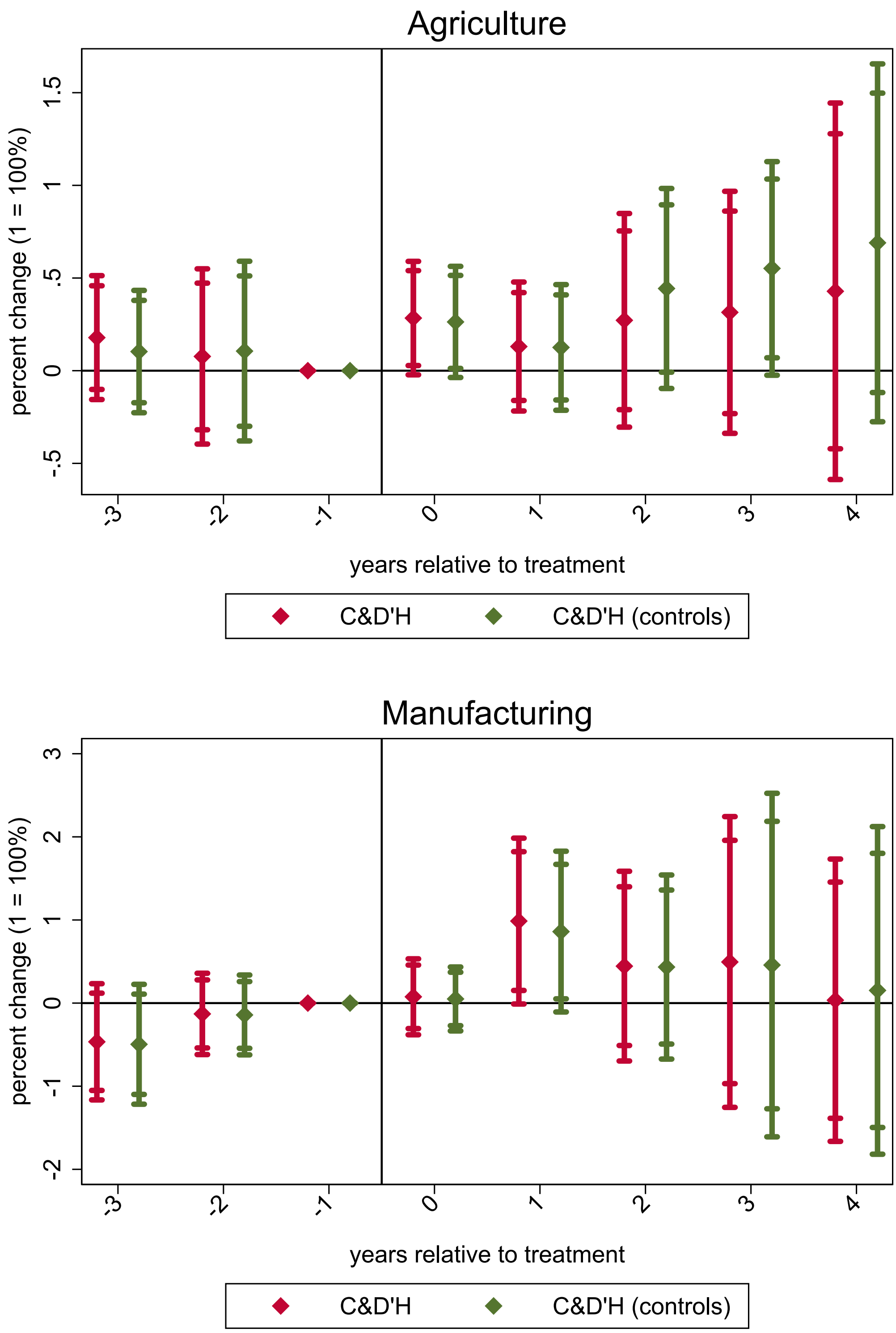


Figure 3.5 Treatment Heterogeneous Effects on Resignations by Sector (percent change)

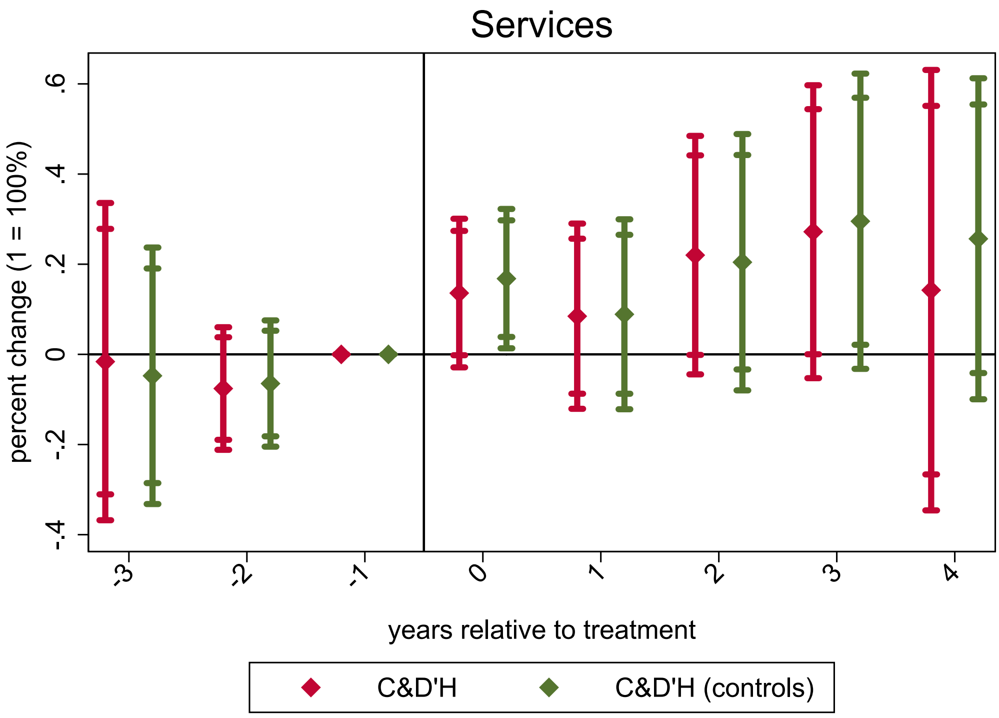
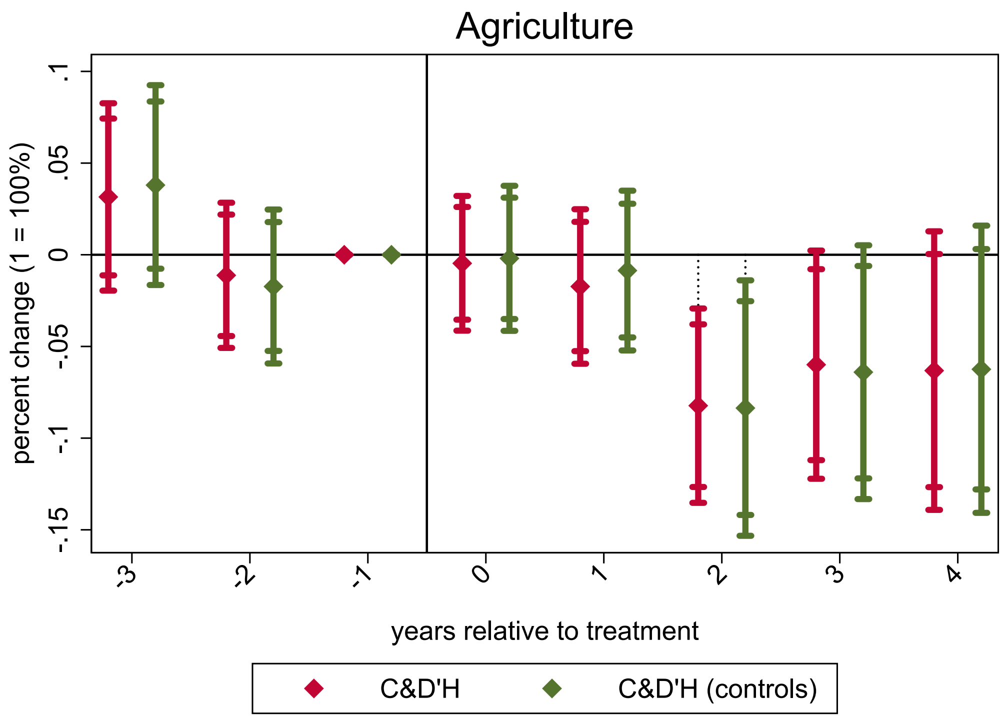


Figure 3.5 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Resignations in the Agriculture, Manufacturing and Service sectors. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.9.

Figure 3.6 Treatment Heterogeneous Effects on Wage by Sector (percent change)



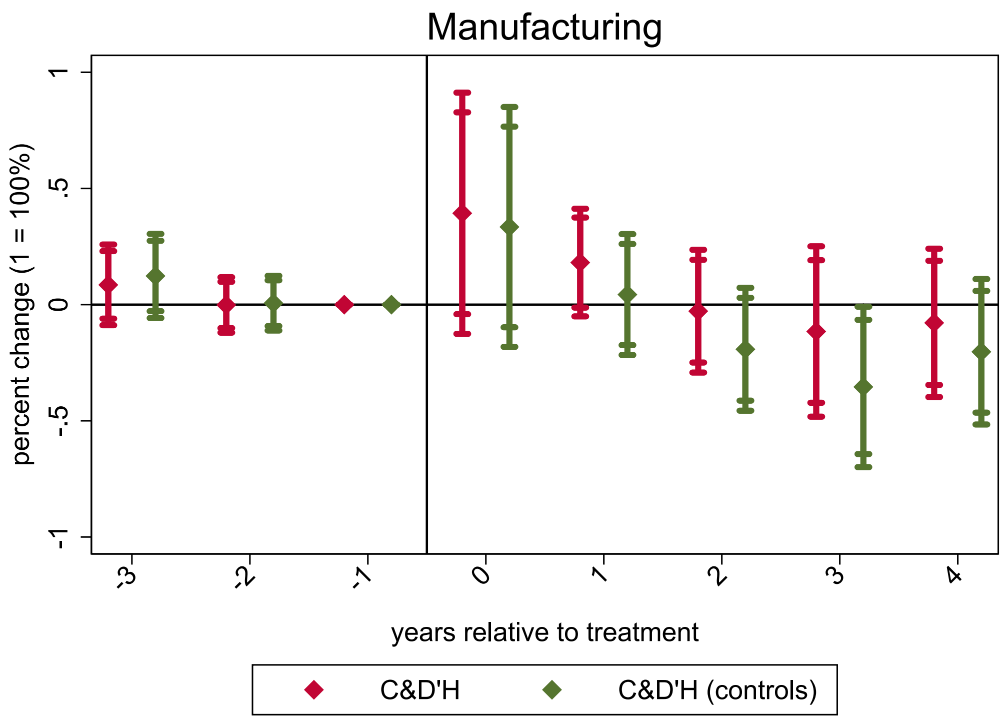


Figure 3.6 Treatment Heterogeneous Effects on Wage by Sector (percent change)

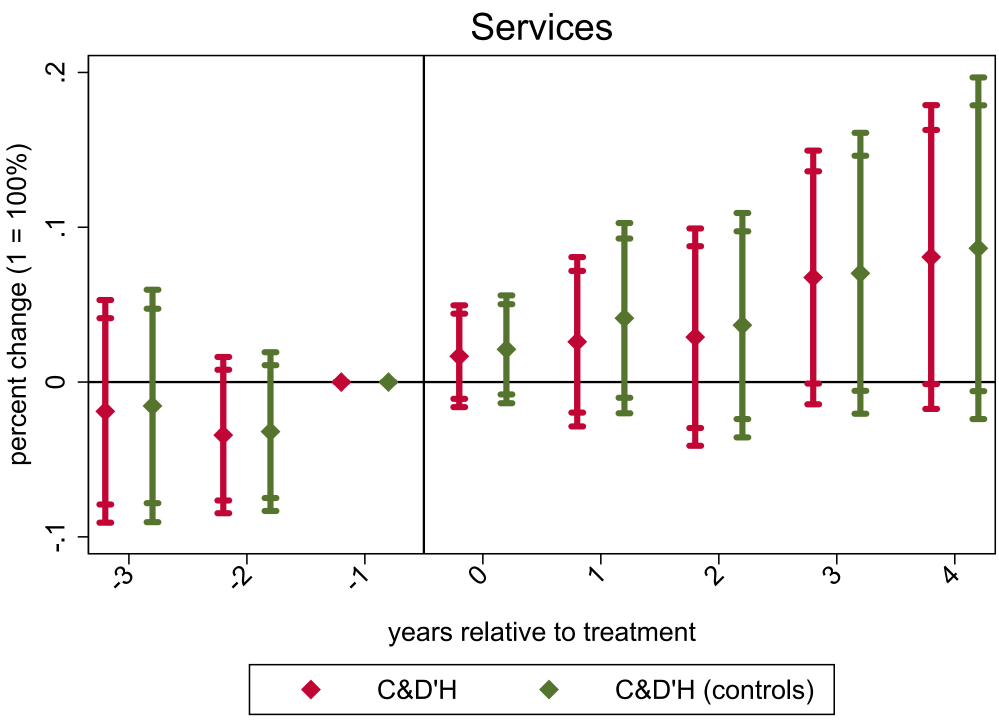
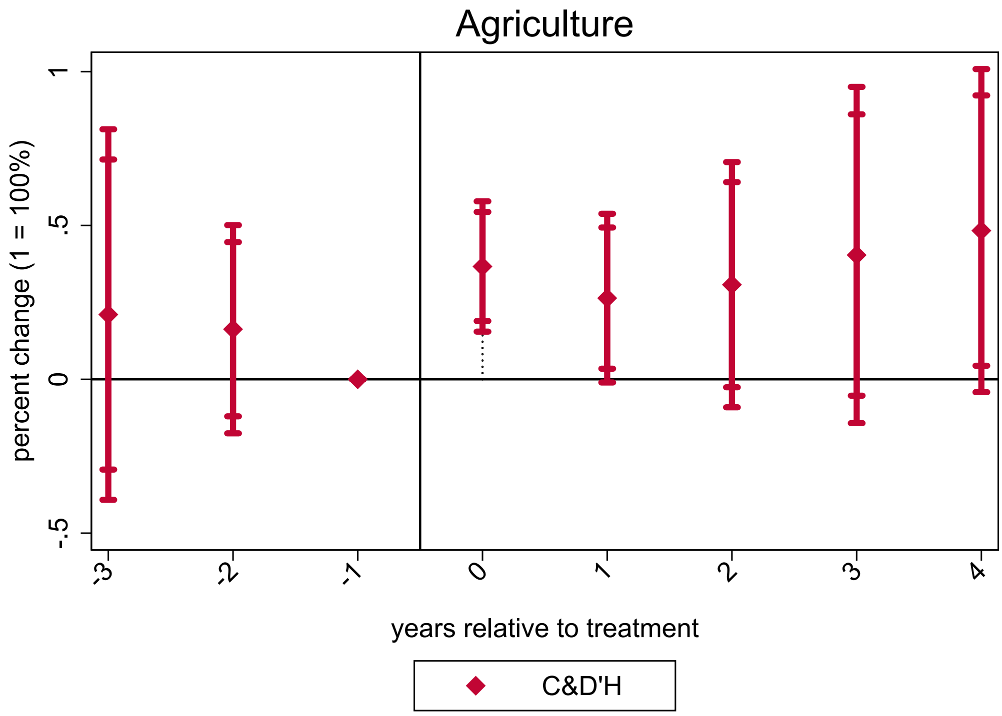


Figure 3.6 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Wage in the Agriculture, Manufacturing and Service sectors. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.10.

Figure 3.7 Treatment Heterogeneous Effects on Entry by Sector (percent change)



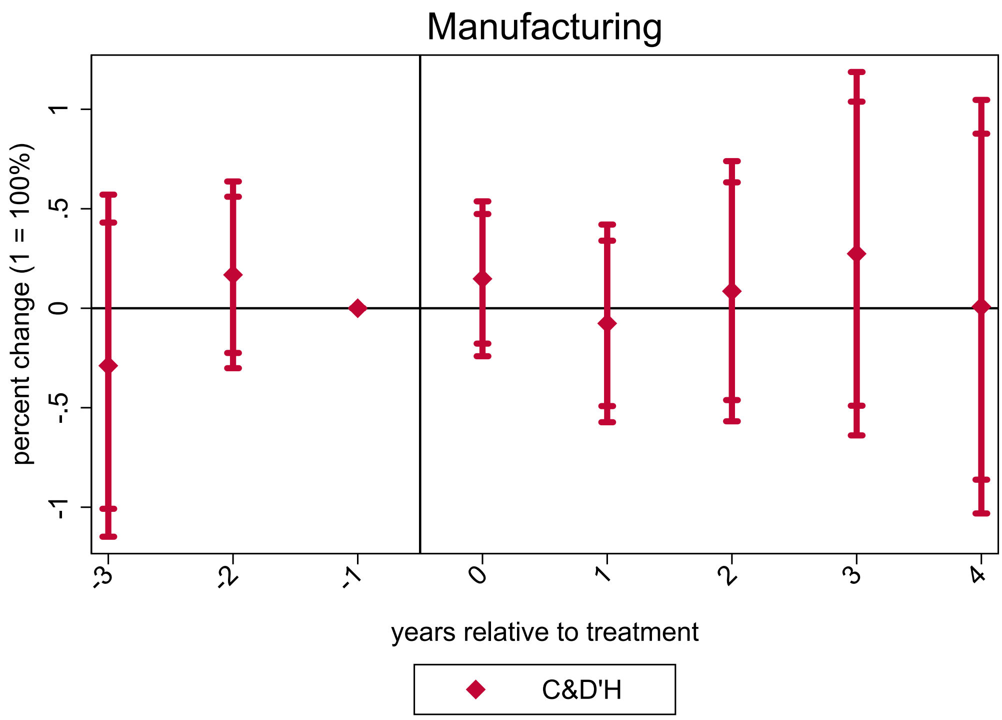


Figure 3.7 Treatment Heterogeneous Effects on Entry by Sector (percent change)

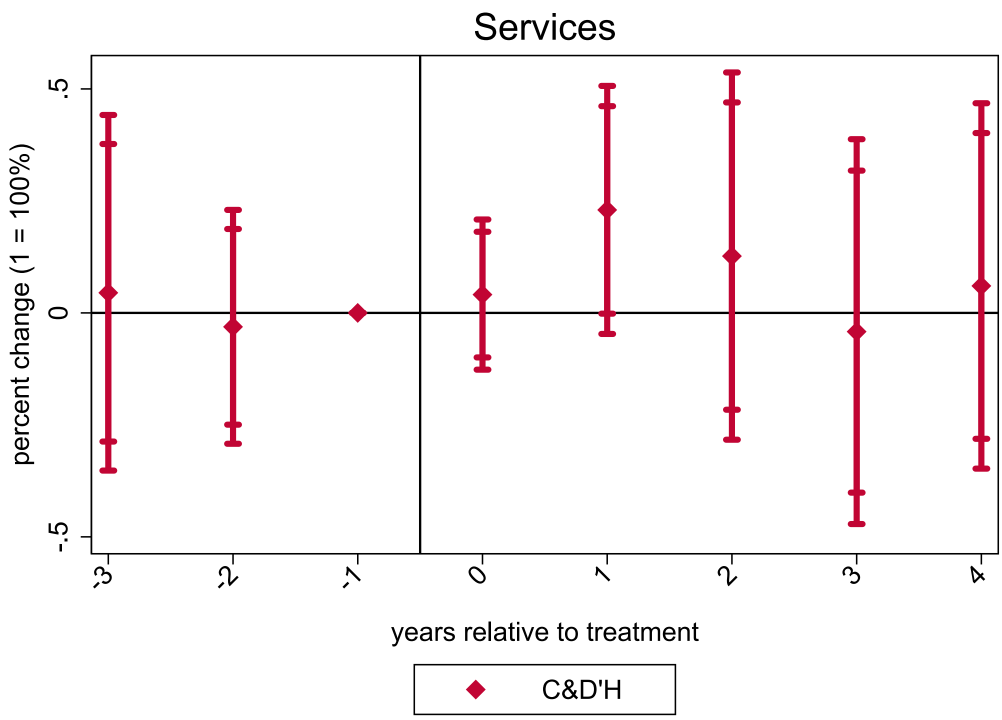
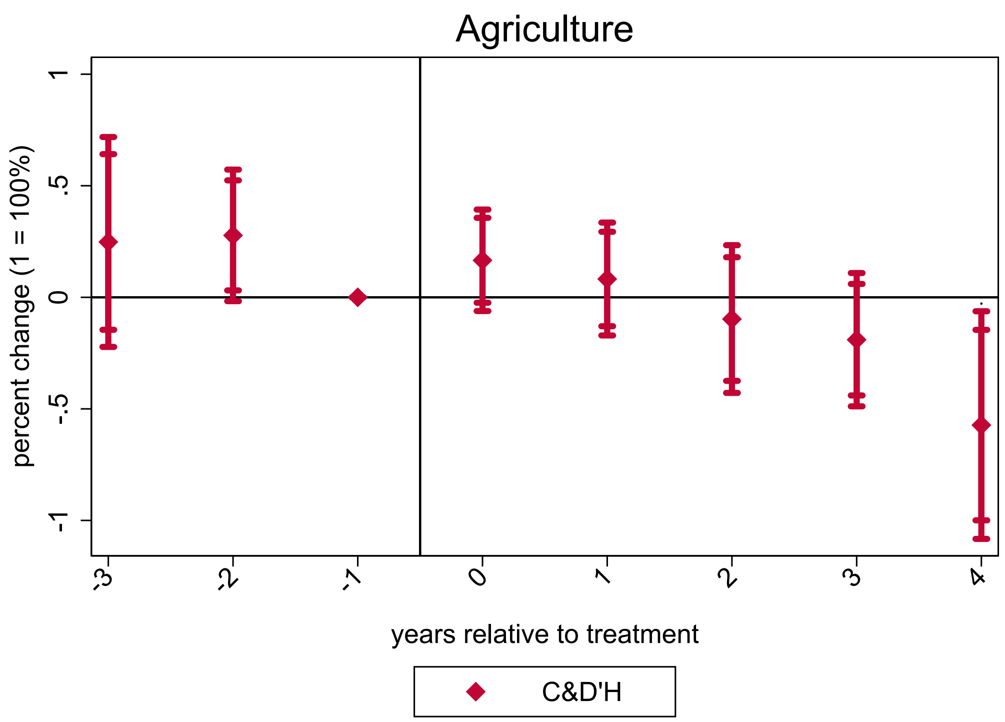


Figure 3.7 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Entry in the Agriculture, Manufacturing and Service sectors. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.11.

Figure 3.8 Treatment Heterogeneous Effects on Exit by Sector (percent change)



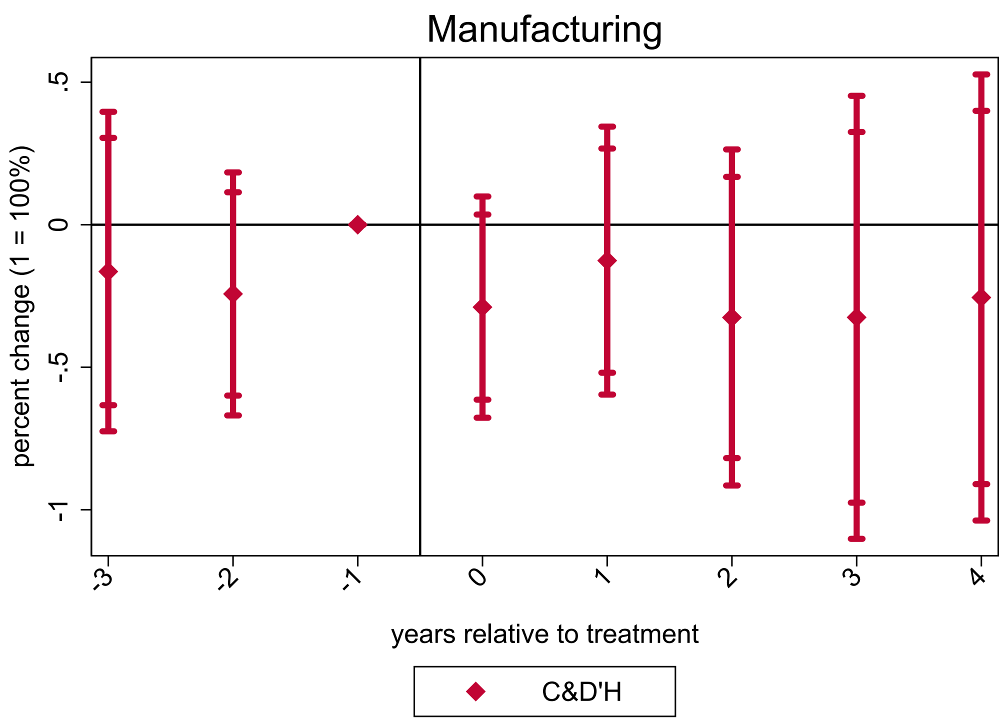


Figure 3.8 Treatment Heterogeneous Effects on Exit by Sector (percent change)

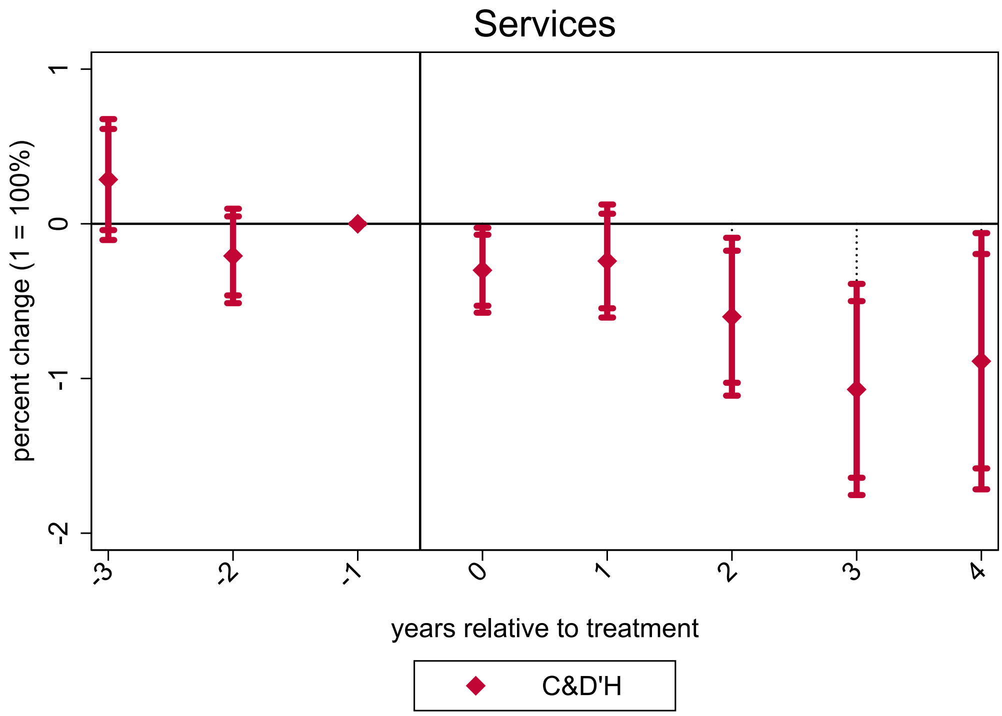


Figure 3.8 ***Notes:*** Event study results from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Exit in the Agriculture, Manufacturing and Service sectors. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous and dynamic effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. Details at appendix table 4.12.

3.7.3 Evidence for Channels

Since agriculture had a more interesting dynamics, I now focus on data about agricultural production, including planted area and yields. I disaggregate between cash crops and food crops[11]. Not only this allows for interesting heterogeneity in the agricultural production, but provide an important evidence against the willing migration of higher skilled agricultural workers.

The first noticeable result is how food crops' planted area are affected by the arrival of a paved road, in Figure 9. Two years after treatment, the planted area was reduced by 11.5% (0.115, ci=95%). An year later, the planted area for food crops was reduced by 21% relative to the planted area one year before treatment. Cash crops were not affected. That generates a shift up to 10 p.p. in the share of cash crops 3 years after treatment (0.100, ci=95%).

Presumably, yields were affected by the reduction in planted area for food crops, which is what I verify in Figure 3.10. Yields fell by 35.5% 2 years after treatment (0.355, ci=95%), and were still down 33.6% 3 years after treatment (0.336, ci=90%). Results for cash crops post-treatment are too uncertain, although it does appear to rise by the end of the period.

Evidence for the fall in planted area and yields are in accordance to a reduction in agricultural employment. I already have some suggestion from the lack of domestic admissions that resigned workers in agriculture have not established elsewhere domestically. The asymmetry between cash crops and food crops' dynamics suggests that resigned workers came mostly from food crops, unless cash crops producers became more productive and displaced workers. Intuitively, it makes sense that food crops were the most affected market. Cash crops are generally export oriented. Their overall transport cost includes road or rail transportation until ports and then shipping. A smaller section of new paved roads should not dramatically alter their incentives. Food crops producers, on the other hand, are generally smaller and provide food for domestic and/or regional markets nearby. An all-weather road could change their situation, either facilitating the export of competitive food products or exposing them to foreign compe-

[11] Their definition are available in the "Data" section, "PAM" subsection.

tition. The latter appears the most convincing explanation for our analysis. Competition from regional markets appear to have generated a negative demand shock for domestic food crops, which reduced their planted area, yields, and forced farms to terminate some workers' contract or they ended up leaving due to lower wage prospects.

The fact that the reduction in employment came most likely from the production of food crops implies that the willing migration channel might not be the best candidate to explain the resignation of agricultural workers. If that was the case, one would expect that cash crops would also be affected by the decision to migrate, specially given that the effect on composition would have to come from higher skilled workers leaving the domestic market. That also points to a negative demand shock on food crops.

To further support the evidence of a negative demand shock on the agricultural sector, I take advantage of the availability of data on the cause of contract termination. If a negative demand shock occurs, one would expect that firms would react first by terminating contracts. While workers could, in principle, also terminate their contracts or opt to not renovate it if wages are falling, one would expect that their reaction would be slower than firms'. Figure 3.11 confirms this intuition. Although it does appear that employees are also increasing contract termination after treatment, which we cannot confirm at 90% c.i. due to high uncertainty, employer contract termination appear to start rising immediately. Two years after treatment, there are 35.854 more resignations by contract termination (c.i.=90%), and that keeps growing to 53.631 (c.i.=90%) and 72.074 (c.i.=95%) in the third and fourth year after treatment, relative to control. If I take point estimates from employee termination seriously, ignoring confidence intervals, it raises much slower, to 16.244 after 4 years, suggesting that either a small fraction of the resignations come from individuals who decided to migrate to better opportunities in foreign or domestic markets, presumably more because of the labor market condition in the domestic agricultural sector than because of the appearance of much better opportunities elsewhere.

The combined evidence appear to support the thesis that the main effect on labor markets happens in the agricultural sector specialized in food crops. Foreign competition shrinks planted area and yields domestically, causing the demand for agricultural workers to fall together with wages, and leading to contract termination that most likely leave these workers unemployed or in informality.

Figure 3.9 Treatment Effects on Planted Area

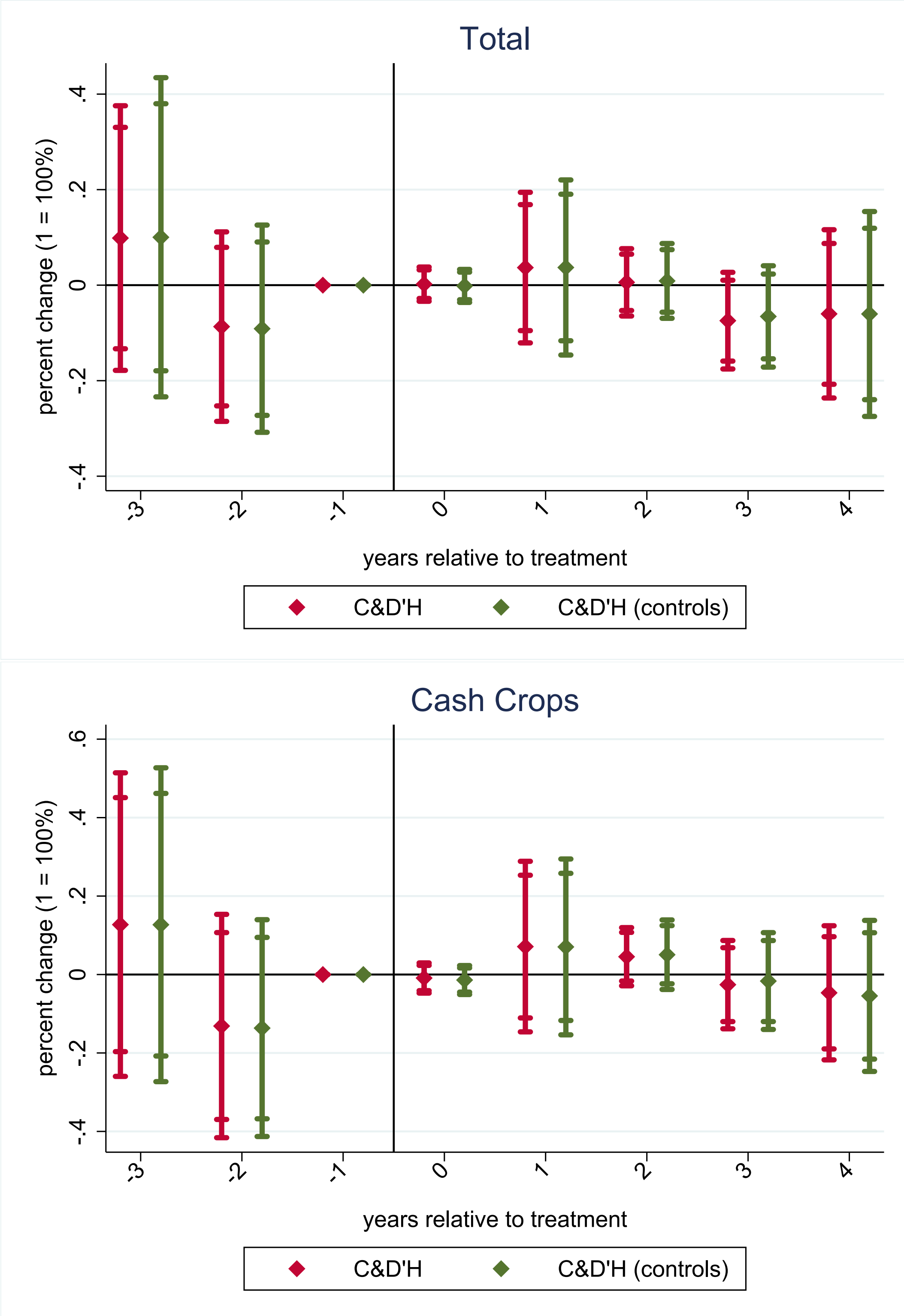


Figure 3.9 Treatment Effects on Planted Area

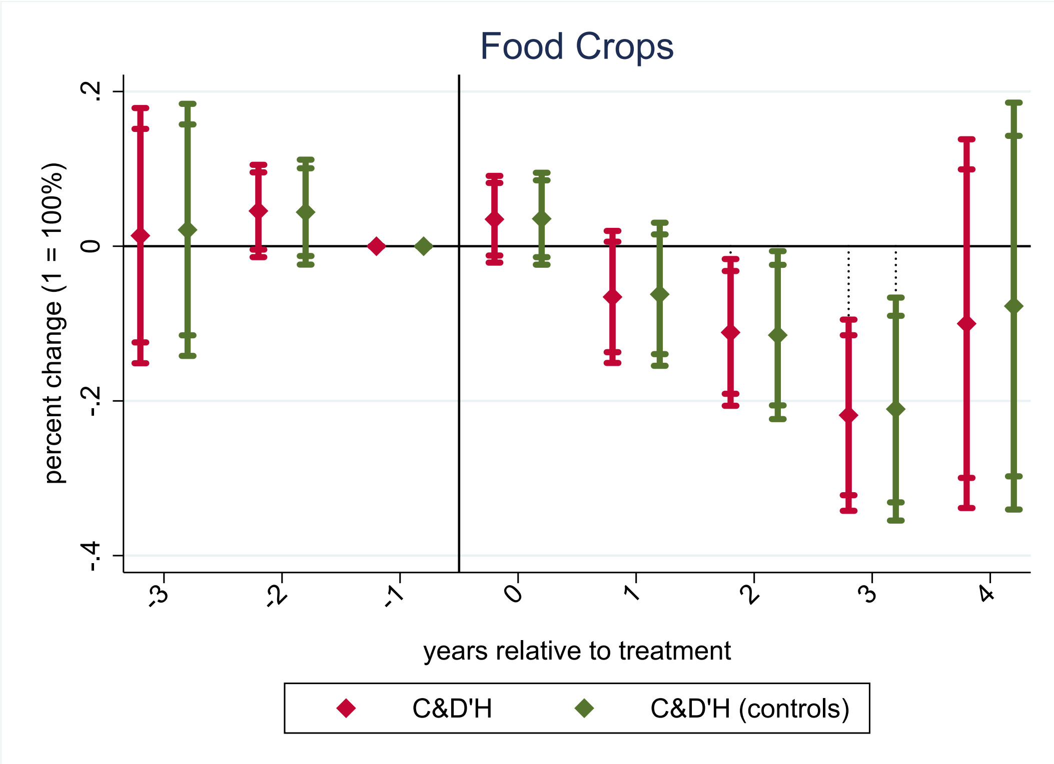


Figure 3.9 ***Notes:*** Event study plots on Planted Area from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Total planted area, Cash Crops planted area and Food Crops planted area. Results for the Share of Cash Crops relative to Food Crops are in percentage points. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. For more details, the full table is available in the Appendix table 4.13.

Figure 3.10 Treatment Effects on Yields (percent change)

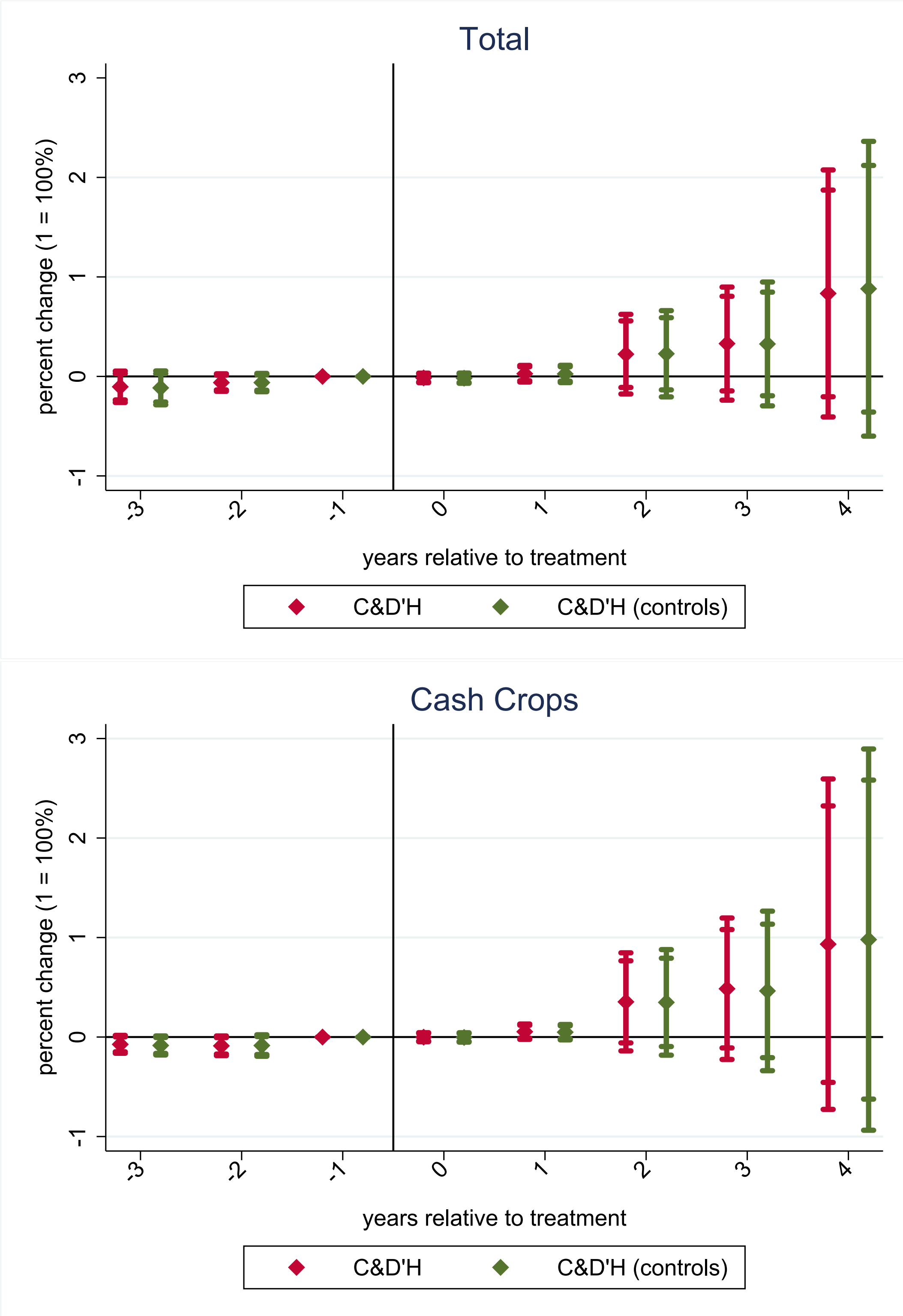


Figure 3.10 Treatment Effects on Yields (percent change)

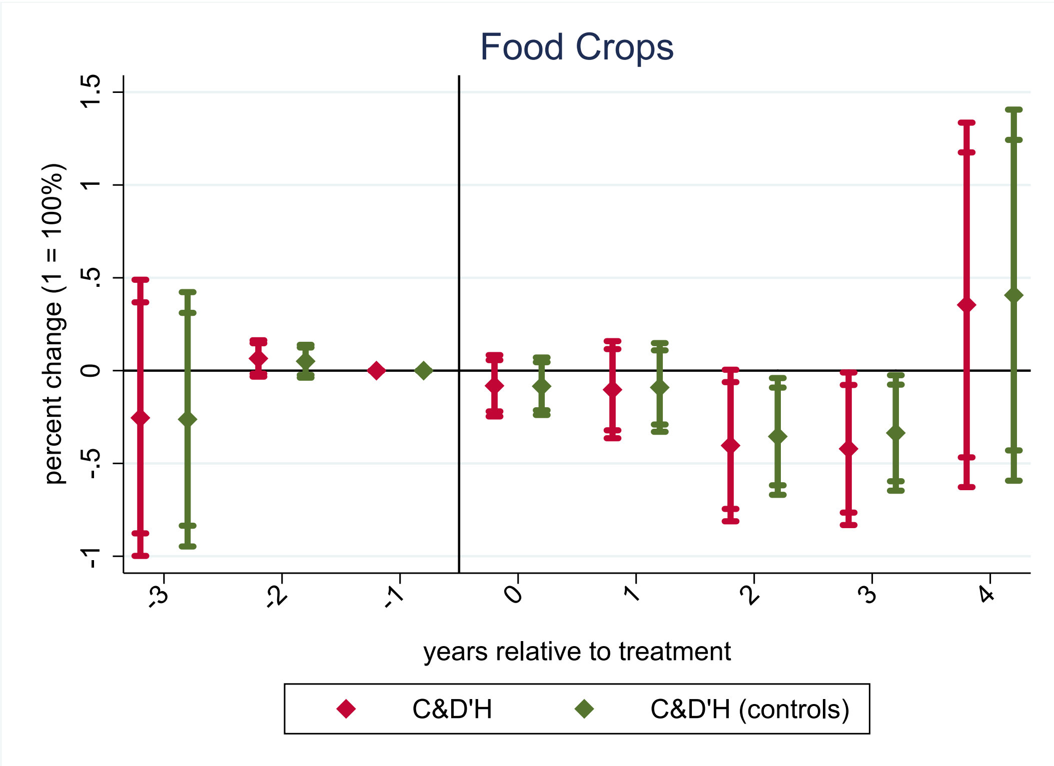


Figure 3.10 ***Notes:*** Event study plots on Yields from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Total Yields, Cash Crops Yields and Food Crops Yields. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. For more details, the full table is available in the Appendix 4.14.

Figure 3.11 Treatment Effects on Resignations in Agriculture by Initiative (levels)

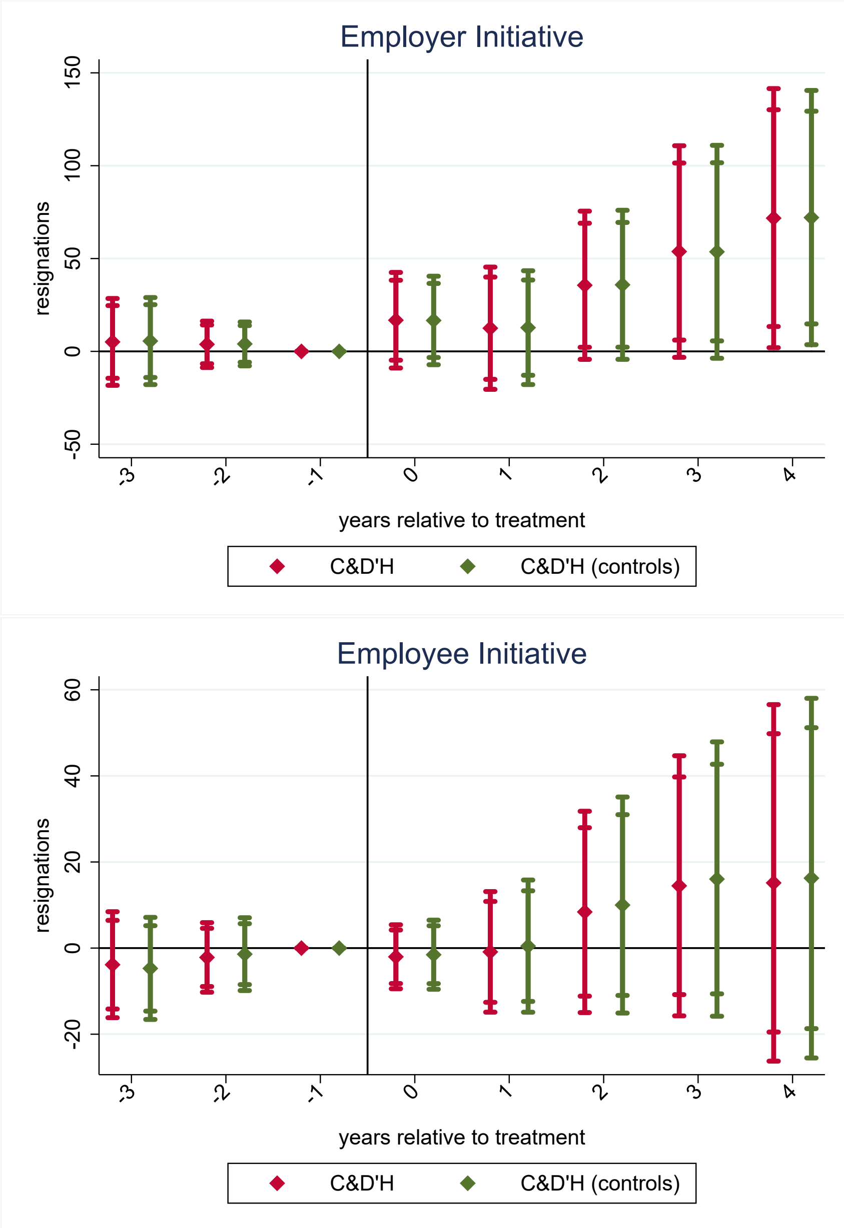


Figure 3.11 ***Notes:*** Event study plots on Planted Area from 3 years before treatment to 4 years after. Results are in percent change relative to the average of the first year of pre-treatment for Total Production Value, Cash Crops Production Value and Food Crops Production Value. Estimates were constructed using the estimator of Chaisemartin & D'Hautefeille, with and without controls, both of which robust to heterogeneous effects. Confidence intervals are reported for 95% and 90% confidence levels. Standard errors are clustered at the municipal level. For more details, the full table is available in the Appendix 4.16.

3.8 Conclusion

This chapter estimated the causal effect of access to a paved feeder road for remote towns in Minas Gerais, Brazil, on outcomes of labor market and firm dynamics.

My main result is that treatment produced an important shock in agriculture, particularly in food crops. The combination of (i) a reduction in 21% in planted area 3 years after treatment; (ii) a reduction of 35.5% in yields 2 years after treatment and still 33.6% lower yields 3 years after treatment; (iii) 35.3% growth in resignations 3 years after treatment; (iv) a reduction of 8.36% and 6.40% in wages for agricultural workers 2 and 3 years after treatment, respectively; (v) an asymmetric impact on food crops and cash crops; and (vi) a larger and faster impact on the employer termination of contacts in relation to employee contract termination, paints a clear picture for a large negative demand shock in the agricultural sector, suggesting food prices were higher domestically relative to foreign markets. Presumably competition was able to bring cheaper food. Since entry is statistically significant but not economically significant, it cannot account as the main destiny of resigned workers. In addition, since admissions are not faster anywhere in the formal sector of the domestic markets, workers either end up unemployed, at informality or are forced to migrate to foreign markets in search of opportunity. Recent literature on the impacts of international trade on domestic labor markets do suggest that the displacement of workers could have long term consequences to those displaced (Dix-Carneiro, 2014; Dix-Carneiro and Kovak, 2019).

Manufacturing remains basically stagnant, although some short term activity shows up in the results. One would not expect much activity anyway, since there are good theoretical arguments suggesting the sector needs agglomeration economies to grow (Krugman, 1991).

Increase import competition in the agricultural sector implies cheaper and/or better products for consumers. If that is the case, as established by our results, I also have good reason to suspect consumers would also benefit from cheaper and better manufacturing products, presumably available at stores and supermarkets.

The formal service sector also appear to slightly shrink. That could be a result from the fall in agricultural production implying less demand for domestic services. But it could also suggest a productivity boost (since wages appear to not be negatively affected), or lesser skilled workers from the service sector could have migrated to foreign markets.

This paper joins the literature estimating causal effects on large infrastructure projects, particularly that of rural/feeder roads, which has been improved recently by better research designs. (Adukia et al., 2020; Aggarwal, 2018; Asher and Novosad, 2020; Shamdasani, 2021). I showed that a demand shock might be responsible to the displacement of agricultural workers, while the literature seems to imply that the main mechanism is that workers willingly migrate for better opportunities.

References

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, 2014, *104* (6), 1667–1697.

Adukia, Anjali, Sam Asher, and Paul Novosad, "Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction," *American Economic Journal: Applied Economics*, 2020, *12* (1), 348–376.

Aggarwal, Shilpa, "Do rural roads create pathways out of poverty? Evidence from India," *Journal of Development Economics*, 2018, *133*, 375–395.

Airey, Tony, "The influence of road construction on the health care behavior of rural households in the Meru district of Kenya," *Transport Reviews*, 1991, *11* (3), 273–290.

Al-Taiar, Abdulah, Allan Clark, Joseph C. Longenecker, and Christopher J.M. Whitty, "Physical accessibility and utilization of health services in Yemen," *International Journal of Health Geographics*, 2010, *9* (38), 1–8.

Alves, Patrick and Alexandre Messa Silva, "Texto para discussão," *IPEA*, 2008, (1325).

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, 2020, *110* (3), 797–823.

Athey, Susan and Guido Imbens, "Design-Based Analysis in Differences-in-Differences Settings with Staggered Adoption," *Journal of Econometrics*, 2022, *226* (1), 62–79.

Banerjee, Abhijit and Esther Duflo, "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," *Review of Economic Studies*, 2014, *81*, 572–607.

Baqaee, David and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Bau, Natalie and Adrien Matray, "Misallocation and Capital Market Integration: Evidence from India," *Econometrica*, 2023, *91* (1), 67–106.

Bils, Mark, Peter Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," *Journal of Monetary Economics*, 2021, *124S*, S39–S56.

Borusyak, Kirill and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," *Working Paper.*

Buera, Francisco J. and Yongseok Shin, "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

_ , Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 2011, *101* (5), 1964–2002.

Buor, Daniel, "Analyzing the primacy of distance in the utilization of health services in the Ahafo-Ano South district, Ghana," *International Journal of Health Planning and Management*, 2003, *18*, 293–311.

Busso, Matias, Lucia Madrigal, and Carmen Pagés, "Productivity and Resource Misallocation in Latin America," *B.E. Journal of Macroeconomics*, 2013, *13* (1), 903–932.

Callaway, Brently and Pedro H.C. Sant'Anna, "Differences-in-Differences with multiple time periods," *Journal of Econometrics*, 2020, *225* (2), 200–230.

Casaburi, Lorenzo, Rachel Glennerster, and Tavneet Suri, "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone," *Working Paper*, 2013, pp. 1–32.

Cavalcanti, Tiago and Paulo Henrique Vaz, "Access to long-term credit and productivity of small and medium firms: A causal evidence," *Economics Letters*, 2017, *150*, 21–25.

Cervero, Robert, "Accessibility and Third World Rural Development: A Case Study of Sumatra," *Review of Urban and Regional Development Studies*, 1990, *2*, 125–138.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis, "Misallocation or Mismeasurement?," *American Economic Journal: Macroeconomics*, forthcoming.

Chomitz, Kenneth M. and David A. Gray, "Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize," *The World Bank Economic Review*, 1996, *10* (3), 487–512.

Cuong, Nguyen Viet, "Estimation of the Impact of Rural Roads on Household Welfare in Viet Nam," *Asia-Pacific Development Journal*, 2011, *18* (2), 105–135.

Dalton, Timothy J., William A. Masters, and Kenneth A. Foster, "Production costs and input substitution in Zimbabwe's smallholder agriculture," *Agricultural Economics*, 1997, *17*, 201–209.

David, Joel M. and Venky Venkateswaran, "The Sources of Capital Misallocation," *American Economic Review*, 2019, *109* (7), 2531–2567.

de Chaisemartin, Clement and Xavier D'Hautefoeille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

de Vries, Gaaitzen J., "Productivity in a Distorted Market: The case of Brazil's retail sector," *The Review of Income and Wealth*, 2014, *60* (3), 499–524.

Dercon, Stephan, Daniel O. Gilligan, John Hoddinott, and Tassew Woldehanna, "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages," *American Journal of Agricultural Economics*, 2009, *91* (4), 1007–1021.

Dias, Danial, Carlos Robalo Marques, and Christine Richmond, "Comparing misallocation between sectors in Portugal," *Economic Bulletin and Financial Stability Report. Article 201602.*, 2016, pp. 27–44.

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885.

_ **and Brian K. Kovak**, "Margins of labor market adjustment to trade," *Journal of International Economics*, 2019, *117*, 125–142.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81*, 1046–1070.

Fan, Shenggen, Bingxin Yu, and Somchai Jitsuchon, "Does Allocation of Public Spending Matter in Poverty Reduction? Evidence from Thailand," *Asian Economic Journal*, 2008, *22* (4), 411–430.

FJP, "Minas Gerais e Suas Regiões de Planejamento: Crescimento Populacional e Distribuição Espacial," 2002. Informativo CEI - Demografia.

Gibson, John and Susan Olivia, "The Effect of Infrastructure Access and Quality on Non-Farm Enterprises in Rural Indonesia," *World Development*, 2010, *38* (5), 717–726.

Gilchrist, Simon, Jae W. Sim, and Zakrajšek, "Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs," *Review of Economic Dynamics*, 2013, *16*, 159–176.

Goodman-Bacon, Andrew, “Differences-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, *225* (2), 254–277.

Haltiwanger, John, Robert Kulick, and Chad Syverson, “Misallocation Measures: The Distortion that Ate the Residual,” *NBER Working Paper*, 2018, (24199), 1–47.

Hine, J., M. Abedin, R. J. Stevens, and T. Anderson, “Does the extension of the rural road network have a positive impact on poverty reduction and resilience for the rural areas served? If so how, and if not why not? A systematic review,” *London: EPPI-Centre, Social Science Research Unit*, 2016, pp. 1–143.

Hsieh, Chang-Tai and Peter Klenow, “Misallocation and Manufacturing TFP in China and India,” *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

IDB, “Inter-American Development Bank Annual Report,” 2021. available at: https://publications.iadb.org/en/inter-american-development-bank-annual-report-2021-year-review.

Jacoby, Hanan G., “Access to Markets and the Benefits of Rural Roads,” *The Economic Journal*, 2000, *110*, 713–737.

Kalemli-Ozcan, Sebnem and Bent E. Sørensen, “Misallocation, Property Rights, and Access to Finance: Evidence from within and across Africa,” in “African Successes, Volume III: Modernization and Development” University of Chicago Press 2016, pp. 183–211.

Krugman, Paul, “Increasing Returns and Economic Geography,” *Journal of Political Economy*, 1991, *99* (3), 483–498.

Lokshin, Michael and Ruslan Yemtsov, “Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor?,” *The World Bank Economic Review*, 2005, *19* (2), 311–333.

Mazlumolhosseini, Ali, "Relationships Between Social and Economic Development and Access to Rural Roads in Developing Countries," *Transportation Research Record*, 1990, *1274*, 179–194.

Michaels, Guy, "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *The Review of Economics and Statistics*, 2008, *90* (4), 683–701.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 2014, *104* (2), 422–458.

Morten, Melanie and Jaqueline Oliveira, "The Effects of Roads on Trade and Migration," *American Economic Journal: Applied Economics*, forthcoming.

Mu, Ren and Dominique Van de Walle, "Rural Roads and Local Market Development in Vietnam," *The Journal of Development Studies*, 2011, *47* (5), 709–734.

Oberfield, Ezra, "Productivity and misallocation during a crisis: Evidence from the Chilean crisis of 1982," *Review of Economic Dynamics*, 2013, *16*, 100–119.

Owuor, G., M. Ngigi, A. S. Ouma, and E. A. Birachi, "Determinants of Rural Poverty in Africa: The Case of Small Holder Farmers in Kenya," *Journal of Applied Sciences*, 2007, *7* (17), 2539–2543.

Porter, Gina, "The impact of road construction on women's trade in rural Nigeria," *Journal of Transport Geography*, 1995, *3* (1), 3–14.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11*, 707–720.

_ **and** _ , "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–174.

Shamdasani, Yogita, "Rural road infrastructure & agricultural production: Evidence from India," *Journal of Development Economics*, 2021, *152*, 1–18.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2022, *225* (2), 175–199.

Vasconcelos, Rafael, "Misallocation in the Brazilian Manufacturing Sector," *Brazilian Review of Econometrics*, 2017, *37* (2), 191–232.

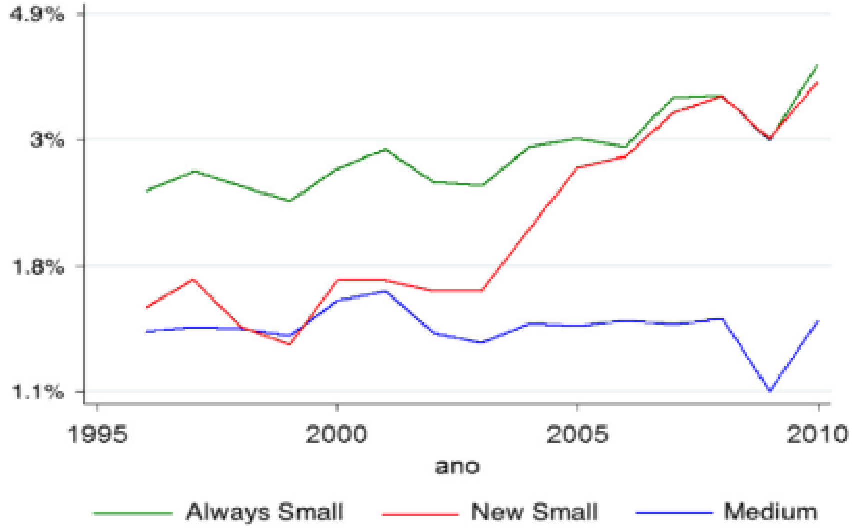
WB, "Safe, Clean and Affordable... Transport for Development," 2008. The World Bank Group's Transport Business Strategy for 2008-2012.

_, "World Bank Annual Report," 2022. available at: https://www.worldbank.org/en/about/annual-report.

Wu, Guiying Laura, "Capital Misallocation in China: Financial frictions or policy distortions?," *Journal of Development Economics*, 2018, *130*, 203–223.

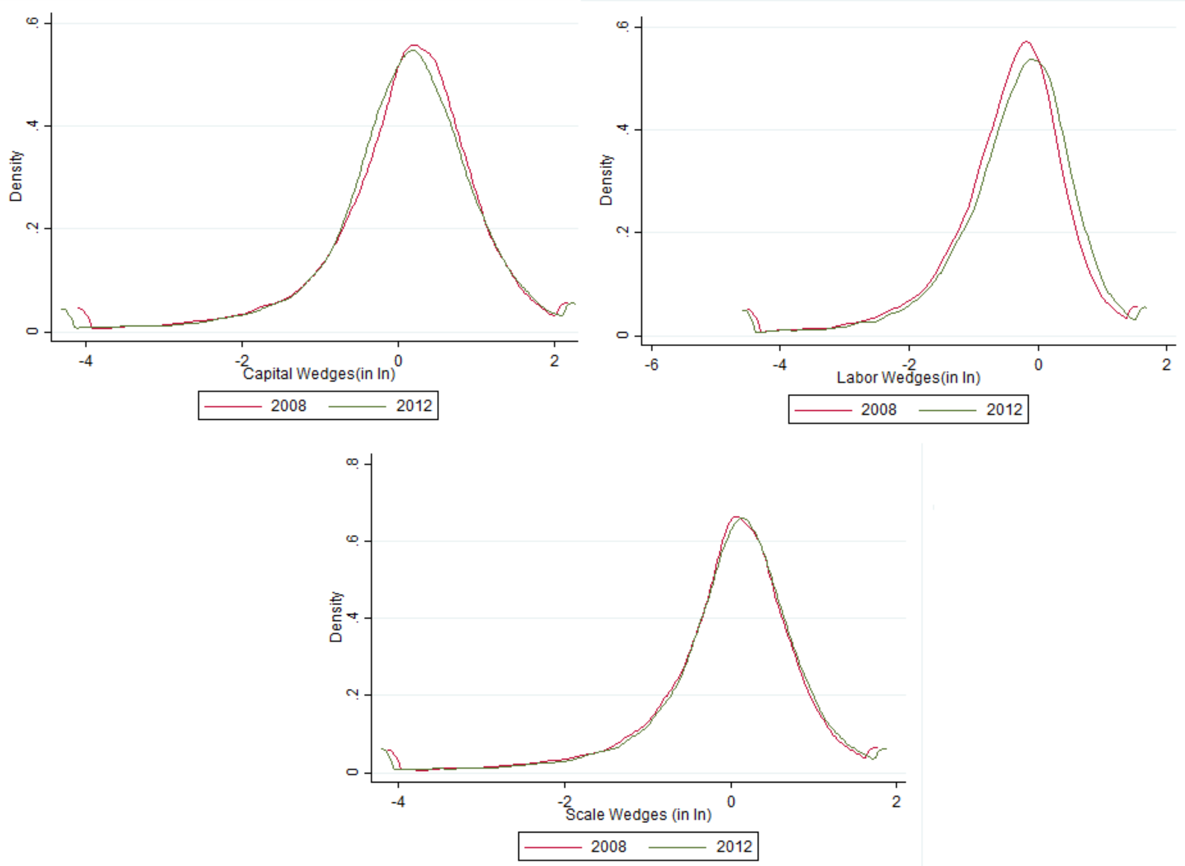
Appendix

Figure 4.1 log[investment/capital stock] new small vs. always small vs. always medium



Source: graph taken from Cavalcanti and Vaz (2017), page 24.

Figure 4.2 Capital, labor and scale wedges for the years 2008 and 2012



Source: Distribution of log capital, labor and scale wedges for the years 2008 and 2012, generated through the model in section 2.3.

Table 4.1 Treatment Effects on Capital, Labor and Scale Wedges ($\sigma = 1.5$)

	New Small x Always Small								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.176***	- 0.165***	- 0.160***	- 0.075	- 0.071	- 0.076	- 0.130**	- 0.121**	- 0.117**
	(0.061)	(0.061)	(0.059)	(0.050)	(0.048)	(0.047)	(0.057)	(0.056)	(0.055)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	45,464	45,460	45,460	45,464	45,460	45,460	45,464	45,460	45,460
R^2	0.428	0.441	0.446	0.528	0.539	0.539	0.437	0.449	0.453
	New Medium x Always Large								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.066*	- 0.069*	- 0.064	0.020	0.017	0.019	- 0.058	- 0.060	- 0.056
	(0.040)	(0.040)	(0.040)	(0.036)	(0.037)	(0.037)	(0.039)	(0.039)	(0.039)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	17,133	17,128	17,128	17,133	17,128	17,128	17,133	17,128	17,128
R^2	0.510	0.532	0.534	0.596	0.611	0.612	0.512	0.533	0.535

Notes: Results from estimations of New small vs. always small and new medium vs. always large. The first column presents results with firm id and year fixed effects. The second column adds industry fixed-effects. The third column add controls: number of employees and state fixed effects.

*p*0.10 p**0.05 p***0.01*

Table 4.2 Treatment Effects on Capital, Labor and Scale Wedges ($\sigma = 1.5$) (Event Study)

	New Small x Always Small					
	Capital Wedge		Labor Wedge		Scale Wedge	
5 years before	- 0.015	0.002	- 0.066	- 0.052	- 0.039	- 0.021
	(0.088)	(0.088)	(0.081)	(0.081)	(0.084)	(0.084)
4 years before	0.022	0.033	- 0.043	- 0.041	- 0.000	0.008
	(0.087)	(0.088)	(0.077)	(0.079)	(0.084)	(0.0853)
3 years before	0.036	0.049	0.009	0.022	0.027	0.039
	(0.101)	(0.102)	(0.098)	(0.099)	(0.098)	(0.099)
2 years before	0.019	0.024	0.008	0.020	0.012	0.017
	(0.065)	(0.064)	(0.060)	(0.060)	(0.0629)	(0.063)
1 year before	0	0	0	0	0	0
year of treatment	- 0.039	-0.040	0.006	0.006	- 0.020	- 0.021
	(0.078)	(0.078)	(0.080)	-0.020	(0.078)	(0.080)
1 year after	- 0.133*	- 0.099	- 0.105	- 0.094	-0.114	-0.086
	(0.078)	(0.077)	(0.075)	(0.074)	(0.077)	(0.076)
2 years after	-0.112	-0.0857	- 0.099	- 0.088	-0.091	-0.070
	(0.072)	(0.071)	(0.071)	(0.069)	(0.071)	(0.070)
3 years after	- 0.177**	-0.152*	- 0.098	- 0.096	- 0.145*	- 0.126
	(0.089)	(0.084)	(0.085)	(0.078)	(0.087)	(0.081)
4 years after	- 0.188*	- 0.149	- 0.096	- 0.080	- 0.143	- 0.111
	(0.103)	(0.099)	(0.096)	(0.090)	(0.099)	(0.094)
5 years after	- 0.279**	- 0.232**	- 0.149	- 0.133	- 0.222*	- 0.182*
	(0.122)	(0.116)	(0.105)	(0.098)	(0.115)	(0.110)
6 years after	- 0.264**	- 0.269**	- 0.117	- 0.129	- 0.203	- 0.205*
	(0.135)	(0.127)	(0.103)	(0.096)	(0.126)	(0.119)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	45,464	45,460	45,464	45,460	45,464	45,460
R^2	0.429	0.446	0.528	0.539	0.437	0.453

Notes: Results for estimations in the event study format for new small vs always small. The first column indicate results using firm id and year fixed effects. The second column indicate results adding controls: number of employees and state fixed effects.
*p*0.10 p**0.05 p***0.01*

Table 4.3 Treatment Effects on Capital, Labor and Scale Wedges ($\sigma = 1.5$) (Event Study)

	New Medium x Always Large					
	Capital Wedge		Labor Wedge		Scale Wedge	
4 years before	0.023	0.019	- 0.020	- 0.015	0.021	0.018
	((0.060))	(0.061)	(0.062)	(0.063)	(0.058)	(0.060)
3 years before	0.017	0.009	0.015	0.021	0.019	0.012
	(0.061)	(0.062)	(0.060)	(0.062)	(0.059)	(0.060)
2 years before	0.002	- 0.002	- 0.026	- 0.022	- 0.001	- 0.005
	(0.049)	(0.050)	(0.0524)	(0.053)	(0.049)	(0.049)
1 year before	0	0	0	0	0	0
Year of treat.	- 0.019	- 0.027	0.045	0.040	- 0.015	-0.022
	(0.041)	(0.042)	(0.040)	(0.040)	(0.040)	(0.041)
1 year after	- 0.046	- 0.052	- 0.003	- 0.000	- 0.042	- 0.047
	(0.039)	(0.041)	(0.038)	(0.039)	(0.039)	(0.040)
2 years after	- 0.109**	- 0.122**	- 0.000	- 0.007	- 0.097*	- 0.108**
	(0.055)	(0.055)	(0.054)	(0.055)	(0.053)	(0.053)
3 years after	- 0.048	- 0.048	0.049	0.050	- 0.035	- 0.035
	(0.052)	(0.052)	(0.049)	(0.0503)	(0.050)	(0.050)
4 years after	- 0.095	- 0.091	- 0.043	- 0.035	- 0.090	- 0.085
	(0.066)	(0.065)	(0.061)	(0.0605)	(0.065)	(0.063)
5 years after	- 0.017	- 0.003	0.020	0.032	- 0.012	0.002
	(0.065)	(0.066)	(0.059)	(0.061)	(0.063)	(0.063)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	17,133	17,128	17,133	17,128	17,133	17,128
R^2	0.511	0.534	0.596	0.612	0.512	0.535

Notes: Results for estimations in the event study format for new medium vs always large. The first column indicate results using firm id and year fixed effects. The second column indicate results adding controls: number of employees and state fixed effects.
*p*0.10 p**0.05 p***0.01*

Table 4.4 Treatment Effects on Capital, Labor and Scale Wedges (*sigma* = 5)

	New Small x Always Small								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.173***	- 0.163***	- 0.158***	- 0.077	- 0.075	- 0.078	- 0.120**	- 0.112**	- 0.109**
	(0.063)	(0.062)	(0.060)	(0.053)	(0.051)	(0.050)	(0.057)	(0.056)	(0.055)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	45,464	45,460	45,460	45,464	45,460	45,460	45,464	45,460	45,460
R^2	0.428	0.441	0.446	0.523	0.532	0.533	0.420	0.432	0.435
	New Medium x Always Large								
	Capital Wedge			Labor Wedge			Scale Wedge		
Eligible*Post	- 0.076*	- 0.080**	- 0.075*	0.080**	0.076*	0.078*	- 0.063*	- 0.067*	- 0.062
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.038)	(0.038)	(0.039)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**
Controls	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**YES**
Obs	17,133	17,128	17,128	17,133	17,128	17,128	17,133	17,128	17,128
R^2	0.507	0.528	0.530	0.563	0.580	0.581	0.502	0.523	0.524

Notes: Results from estimations of New small vs. always small and new medium vs. always large. The first column presents results with firm id and year fixed effects. The second column adds industry fixed-effects. The third column add controls: number of employees and state fixed effects.
*p*0.10 p**0.05 p***0.01*

Table 4.5 Treatment Effects on Capital, Labor and Scale Wedges ($\sigma = 5$) (Event Study)

	New Small x Always Small					
	Capital Wedge		Labor Wedge		Scale Wedge	
5 years before	- 0.016	0.007	- 0.061	- 0.065	- 0.039	- 0.017
	(0.088)	(0.087)	(0.091)	(0.092)	(0.084)	(0.084)
4 years before	0.013	0.032	- 0.008	- 0.025	-0.002	0.012
	(0.087)	(0.088)	(0.091)	(0.091)	(0.084)	(0.085)
3 years before	0.031	0.050	0.011	0.008	0.028	0.044
	(0.100)	(0.102)	(0.110)	(0.110)	(0.097)	(0.098)
2 years before	0.010	0.018	0.010	0.017	0.013	0.019
	(0.065)	(0.065)	(0.063)	(0.063)	(0.063)	(0.063)
1 year before	0	0	0	0	0	0
Year of treat.	- 0.048	- 0.047	0.015	0.011	- 0.011	- 0.012
	(0.080)	(0.080)	(0.086)	(0.087)	(0.078)	(0.078)
1 year after	- 0.134*	- 0.095	- 0.092	- 0.096	- 0.112	- 0.082
	(0.080)	(0.079)	(0.092)	(0.092)	(0.077)	(0.076)
2 years after	- 0.118	- 0.085	- 0.089	- 0.094	- 0.088	- 0.064
	(0.074)	(0.073)	(0.078)	(0.077)	(0.072)	(0.071)
3 years after	- 0.188**	- 0.156*	- 0.088	- 0.097	- 0.138	- 0.117
	(0.090)	(0.084)	(0.095)	(0.090)	(0.087)	(0.082)
4 years after	- 0.190*	- 0.148	- 0.073	- 0.068	- 0.133	- 0.101
	(0.106)	(0.101)	(0.102)	(0.098)	(0.101)	(0.096)
5 years after	- 0.267**	- 0.217*	- 0.119	- 0.122	- 0.200*	- 0.159
	(0.126)	(0.120)	(0.112)	(0.109)	(0.117)	(0.112)
6 years after	- 0.255*	- 0.259**	- 0.147	- 0.164	- 0.190	- 0.189
	(0.136)	(0.129)	(0.108)	(0.103)	(0.126)	(0.119)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	45,464	45,460	45,464	45,460	45,464	45,460
R^2	0.428	0.446	0.503	0.514	0.420	0.420

Notes: Results for estimations in the event study format for new small vs always small. The first column indicate results using firm id and year fixed effects. The second column indicate results adding controls: number of employees and state fixed effects.
*p*0.10 p**0.05 p***0.01*

Table 4.6 Treatment Effects on Capital, Labor and Scale Wedges ($\sigma = 3$) (Event Study)

	New Medium x Always Large					
	Capital Wedge		Labor Wedge		Scale Wedge	
4 years before	0.013	0.007	0.034	0.042	0.015	0.010
	(0.061)	(0.062)	(0.068)	(0.069)	(0.059	(0.060)
3 years before	0.019	0.011	0.024	0.027	0.016	0.009
	(0.061)	(0.062)	(0.068)	(0.069)	(0.059)	(0.0608)
2 years before	0.001	- 0.004	0.013	0.023	- 0.001	- 0.005
	(0.050)	(0.051)	(0.059)	(0.060)	(0.049)	(0.050)
1 year before	0	0	0	0	0	0
Year of treat.	- 0.021	- 0.032	0.114**	0.114**	- 0.014	- 0.024
	(0.042)	(0.042)	(0.054)	(0.054)	(0.040)	(0.041)
1 year after	- 0.059	- 0.069	0.098*	0.106**	- 0.049	- 0.056
	(0.041)	(0.042)	(0.051)	(0.052)	(0.039)	(0.041)
2 years after	- 0.111**	- 0.124**	0.062	0.054	- 0.099*	- 0.111**
	(0.055)	(0.055)	(0.061)	(0.062)	(0.054)	(0.054)
3 years after	- 0.065	- 0.068	0.142**	0.145**	- 0.044	- 0.045
	(0.052)	(0.053)	(0.057)	(0.058)	(0.050)	(0.050)
4 years after	- 0.130*	- 0.127*	0.080	0.084	- 0.114*	- 0.110*
	(0.068)	(0.066)	(0.067)	(0.067)	(0.065)	(0.064)
5 years after	- 0.023	- 0.009	0.083	0.095	- 0.019	- 0.004
	(0.066)	(0.066)	(0.066)	(0.067)	(0.062)	(0.063)
Unit FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year FE	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Industry FE	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Controls	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Obs	17,133	17,128	17,133	17,128	17,133	17,128
R^2	0.507	0.530	0.563	0.581	0.502	0.525

Notes: Results for estimations in the event study format for new medium vs always large. The first column indicate results using firm id and year fixed effects. The second column indicate results adding controls: number of employees and state fixed effects.
*p*0.10 p**0.05 p***0.01*

Table 4.7 Treatment Effects on Total Admissions, Resignations and Wages

	dep. variable: Admissions		*dep. variable: Resignations*		*dep. variable: Wage*	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	- 0.737	- 0.347	0.005	- 0.042	- 0.034	- 0.024
	(0.851)	(1.090)	(0.117)	(0.095)	(0.028)	(0.029)
2 years before	0.371	0.204	- 0.015	- 0.004	- 0.028	- 0.027
	(0.526)	(0.612)	(0.101)	(0.087)	(0.024)	(0.024)
1 year before	0	0	0	0	0	0
year of treatment	- 0.177	- 0.857	0.179**	0.185**	0.027*	0.027*
	(0.328)	(0.579)	(0.068)	(0.067)	(0.014)	(0.016)
1 year after	- 0.503	- 1.204	0.186*	0.169*	0.049*	0.056*
	(0.726)	(1.043)	(0.115)	(0.098)	(0.030)	(0.032)
2 years after	- 0.022	- 0.375	0.213	0.208	0.032	0.042
	(0.782)	(1.261)	(0.163)	(0.148)	(0.039)	(0.038)
3 years after	- 0.122	0.395	0.345*	0.353**	0.065	0.042
	(1.117)	(2.101)	(0.179)	(0.187)	(0.053)	(0.038)
4 years after	- 0.170	1.299	0.366	0.336	0.068	0.078
	(1.095)	(2.150)	(0.258)	(0.264)	(0.059)	(0.053)
Controls		✓		✓		✓
Obs (max/min)	830 / 172	830 / 172	830 / 172	830 / 172	893 / 317	893 / 317
Switchers (max/min)	203 / 105	203 / 105	203 / 105	203 / 105	203 / 174	203 / 174
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.8 Treatment Effects on Firm Size, Entry and Exit (Percent Change, 1 = 100%)

	dep. variable: Firm Size		*dep. variable: Entry*	*dep. variable: Exit*
	C&D'H	C&D'H Controls	C&D'H	C&D'H
3 years before	- 0.028	- 0.062	0.056	0.188
	(0.047)	(0.052)	(0.143)	(0.145)
2 years before	0.033	0.024	0.044	0.000
	(0.022)	(0.022)	(0.099)	(0.086)
1 year before	0	0	0	0
year of treatment	0.046*	0.040	0.173**	- 0.057
	(0.029)	(0.028)	(0.061)	(0.075)
1 year after	0.047*	0.028	0.202**	- 0.070
	(0.028)	(0.027)	(0.103)	(0.091)
2 years after	0.067	0.026	0.181	- 0.313**
	(0.063)	(0.067)	(0.139)	(0.135)
3 years after	0.069	- 0.003	0.158	- 0.524**
	(0.101)	(0.095)	(0.149)	(0.154)
4 years after	0.054	0.000	0.208*	- 0.591**
	(0.105)	(0.098)	(0.138)	(0.178)
Controls		✓		
Obs (max/min)	14,081 / 1,636	14,081 / 1,477	911 / 319	911 / 319
Switchers (max/min)	4,390 / 1,485	3,924 / 1,335	203 / 150	203 / 150
Bootstrap Reps.	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.9 Treatment Effects on Resignations by Sector

	dep. variable: Resignations					
	Agriculture		Manufacturing		Services	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	0.178	0.103	- 0.465	- 0.495	- 0.016	- 0.047
	(0.177)	(0.176)	(0.326)	(0.353)	(0.165)	(0.160)
2 years before	0.076	0.106	- 0.130	- 0.142	- 0.075	- 0.064
	(0.250)	(0.249)	(0.230)	(0.236)	(0.072)	(0.068)
1 year before	0	0	0	0	0	0
year of treatment	0.283*	0.263*	0.076	0.049	0.136	0.168**
	(0.158)	(0.153)	(0.224)	(0.187)	(0.080)	(0.081)
1 year after	0.130	0.125	0.987*	0.860*	0.084	0.089
	(0.170)	(0.168)	(0.502)	(0.494)	(0.100)	(0.111)
2 years after	0.272	0.443	0.444	0.434	0.220	0.204
	(0.280)	(0.264)	(0.601)	(0.631)	(0.137)	(0.154)
3 years after	0.315	0.551*	0.494	0.457	0.272	0.295*
	(0.304)	(0.298)	(0.959)	(1.179)	(0.176)	(0.170)
4 years after	0.428	0.689	0.035	0.152	0.142	0.256
	(0.549)	(0.486)	(0.946)	(1.137)	(0.229)	(0.212)
Controls		✓		✓		✓
Obs (max/min)	806 / 166	806 / 162	583 / 126	583 / 111	839 / 275	839 / 275
Switchers (max/min)	199 / 103	199 / 100	155 / 84	155 / 77	203 / 174	203 / 174
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.10 Treatment Effects on Wages by Sector

	dep. variable: Wages					
	Agriculture		Manufacturing		Services	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	0.031	0.038	0.085	0.123	- 0.019	- 0.015
	(0.027)	(0.027)	(0.092)	(0.086)	(0.038)	(0.038)
2 years before	- 0.011	- 0.017	- 0.001	0.006	- 0.034	- 0.032
	(0.020)	(0.020)	(0.062)	(0.062)	(0.025)	(0.027)
1 year before	0	0	0	0	0	0
year of treatment	- 0.005	- 0.002	0.393	0.334	0.017	0.021
	(0.019)	(0.019)	(0.269)	(0.257)	(0.017)	(0.017)
1 year after	- 0.017	- 0.009	0.181	0.043	0.026	0.041
	(0.021)	(0.022)	(0.116)	(0.137)	(0.028)	(0.033)
2 years after	- 0.082**	- 0.084**	- 0.028	- 0.192	0.029	0.037
	(0.027)	(0.032)	(0.132)	(0.136)	(0.035)	(0.038)
3 years after	- 0.059*	- 0.064*	- 0.115	- 0.354**	0.067	0.070
	(0.031)	(0.035)	(0.181)	(0.185)	(0.041)	(0.046)
4 years after	- 0.063	- 0.062	- 0.078	- 0.203	0.080	0.086
	(0.039)	(0.040)	(0.157)	(0.155)	(0.049)	(0.056)
Controls		✓		✓		✓
Obs (max/min)	792 / 114	792 / 114	583 / 126	583 / 111	893 / 317	893 / 317
Switchers (max/min)	195 / 59	195 / 59	155 / 84	155 / 77	203 / 174	203 / 174
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.11 Treatment Effects on Entry by Sector

	dep. variable: Entry		
	Agriculture	Manufacturing	Services
	C&D'H	C&D'H	C&D'H
3 years before	0.210	- 0.289	0.045
	(0.307)	(0.434)	(0.201)
2 years before	0.163	0.168	- 0.031
	(0.179)	(0.238)	(0.123)
1 year before	0	0	0
year of treatment	0.367**	0.148	0.041
	(0.112)	(0.210)	(0.080)
1 year after	0.264*	- 0.076	0.230
	(0.147)	(0.260)	(0.139)
2 years after	0.307	0.085	0.127
	(0.212)	(0.338)	(0.201)
3 years after	0.404	0.274	- 0.042
	(0.290)	(0.463)	(0.210)
4 years after	0.483*	0.008	0.060
	(0.280)	(0.550)	(0.213)
Controls			
Obs (max/min)	888 / 313	644 / 235	911 / 319
Switchers (max/min)	200 / 146	153 / 97	203 / 150
Bootstrap Reps.	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.12 Treatment Effects on Exit by Sector

	dep. variable: Exit		
	Agriculture	Manufacturing	Services
	C&D'H	C&D'H	C&D'H
3 years before	0.248	- 0.164	0.286
	(0.236)	(0.299)	(0.224)
2 years before	0.277*	- 0.164	- 0.207
	(0.156)	(0.299)	(0.151)
1 year before	0	0	0
year of treatment	0.166	- 0.289	- 0.300**
	(0.112)	(0.189)	(0.139)
1 year after	0.082	- 0.126	- 0.240
	(0.138)	(0.239)	(0.187)
2 years after	- 0.097	- 0.325	- 0.600**
	(0.174)	(0.299)	(0.260)
3 years after	- 0.189	- 0.325	- 1.070
	(0.162)	(0.377)	(0.350)
4 years after	- 0.572**	- 0.255	- 0.888**
	(0.269)	(0.397)	(0.443)
Controls			
Obs (max/min)	888 / 313	644 / 235	911 / 319
Switchers (max/min)	201 / 146	153 / 97	203 / 150
Bootstrap Reps.	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.13 Treatment Effects on Planted Area

	dep. variable: Planted Area					
	Total		Cash Crops		Food Crops	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	0.099	0.100	0.127	0.127	0.014	0.021
	(0.160)	(0.166)	(0.192)	(0.211)	(0.085)	(0.083)
2 years before	- 0.087	- 0.091	- 0.131	- 0.137	0.045	0.044
	(0.108)	(0.110)	(0.136)	(0.144)	(0.031)	(0.032)
1 year before	0	0	0	0	0	0
year of treatment	0.002	- 0.001	- 0.009	- 0.014	0.035	0.035
	(0.016)	(0.018)	(0.020)	(0.020)	(0.027)	(0.029)
1 year after	0.037	0.037	0.071	0.070	- 0.065	- 0.062
	(0.087)	(0.086)	(0.109)	(0.116)	(0.045)	(0.045)
2 years after	0.006	0.009	0.045	0.050	- 0.111**	- 0.115**
	(0.033)	(0.039)	(0.039)	(0.044)	(0.052)	(0.053)
3 years after	- 0.074	- 0.065	- 0.026	- 0.017	- 0.218**	- 0.210**
	(0.048)	(0.051)	(0.056)	(0.062)	(0.068)	(0.069)
4 years after	- 0.060	- 0.060	- 0.047	- 0.054	- 0.100	- 0.077
	(0.079)	(0.098)	(0.087)	(0.095)	(0.125)	(0.145)
Controls		✓		✓		✓
Obs (max/min)	829 / 171	806 / 162	829 / 171	806 / 162	829 / 171	806 / 162
Switchers (max/min)	202 / 104	199 / 100	202 / 104	199 / 100	202 / 104	199 / 104
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.14 Treatment Effects Yields

	dep. variable: Yields					
	Total		Cash Crops		Food Crops	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	- 0.104	- 0.115	- 0.073	- 0.084	- 0.254	- 0.262
	(0.088)	(0.081)	(0.049)	(0.047)	(0.357)	(0.364)
2 years before	- 0.062	- 0.062	- 0.088	- 0.085	0.065	0.050
	(0.047)	(0.050)	(0.056)	(0.055)	(0.045)	(0.043)
1 year before	0	0	0	0	0	0
year of treatment	- 0.015	- 0.018	- 0.001	- 0.004	- 0.081	- 0.084
	(0.024)	(0.024)	(0.023)	(0.024)	(0.085)	(0.084)
1 year after	0.027	0.025	0.054	0.049	- 0.103	- 0.090
	(0.044)	(0.041)	(0.039)	(0.037)	(0.131)	(0.124)
2 years after	0.224	0.228	0.353	0.348	- 0.403*	- 0.354**
	(0.219)	(0.251)	(0.259)	(0.270)	(0.204)	(0.165)
3 years after	0.330	0.326	0.485	0.463	- 0.421**	- 0.336**
	(0.314)	(0.359)	(0.381)	(0.394)	(0.204)	(0.161)
4 years after	0.834	0.881	0.933	0.979	0.354	0.406
	(0.703)	(0.826)	(0.887)	(0.936)	(0.467)	(0.559)
Controls		✓		✓		✓
Obs (max/min)	830 / 172	807 / 163	830 / 172	807 / 163	830 / 172	807 / 163
Switchers (max/min)	203 / 105	200 / 101	203 / 105	200 / 101	203 / 105	200 / 101
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.15 Treatment Effects on Production Value

	dep. variable: Production Value					
	Total		Cash Crops		Food Crops	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	- 0.290**	- 0.315**	- 0.283**	- 0.324**	- 0.307	- 0.294
	(0.141)	(0.143)	(0.105)	(0.118)	(0.304)	(0.309)
2 years before	- 0.095	- 0.092	- 0.108	- 0.091	- 0.064	- 0.092
	(0.089)	(0.085)	(0.117)	(0.120)	(0.052)	(0.056)
1 year before	0	0	0	0	0	0
year of treatment	- 0.031	- 0.040	- 0.004	- 0.013	- 0.095	- 0.094
	(0.080)	(0.079)	(0.100)	(0.110)	(0.074)	(0.080)
1 year after	- 0.096	- 0.102	- 0.052	- 0.073	- 0.202	- 0.172
	(0.079)	(0.077)	(0.073)	(0.078)	(0.203)	(0.182)
2 years after	0.071	0.107	0.232**	0.248**	- 0.319	- 0.231
	(0.097)	(0.092)	(0.077)	(0.097)	(0.277)	(0.222)
3 years after	0.144	0.158	0.288**	0.259**	- 0.204	- 0.087
	(0.095)	(0.100)	(0.119)	(0.110)	(0.264)	(0.205)
4 years after	0.503**	0.567**	0.618**	0.689**	0.225	0.273
	(0.226)	(0.281)	(0.261)	(0.325)	(0.295)	(0.344)
Controls		✓		✓		✓
Obs (max/min)	830 / 172	807 / 163	830 / 172	807 / 163	830 / 172	807 / 163
Switchers (max/min)	203 / 105	200 / 101	203 / 105	200 / 101	203 / 105	807 / 163
Bootstrap Reps.	500	500	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*

Table 4.16 Treatment Effects on Resignations in Agriculture by Initiative (levels)

	dep. variable: Resignations			
	Employer		Employee	
	C&D'H	C&D'H Controls	C&D'H	C&D'H Controls
3 years before	5.098	5.569	- 3.856	- 4.721
	(11.929)	(11.956)	(6.281)	(6.049)
2 years before	3.802	4.018	- 2.169	- 1.404
	(6.376)	(6.023)	(4.122)	(4.314)
1 year before	0	0	0	0
year of treatment	16.786	16.657	- 2.017	- 1.536
	(13.130)	(12.159)	(3.792)	(4.096)
1 year after	12.483	12.787	- 0.877	0.459
	(16.797)	(15.637)	(7.146)	(7.834)
2 years after	35.613*	35.854*	8.400	9.997
	(20.367)	(20.479)	(11.930)	(12.802)
3 years after	53.772*	53.631*	14.467	16.034
	(29.074)	(29.254)	(15.413)	(16.257)
4 years after	71.759**	72.074**	15.147	16.244
	(35.595)	(34.929)	(21.129)	(21.311)
Controls		✓		✓
Obs (max/min)	789 / 270	789 / 270	541 / 171	541 / 171
Switchers (max/min)	173 / 117	173 / 117	126 / 70	126 / 70
Bootstrap reps.	500	500	500	500

Notes: This table provides treatment effects by year. Effects should be interpreted as percent change relative to one year before treatment. Controls are workers characteristics: education, gender, age, work experience in current firm and weekly working hours. Estimates are computed for each yearly comparison of treated against controls. Observations are reported for the maximum and minimum numbers, and include both switchers (treated) and control groups. Switchers report the maximum and minimum number of treated units. Maximum numbers for both observations and switchers typically refer to the year of treatment, and minimum numbers are for the extremities, typically 3 years before or 4 years after. Larger standard errors reflect a smaller number of observations. Each estimation is done using 500 bootstrap repetitions. Standard errors are reported in brackets.
*p*0.10 p**0.05*