UNIVERSITAT DE BARCELONA

Essays on Misallocation of Resources: Evidence from Cuba

Ricardo González Aguila

PhD in Economics

Essays on Misallocation of Resources: Evidence from Cuba

Ricardo González Aguila

UNIVERSITAT DE BARCELONA

PhD in Economics

Thesis title:

Essays on Misallocation of Resources: Evidence from Cuba

PhD candidate:
Ricardo González Aguila

Advisor:
Sergi Basco

Date:
January 2025

LIBERTAS PERFVNDET · OMNIA LVCE
UNIVERSITAT DE BARCELONA

To Yaima and Martin.

Acknowledgements

Completing this work marks the end of a significant stage in my life. Reaching this point would not have been possible without the support of many.

On a personal level, I am deeply grateful to my supervisor, Sergi Basco, for his time, dedication, and kindness, for teaching me the challenging art of research, and for encouraging me when I needed it most. His professionalism is truly inspiring, and I could not have asked for a better supervisor.

At an institutional level, I will always be grateful to the Economics program at the University of Barcelona for welcoming me, for fostering a stimulating and collaborative learning environment, and for funding my studies. My sincere appreciation goes not only to the academic committee and its two coordinators, Elisabet Viladecans and Ció Patxot, but also to one of the program's pillars, Jordi Roca, whose constant support I truly appreciate.

I am also grateful to the Economic Theory Department, where over the years I have genuinely felt at home. Special thanks to those professors who contributed to my work through the internal seminar: Xavier Raurich, Marc Teignier, Rafael Serrano, Vahagn Jerbashian, Bruno Delalibera, Montserrat Vilalta, and Alessandro Di Nola. I sincerely appreciate your insightful recommendations and guidance.

My thesis has also greatly benefited from the support of many dear colleagues, whose insights have been invaluable. I am especially grateful to Yaima Doimeadiós, Leandro Zipitría, Francisco Borraz, Ileana Díaz, Carlos Lage, Ricardo Torres, Omar Everleny, and Sofía Albizu-Campos.

I would also like to extend my gratitude to many of my professors, whom I admire and respect, and whose influence is reflected, in one way or another, in my thesis: Vilma Hidalgo, Graciela Sanroman, Juan Triana, Fernando Borraz, Emmanuel Guerre, and Stepana Lazarova. Much of what I aspire to be comes from them.

Of course, I would like to extend a special thank you to my alma mater, the University of Havana—both the Faculty of Economics and the Center for the Study of the Cuban Economy (CEEC). These institutions have shaped me both personally and professionally, and I will always hold them close to my heart. To my colleagues at CEEC, thank you for being by my side throughout this journey. Moreover, I have reached this point thanks to many of my former students, who are now talented colleagues and friends: Daibel, Olivia, Leandro, Elisabet, and Ana Laura. Thank you, guys.

I am particularly grateful to the Department of Economics (Decon) at UdelaR in Uruguay for all the support I have received throughout the different stages of my academic journey, with special thanks to Rosario Domingo.

Without the emotional support of my friends, reaching this milestone would have been far more difficult. To my lifelong friends—Yordi, Wilfre, Jose, Roy, Katty, Ana, Ale, Silvia, María, Lily, Giulio, Jared, Daniel, Victor Rimbau, and especially Uri—thank you all for being part of my life. And to my newer friends, the amazing crew from the coolest PhD office at UB—Wilma, Mariona, Natasha, Hamza, Tahha, Carlos, Ruixi, Ha, and Edu—thank you all for being part of this journey.

Without the unwavering support of my family, reaching this point would have been simply impossible. Thank you, Mom, who has always unconditionally put me before herself. To my father, for the inspiration—I know how much you would have loved to be here, Dad (I'm sorry I didn't get here sooner). Thanks to my superhero, Rafa, for carrying my burdens, and to Angy, for her endless kindness. To María Elena and Doime, for their inspiring examples. To Yaima, for being by my side every step of the way—for bringing me this far and for everything we have built together, this thesis is also dedicated to you. Of course, thanks to Martin—*mi mapa, mi refugio y mi personita favorita. Gracias por hacer mi vida feliz.*

Abstract

This thesis addresses the issue of resource misallocation in transition economies, using Cuba as a case study. Although the country undertook some market-based reforms in the mid-1990s, it was not until 2011 that a more significant process of economic opening began, marked by the gradual expansion of the private sector. However, despite the undeniable changes compared to the past, the 2011 reform was far less comprehensive than those implemented in China and Vietnam—a factor that contributed to somewhat disappointing results after a decade.

My work examines several unresolved issues that have ultimately hindered a more significant reduction in misallocation, which, to some extent, helps explain the limited effectiveness of the reform efforts. The study identifies general distortions— the root causes of misallocation (Chapter 2)— as well as more specific causes related to the foreign exchange market (Chapter 3) and the labor market (Chapter 4). Each chapter provides a conceptual, theoretical, and empirical discussion of these issues within the Cuban context.

Chapter 2 builds on the seminal work of Hsieh and Klenow (2009) to measure the welfare costs of misallocation driven by the effects of heterogeneous distortions. The model is applied to Cuban manufacturing industry data from 2004 to 2017. The findings indicate that if all distortions were hypothetically removed, aggregate TFP would increase by 45% to 75% over the study period. This underscores the existing reserves that could be leveraged through structural policies to induce productivity gains. Furthermore, the chapter identifies potential mechanisms that could act as drivers for such improvements.

Chapter 3, on the other hand, examines distortions in the foreign exchange market. For years, Cuba maintained a highly peculiar exchange rate system, characterized by multiple exchange rates, monetary segmentation, and an overvalued official currency. The latter created a system of currency transfers (subsidies and taxes), which became the root of multiple efficiency problems, with a particular impact on the competitiveness of the export sector.

In this chapter, the model of Hsieh and Klenow (2009) is extended to incorporate the exchange rate. The extended model is applied to Cuba using 2015 data, with a persistent exchange rate gap of 2,400%. The findings reveal

a greater dispersion of TFPR when the exchange rate gap is accounted for, and TFP gains that are 50 percentage points higher compared to the baseline model (excluding the gap). These results suggest that exchange rate policy is a key factor for economic renewal.

Finally, Chapter 4 investigates the efficiency implications of redundant employment, a persistent issue in the Cuban labor market. State-owned enterprises (SOEs) do not have control over hiring and dismissing their workers—an institutional constraint with significant implications for SOE' profitability, competitiveness, and short-term financial health.

The chapter employs a two-stage optimal choice model, applied to data from 2009 to 2013. The results show that in 2009, 93% of manufacturing SOEs in Cuba were affected by this issue, with an estimated 66% of state employment in this sector classified as redundant. However, by 2013, a significant reduction of 15 percentage points in the redundancy rate was observed. This decline is attributed to a combination of rising prices, increased availability of raw materials, productivity growth, and improved financial performance during the period. The results suggest that restoring firms' autonomy to determine their workforce size could serve as a stimulus for productivity. However, they also indicate that reforms in other areas, such as pricing policies, could help mitigate the consequences of redundant employment.

Keywords: *Total Factor Productivity, Real Exchange Rate Misalignment, Labor Redundancy, Manufacturing Industry*

Contents

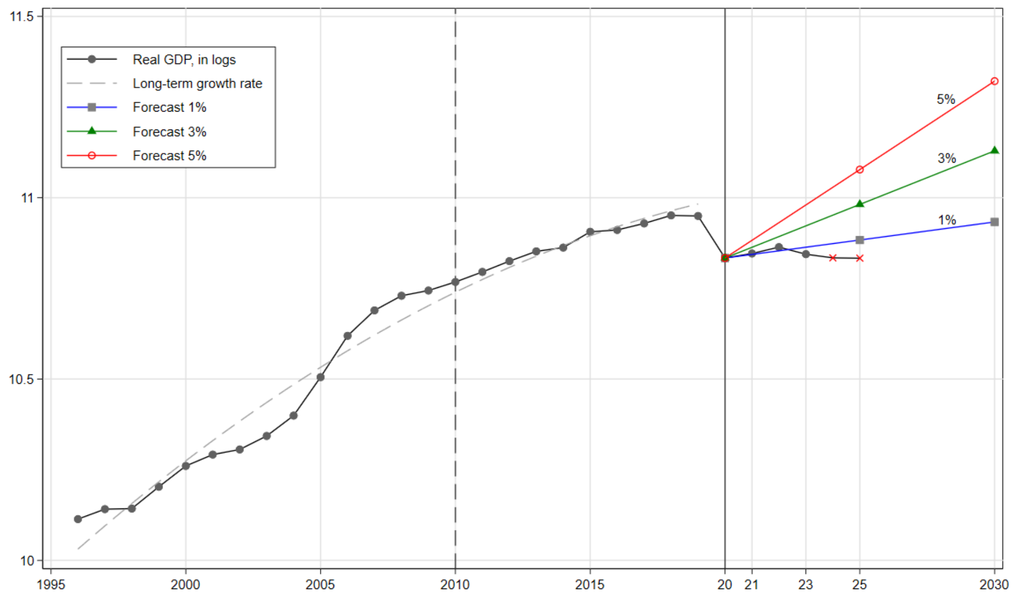
1 Introduction

Resource misallocation has become a central theory in explaining the significant differences in per capita income between countries (Jones, 2016). As emphasized by the extensive literature on the subject, developing economies not only lag behind advanced economies in terms of technological progress but also suffer from far more pronounced inefficiencies in the allocation of their scarce resources. These inefficiencies often arise from policies and/or institutions that distort relative prices in a heterogeneous manner (Restuccia and Rogerson, 2008, 2017). Consequently, they generate aggregate welfare losses, with observed output consistently falling short of the potential output of the economy.

This thesis addresses the issue of misallocation in transition economies. A transition economy is one that undergoes structural transformations, shifting its foundations from a centrally planned system to a market-based economy (Roland, 2000). Misallocation is an inherent feature of these economies, arising from the —often prolonged— coexistence of inherited centrally planned institutions and nascent, underdeveloped market institutions. The resulting distortions emerge within an institutional framework that retains many features of the previous system, such as the dominant role of the state, public ownership, and systematic interventions in goods and factor markets.

In this study, I examine the particular case of Cuba. The Cuban experience represents a late transition, one that has received considerably less attention in the literature compared to other cases, such as those of the former Soviet Union, China, or Vietnam. Although the country implemented some market-oriented reforms in the mid-1990s (CEPAL, 1999), a more substantial economic opening process did not begin until 2011, marked by the gradual expansion of the private sector. This expansion is reflected in employment trends; for instance, the share of the total labor force employed in the non-state sector increased from 18.4% in 2010 to 37.6% in 2023 (ONEI, 2023). While this figure is not particularly high, it represents a significant shift compared to the past.

However, the results achieved after more than a decade have been quite disappointing. Figure 1.1 illustrates this, highlighting three key points. First, long-term growth slowed over the past decade, reflecting persistent structural weaknesses. While GDP grew at an average annual rate of 2.8% in the first five-year period, it dropped to just 1.1% in the second ONEI (2022a), despite a series of implemented reforms. Second, the economy has remained stagnant since the onset of the COVID-19 crisis in 2020, when GDP plummeted by 10.9% ONEI (2022a). Third, the outlook for recovery appears increasingly bleak, given the sluggish growth trajectory—hovering around 1%—and the current geopolitical conditions, which are anything but favorable for the country.



Note: Own calculations based on ONEI, 2022a, Statistical Series, Table 5.4 . × projections by CEPAL (2024)

Figure 1.1: Real GDP, Cuba, 1995–2030.

There is a certain consensus that these results stem from a multiplicity of causes: (i) incomplete—and in some cases failed—reforms, (ii) U.S. sanctions and their geopolitical implications, and (iii) the recurrent occurrence of adverse external shocks, especially over the past five years. These three factors have contributed to a significant deterioration in the country's already low living standards.

Within the first group of causes, some studies suggest that Cuba's approach to reform has been less comprehensive than those of China and Vietnam. Alonso and Vidal (2020) argue that conflicts of interest and divergent objectives have led Cuban authorities to adopt a cautious stance and remain reluctant to implement bold measures. Policymakers have even stalled or reversed previously achieved pro-reform political consensuses—effectively "trapping" the reform and preventing the acceleration of change. As a result, many inherited problems have remained unresolved, ultimately leading to stagnation in productivity, as documented by Doimeadios (2012); Palacios (2013); Vidal (2017).

This thesis is based on the hypothesis that at least some of the less-than-satisfactory outcomes of the Cuban reform process can be attributed to resource misallocation. Misallocation in the Cuban context is a pervasive, cross-cutting issue, with distortions present across all markets. Many of the problems inherited from the previous system remain highly prevalent: (i) soft budget constraints, (ii) political burdens, (iii) price regulations, and (iv) the lack of SOE autonomy -institutional characteristics that are deeply embedded in the still dominant state enterprise system (Diaz et al., 2024). Moreover, considering the numerous imbalances and constraints the country currently faces, implementing structural policies to reduce misallocation is one of the few viable drivers of economic recovery in the short to medium term.

The policies implemented during the reform process did not aim to create a uniform set of "rules of the game" for all market participants. Instead, they tended to favor certain groups over others. This heterogeneity is evident at various levels—for instance, between state and private agents, politically connected and non-connected agents, large and small firms, and exporters and non-exporters. As a result, a heterogeneous system of barriers and benefits has been reinforced, creating advantages for some agents and disadvantages for others. Over time, these dynamics have strengthened rent-seeking behavior and entrenched extractive institutions. Understanding how such heterogeneous systems of advantages and disadvantages are created—and working towards leveling the playing field—becomes a key objective in reducing misallocation.

In line with the above, this thesis aims to quantitatively assess the costs of resource misallocation in Cuba (Chapter 2) and explores its links to the foreign exchange market (Chapter 3) as well as the labor market (Chapter 4).

More specifically, Chapter 2 presents an estimation of the welfare cost of misallocation in terms of total factor productivity (TFP), based on the seminal work of Hsieh and Klenow (2009). In the latter, the effects of misallocation are measured through gaps in the marginal revenue product of factors (MRP) across firms. One key advantage of this approach is that it does not require identifying a direct source of misallocation; instead, it indirectly captures overall distortions, which translate into wedges on MRP (Restuccia and Rogerson, 2017). The model is applied to Cuban manufacturing industry data for the period 2004–2017. Although Hsieh and Klenow (2009) is a widely used analytical framework, no such study had been conducted for the Cuban case until now, making this one of the key contributions of my research.

The findings indicate that fully eliminating heterogeneous distortions in Cuba would result in a substantial increase in aggregate TFP, with estimated gains ranging from 45% to 75% over the study period. This highlights the significant untapped potential that still exists to boost economic growth in the country. The results also suggest that if distortions were fully removed, more than half of state-owned enterprises (SOEs) would scale down their production by up to 50%. This finding supports anecdotal and suggestive evidence that frequently points to the 'gigantism' of Cuban SOEs and its negative impact on productivity.

Additionally, although misallocation remains high, a decline is observed after 2011, coinciding with the implementation of structural reforms. This reduction may explain approximately one-third of the observed growth in manufacturing real value added during the period, once again highlighting the importance of structural reforms in accelerating Cuba's economic growth. The study also explores potential mechanisms through which these reforms may operate. In particular, it finds that although foreign firms are about 35% more productive in terms of output than SOEs, they face greater implicit barriers that hinder their growth and expansion. This provides quantitative evidence supporting the widely held belief in Cuba that reducing barriers for non-state economic actors is key to fostering short-term growth.

Chapter 3, on the other hand, examines the relationship between real exchange rate misalignment and resource misallocation. Specifically, it explores how exchange rate overvaluation affects long-term output through the misallocation channel. I propose a new mechanism that operates through exchange rate transfers (subsidies and taxes). To this end, I extend the model of Hsieh and Klenow (2009) by incorporating the exchange rate. The findings indicate that failing to account for exchange rate misalignment results in an underestimation of the true cost of misallocation. Consequently, the results suggest that the misallocation literature should be revisited in economies where exchange rate misalignment is significant.

The extended model is then applied to Cuba's 2015 data, which represents an extreme case of misalignment due to the presence of multiple exchange rates, with a remarkable exchange rate gap of 2,400%. In the specific case of Cuba, a greater dispersion of marginal revenue products (MRPs) is observed when incorporating the influence of the exchange rate gap, and it is documented that TFP gains increase from 54% in the baseline model (without the gap) to 106% in the counterfactual model (with the gap). This striking difference of more than 50 percentage points is attributed to the effect of the 2400% spread. The results suggest that reforms in the foreign exchange market are a key factor in fostering economic growth.

Finally, Chapter 4 investigates the effects of labor redundancy on resource allocation, a common issue in the labor markets of many transition economies. Governments often have both the incentives and the means to dictate the minimum number of workers that each state entity must hire, driven, for instance, by the aim of keeping unemployment levels low. The inability of SOEs to autonomously determine their workforce size and its composition results in redundant employment—an example of a broader phenomenon referred to in the transition literature as 'policy burden' (Lin et al., 1998; Lin and Tan, 1999; Lixing, 2008; Liao et al., 2023). In a nutshell, SOEs are compelled to comply with political mandates, which impose additional costs, leading to declines in profitability and competitiveness.

To quantify the extent of labor redundancy and its relationship with other economic and financial variables, I employ a two-stage decision model, which is subsequently estimated using firm-level manufacturing data from Cuba between 2009 and 2013. The results show that, in 2009, 93% of all manufacturing SOEs in Cuba had excess workers, with an estimated 66% of total state employment classified as redundant. However, by 2013, the redundancy rate had decreased by 15 percentage points, a decline attributed to a combination of rising prices, improved availability of raw materials, productivity gains, and better financial performance during the period. These findings highlight that, although labor redundancy remains a persistent challenge for the Cuban economy, market-oriented reforms that boost productivity and strengthen financial institutions can help mitigate its effects.

To conclude, the findings of this thesis suggest that Cuba still has untapped efficiency reserves. Reducing misallocation represents a key driver of economic growth, particularly in the short term. Although the study covers data up to 2017, many of the issues discussed are more relevant today than ever. Institutional and incentive-related challenges persist, along with distortions in the foreign exchange and labor markets, as well as in many other areas that urgently require structural reform. Given the current unfavorable international geopolitical climate, mounting macroeconomic imbalances, and the growth trap in which the Cuban economy remains stuck, the best—perhaps the only—strategy for the current administration is to pursue reforms based on a more efficient use of resources. Reducing misallocation is not an easy path, but it is an unavoidable one if the goal is to improve the living standards of Cuban society.

2 Misallocation of Resources in a Transition Economy

2.1 Introduction

Research on economic development consistently highlights that disparities in per capita income between rich and poor countries are only partially explained by differences in capital, labor, or natural resources, and are largely driven by productivity gaps (Hall and Jones, 1999; Caselli, 2005). One increasingly relevant explanation for cross-country productivity differences is resource misallocation. This framework underscores that policies and institutions which distort the relative prices faced by heterogeneous producers hinder the allocation of resources to their most efficient use, thereby reducing aggregate productivity relative to its potential level (Restuccia and Rogerson, 2008, 2017).

While misallocation is an issue affecting both developed and developing economies, some quantitative evidence suggests its effects are more severe and detrimental in developing economies (Hsieh and Klenow, 2009), where institutional capacities are relatively weaker. Likewise, it represents a heavier burden for transition economies. I define a transition economy as one undergoing a profound structural transformation, where its foundations shift from a centrally planned system to a market-based economy (Roland, 2000). In such economies, the rules and policies inherited from the previous institutional regime play a central role. Even with private property and functioning markets, the state often maintains a dominant interventionist role, market institutions remain weak, and public ownership, along with misaligned incentives, continues to prevail. These distortions fuel extractive or rent-seeking behavior, leading to more inefficient resource allocation, stalled investment and technological progress, and stagnant productivity.

In this paper, I aim to provide some quantitative evidence on the potential cost of resource misallocation in Cuba. The Cuban case represents a late transition experience that is underexplored in the literature. Unlike China's reforms, which began in 1978, Vietnam's in 1986, or those of post-Soviet countries in 1991, Cuba's market-based reform process did not begin until 2011.[1] The resulting economic model exhibits several idiosyncratic characteristics that make it particularly interesting for misallocation theory. For example, while (small) private property and markets were more significantly promoted with the 2011 reforms, the government never encouraged the unification of private and state markets (not even gradually, as in the cases of China and Vietnam). This resulted in a permanent segmentation of relative prices between state and private agents, leading to misallocation of resources and other supply-side issues. Much of the widespread scarcity (exacerbated by the reforms) and the expansion of parallel markets for factors and goods can be traced back to this issue.

Another institutional peculiarity of the Cuban case relates to its monetary environment. During the period covered by the study, firms operated under a monetary institutional arrangement characterized by the coexistence of two domestic currencies,[2] the use of exchange controls, as well as multiple exchange rates with a remarkable 2,400% spread between official and parallel rates. The price segmentation between the state and private sectors forced a situation where, for the same international transaction, state-owned enterprises (SOEs) used the overvalued official exchange rate of 1 Cuban peso per US dollar, while private agents relied on the depreciated parallel rate of 24 Cuban pesos per US dollar. The heterogeneity in the relative exchange rate faced by firms represents a salient source of misallocation in Cuba.

Furthermore, Cuba's status as a sanctioned country is another idiosyncratic factor at the root of misallocation, a condition that acts as a barrier to its integration into the global economy. Sanctions often push the economy toward a form of autarky—frequently justified on national security grounds—forcing the country to allocate scarce resources to the production of goods and services that would not otherwise be produced in a freely trading economy. Moreover, the impact of sanctions is not uniform across firms. First, they disproportionately affect SOEs compared to private ones. Second, even among SOEs, there is significant heterogeneity, with some being blacklisted in ways that effectively block any engagement with international counterparts.

[1] I recognize the mid-1990s reforms as an important step toward opening up the economy following the 1991–1993 crisis. However, their scope was far more limited compared to the 2011 reforms. Not only were the 1990s reforms largely undone in the 2000s, but their quantitative impact was also much less significant. For instance, by 2009, 82% of employment was still concentrated in the state sector, and private workers—excluding cooperatives—numbered 590,000 (**?**). A decade later, state employment accounted for 66%, and private workers exceeded one million (idem).

[2] The Cuban Peso (CUP) and the Cuban Convertible Peso (CUC).

The specific goal of this paper is to provide a first quantification of the overall effect of misallocation on productivity (TFP) in Cuba. I adopt the indirect approach introduced in the seminal work of Hsieh and Klenow (2009). The main advantage of this approach is that it allows to quantify the extent of misallocation without explicitly specifying any underlying cause (Restuccia and Rogerson, 2017).[3] In particular, in this paper, I follow Dias et al., 2016, which incorporate intermediate inputs into the Hsieh-Klenow framework. I choose to incorporate intermediate input because material shortages (such as raw materials, fuels and energy) constitute the most critical constraints facing Cuban companies, often resulting in significant idle capacity.

I apply the model to data from the Cuban manufacturing industry for the period 2004–2017. The dataset primarily consists of data from SOEs, along with a small sample of approximately 30 foreign companies (joint ventures) for the sub-period 2005–2006. Unfortunately, data from the private sector, such as that from manufacturing cooperatives, is not available in these years. The data set also covers a period after 2011, coinciding with the implementation of structural reforms. I derive a measure of TFPR for each year, which, within the HK methodological framework, represents a weighted measure of marginal product gaps across firms (commonly referred to as wedges). I use TFPR for two purposes: first, to simulate TFP gains under hypothetical reductions in misallocation; and second, to correlate it with various policies to examine whether these policies increased or reduced TFPR variation.

I find that reducing resource misallocation would lead to a significant increase in aggregate total factor productivity (TFP). Quantitatively, if resources were allocated efficiently (i.e., marginal products across firms within the same sector were equalized), it would result in TFP increases of 45% to 75% over the study period. These estimates likely represent a lower bound, as they exclude the private sector, which I suspect would amplify TFPR dispersion even further. Findings also suggest that more than half of the firms would scale down their production by up to 50% if misallocation were eliminated. This means that many firms are currently overproducing, aligning with anecdotal evidence often pointing to the "gigantism" of Cuban state-owned enterprises. I also document a decline in misallocation over time, consistent with the structural reforms introduced after 2011. This reduction may account for roughly one-third of the observed growth in manufacturing real value added during the period. The findings also reveal that while foreign firms are about 35% more productive in quantity than state-owned enterprises, they face greater implicit

[3] I leave for future research to identify how different particularities of the Cuban case contribute to the overall distortion in the allocation of resources.

constraints that hinder their growth and expansion. This fact is in line with their limited presence in the Cuban economy and the extensive literature that consistently underscores the institutional restrictions they face (Triana, 2021).

The study of TFP and its explanatory factors in Cuba is not new. Empirical research based on growth accounting has consistently shown that long-term economic growth has declined significantly, largely due to insufficient factor accumulation and persistent stagnation in TFP (Doimeadios, 2012; Palacios, 2013; Vidal, 2017). Productivity stagnation has been—directly or indirectly—attributed to various causes, including the degree of centralization in state resource allocation (Doimeadios, 2012; Palacios, 2013); poor export performance (Quiñones and Torres, 2015; Romero, 2015; Garcia and Sovilla, 2017); growth-reducing structural change (Torres, 2013); high levels of technical inefficiency (Doimeadios and Sanchez, 2015); distortions in human capital (Cribeiro, 2012); monetary duality and other macro-level institutional problems (**?**Hidalgo and Cribeiro, 2015; Vidal, 2017); barriers to innovation (Diaz, 2018); and (political) conflicts of interest that hinder deeper reforms (Alonso and Vidal, 2020).

Nevertheless, while misallocation is a concept implicitly present in the spirit of many of these studies—and some even provide invaluable examples and suggestive evidence of its role—no research to date has placed misallocation at the core of its analysis. Consequently, there is a notable lack of quantitative evidence on the welfare costs that misallocation imposes on society. This gap is particularly striking for a centralized economy, characterized by numerous distortions and inefficiencies. This paper aims to contribute to filling this gap in the literature by providing the first quantitative analysis of misallocation in Cuba. While comparable studies have been conducted in other emerging and transition economies (e.g., China and Vietnam), exploring the case of Cuba provides valuable evidence on an increasingly relevant topic.

Compared to measurements for China (Hsieh and Klenow, 2009) and Vietnam (Nguyen and Nguyen, 2020), I document lower levels of misallocation in my case of study—a result that deeply surprised me, as it runs counter to my expectations. Although this finding is not entirely comparable due to methodological differences between studies—for instance, differences in approach (value-added vs. gross output), productive units (plants vs. establishments), or agent types (private vs. state-owned)—my study opens the door to new research questions to further investigate this result. I leave, for example, open the question of to what extent the closed-economy HK model can capture the welfare costs caused by exchange rate misalignment, which is particularly relevant to the Cuban case but typically fall within the scope of open-economy research. Therefore, what I suggest is that examining Cuba is not only important because it constitutes an "interesting case," but also because it can pave the way for new lines of research within the misallocation literature.

The paper is organized as follows. Section 2.2 examines the various sources of misallocation in Cuba, with an emphasis on its structural drivers. Section 4.3 introduces a simple model to quantitatively assess its effects on TFP. Section 2.4 describes the data, while Sections 2.5 and 2.6 analyze the results. Finally, Section 2.7 summarizes the main findings and offers conclusions.

2.2 Potential Sources of Misallocation in Cuba

In this section, I provide an overview of the main sources of misallocation in Cuba. I disentangle its multiple interconnected causes and map the channels that drive this phenomenon in my case study.

Institutional problems are at the core of misallocation in Cuba. Despite the gradual implementation of market reforms since 2011 and the subsequent transition toward a mixed economy, the country's economic foundations remain deeply rooted in the principles of a planned economy. The state continues to dominate by substituting market mechanisms and centrally allocating resources. Although many markets have emerged, widespread government intervention and pervasive institutional distortions undermine their proper functioning, leading to significant resource misallocation. These distortions are systemic, affecting all markets, including those for final goods and services, factor markets, the monetary market, and the foreign exchange market. As a result, they not only weaken price signals but also generate numerous incentive-related challenges within firms, further exacerbating inefficiencies.

Stemming from this general idea, Figure 2.1 separately identifies three channels that underlie the root of the problem. These channels are simply different expressions of the same underlying issue described above.

The first is government intervention, which in turn, can be divided into two types: allocative and non-allocative interventions. Regarding the former, the Cuban government systematically intervenes by directly allocating scarce productive inputs to selected "strategic" firms or industries. The selection criteria are discretionary and often driven by social and political motivations. Through the economic plan, essential resources (e.g., fuel, energy, and basic raw materials), are distributed among state companies, and investments are executed according to the plan's directives. The plan also influences the allocation of workers to companies, which often operate with limited autonomy in hiring

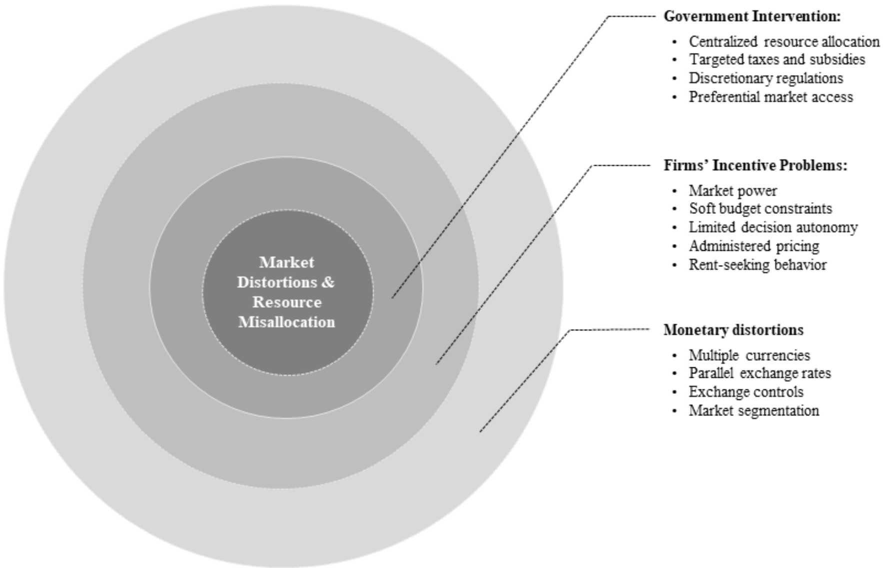


Figure 2.1: Three Channels of Resources Misallocation in Cuba

and dismissal decisions. This imposes policy burdens on firms—a typical issue in transition economies (see, e.g., Lin and Tan (1999); Lin and Li (2008)). Although the plan internalizes relative profitability information from markets, extra-economic criteria play a substantial role in allocations, often incorporating factors related to political priorities, social justice, or developmental goals.

Another way the government intervenes is through the discretionary application of taxes and subsidies on specific producers . An example of this is the policy of public transfers to state-owned enterprises (SOEs). The resources allocated to this policy appear directly in the state budget as a current expenditure item. These transfers have been notably high in Cuba, ranging from 10% to 25% of current GDP since the mid-1990s (González, 2021). The subsidies to state-owned enterprises (SOEs) are justified (not solely, but primarily) by price regulation, a widespread practice in Cuba.[4]

A noteworthy aspect in this regard is the heterogeneity in subsidy allocation among otherwise similar producers within the same industry. Although legislation exists to regulate and standardize the allocation of subsidies (see, for example, MFP (2013a)), the rate at which the same product is subsidized varies significantly from one enterprise to another. Ultimately, these rate dif-

[4]The idea is similar to that of politically mandated prices in markets economies, where the government regulates a monopoly's price by setting it below the average cost. This forces the monopoly to operate at a loss unless the government provides subsidies.

ferences reflect cost disparities among SOEs, as the subsidy is determined not by the (average) market costs but by each enterprise's individual costs. Instead of subsidizing the difference between the market price of the product and its regulated price, the government compensates for the gap between each firm's individual output price and the regulated price. In doing so, it not only covers the cost of the policy but also subsidizes the differences in relative costs across firms. In other words, it channels funds to enterprises with lower relative productivity/profitability.

On the other hand, regarding non-allocative interventions, a typical example is the issuance of specific regulations for certain business groups or, alternatively, special legal provisions that exempt them from compliance. These practices create favorable (or at least preferential) conditions for beneficiary firms, allowing them to access resources at lower costs than their competitors. Additionally, certain (foreign) companies are occasionally granted special licenses that restrict market competition, yielding extraordinary benefits. In sum, all these discretionary practices contribute to resource misallocation, ultimately hindering aggregate productivity growth.

The second channel identified in Figure 2.1 focuses on the horizontal institutional features inherited from centralized planning, which reduce the flexibility of the business environment and undermine the efficient performance of companies. In other words, resource misallocation also arises from the "rules of the game" set by the government, which shape economic incentives. Market power is a typical example of this dynamic in Cuba. Part of the monopoly power enjoyed by some SOEs stems from entry barriers legally imposed by the government. Although these barriers were significantly reduced with the 2011 reforms and, more notably, in 2021, when small private investments were granted legal recognition, they still persist in some sectors considered strategic. This deters more productive firms from entering and limits the competition faced by incumbent SOEs, primarily to imports and, in some cases, foreign companies authorized to invest.

Another example of misallocation caused by incentive problems –rooted in the institutional features inherited from centralized planning– is the phenomenon of soft budget constraints (SBCs) (Kornai, 1980; Kornai et al., 2003). SBCs represent a syndrome affecting firms in such economic regimes, where the state adopts a paternalistic approach, preventing bankruptcy through various instruments, even when firms incur chronic losses. This implies the existence of a third party (agent) willing to cover the firms' financial deficits, either fully or partially, thereby freeing them from the obligation to reduce or cease operations if the deficit persists (Kornai et al., 2003, 1096–98). SBCs create an incentive problem for firms, as they anticipate being refinanced.

The presence of the SBC syndrome in Cuba hardly requires extensive validation (see Kunzmann (2021); Gonzalez and Torres (2024) for a review on this topic). The systematic existence of SOEs operating at a loss is a common occurrence. In January 2022, the Cuban government revealed that 457 of these enterprises—about 27% of the total—were running at a loss (Tamayo, 2022). While these high numbers may partially reflect the severe economic crisis Cuba has faced since 2019, they cannot be entirely attributed to it. There is additional evidence pointing in the same direction in earlier years (see, for instance, Castro, 2015 and Kunzmann, 2021, 190).

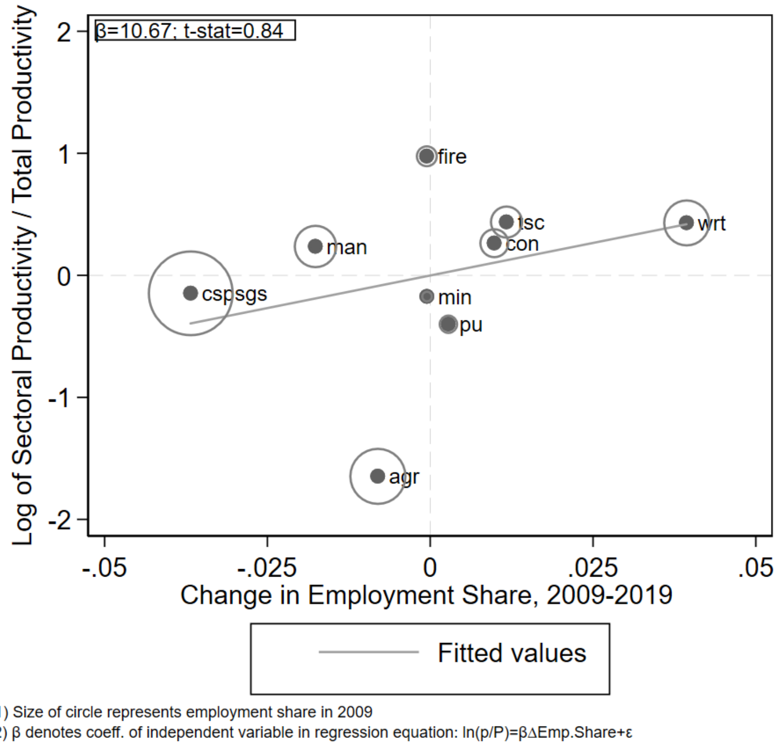
To prevent these enterprises from collapsing, state financial institutions, the public budget, and other indirect mechanisms are employed (some of them are described in Gonzalez and Torres, 2024). Regardless of the specific form these transfers take, the conclusion remains the same: resources with a high opportunity cost are channeled into unviable enterprises.

Finally, the last channel identified in Figure 2.1 relates to the monetary system. For over two decades, Cuba operated a dual monetary system grounded in two domestic currencies: the Cuban peso (CUP) and the convertible Cuban peso (CUC), with multiple exchange rates. Households (and later the private sector) operated in their international transactions with a fixed parallel exchange rate of 24 CUP per USD,[5] while SOEs and foreign companies used the official exchange rate of 1 CUP per USD. The segmentation of the exchange rate between households and firms was enforced by legally restricting companies from accessing the parallel market, and by the limited interaction between private and state-owned firms.

The consequences of this monetary arrangement have been extensively characterized by the literature: i) it distorted relative prices; ii) obscured business and national accounting; iii) caused an implicit cross-system of taxation and subsidies; and iv) created a complex and non-transparent institutional environment for monetary and fiscal policy coordination (Doimeadios and Hidalgo, 2011, 109–15; Vidal, 2008; Di Bella and Wolfe, 2008, 50). A key consequence of this system was significant and heterogeneous price distortions at the firm level. The overvalued official exchange rate implicitly subsidized imported material inputs while simultaneously taxing exported goods, creating a cross-system of taxes and subsidies among heterogeneous producers. This mechanism became a major driver of resource misallocation in the economy.

I conclude this section by presenting some suggestive evidence on resource

[5] The exchange rate was pegged at this value in the mid-2000s; until then, it operated as a banded floating exchange rate.



Source: Authors' calculations based on Tables 5.4 and 7.3, ONEI, 2022. **agr** - Agriculture, Hunting, Forestry and Fishing; **min** – Mining, and Quarrying; **man** – Manufacturing; **pu** - Public Utilities (Electricity, Gas, and Water); **con** - Construction; **wrt** - Wholesale and Retail Trade, Hotels, and Restaurants; **tsc** - Transport, Storage, and Communications; **fire** - Finance, Insurance, Real Estate, and Business Services; **cspsgs** - Community, Social, Personal, and Government Services.

Figure 2.2: Correlation between sectoral productivity and change in employment shares in Cuba, 2009–2019

allocation during the reform period. Figure 2.2 illustrates changes in employment between 2009 (the baseline year before the reforms) and 2019 on the horizontal axis, with relative productivity (sectoral productivity over the economy-wide average) on the vertical axis. The figure reveals that, in 2009, most employment (represented by circle size) was concentrated in sectors with below-average productivity, such as social services and agriculture—evidence consistent with the narrative in this paper. However, during the reform years, employment shifted toward activities with relatively higher productivity, such as tourism. Nonetheless, the slope of the fitted regression line suggests that this correlation is not particularly strong and, in fact, turned out to be statistically insignificant. That said, although weak, this constitutes suggestive evidence that the reforms might have contributed to improving resource allocation—an issue I explore in greater detail in subsequent sections.

2.3 A Theoretical Model of Misallocation with Intermediate Inputs

I have previously discussed the country-specific mechanisms through which distortions and frictions could influence the optimal allocation of inputs at the firm level in Cuba. In this section, I outline the methodology used to analyze the relationship between resource misallocation, aggregate value added, and aggregate productivity. I examine this relationship through the lens of Hsieh and Klenow (2009); however, given the importance of intermediate inputs in my case study, I adopt the extended theoretical framework that incorporates this production factor, as developed in the literature by, among others, Bils et al. (2021); Dias et al. (2016). In particular, I follow this latter work.

The economy produces a single final good Y with a representative firm operating in a perfectly competitive final output market. The firm aggregates the output Y_s of S industries in the economy using a Cobb-Douglas production technology:

$$Y = \prod_{s=1}^{S} (Y_s)^{\theta_s}, \tag{2.1}$$

where θ_s represents the industry share and $\sum_{s=1}^{S} \theta_s = 1$. Given cost minimization, industry shares, θ_s, are given by:

$$\theta_s = \frac{P_s Y_s}{PY}, \tag{2.2}$$

where P_s is the price of industry gross output Y_s, and P is the price of the final good. The final good Y is assumed to be the numeraire, so I can set $P = 1$.

Industry-level gross output, Y_s, is represented as a CES aggregate of M_s differentiated products:

$$Y_s = \left(\sum_{i=1}^{M_s} (Y_{si})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{2.3}$$

where Y_{si} represents the gross output of firm i, and the parameter σ measures the elasticity of substitution between varieties of differentiated goods.

The gross output for each differentiated good is represented by a Cobb-Douglas production technology:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{\beta_s} X_{si}^{1-\alpha_s-\beta_s}, \tag{2.4}$$

where A_{si}, K_{si}, L_{si}, and X_{si} represent firm i's total factor productivity, capital stock, labor, and intermediate inputs, respectively. The elasticity of output with respect to the factors varies across industries but is the same for all firms within the same industry.

To capture the effect of distortions, wedges are introduced into the model. First, an output wedge, τ^y, represents any distortion that increases the marginal product of all inputs proportionally. Second, two input wedges, τ^l and τ^x, capture distortions that increase the marginal product of labor (or materials) relative to other inputs in the model. Wedges are typically modeled as taxes on revenues for output, or as taxes on labor and material use for input factors. Therefore, profits are given by:

$$\pi_{si} = (1-\tau_{si}^y)P_{si}Y_{si} - R_sK_{si} - (1+\tau_{si}^l)W_sL_{si} - (1+\tau_{si}^x)m_sX_{si} \tag{2.5}$$

where R_s, W_s, and m_s denote the cost of capital, wages, and the price of intermediate inputs, respectively.

From the (optimal) factor demands, I reveal the expressions for the τs from the factor ratios and the capital share.

$$\begin{aligned}(1+\tau_{si}^x) &= \frac{1-\alpha_s-\beta_s}{\alpha_s}\cdot\frac{R_sK_{si}}{m_sX_{si}}\\(1+\tau_{si}^l) &= \frac{\beta_s}{\alpha_s}\cdot\frac{R_sK_{si}}{W_sL_{si}}\\(1-\tau_{si}^y) &= \frac{\sigma}{\sigma-1}\cdot\frac{1}{\alpha_s}\cdot\frac{R_sK_{si}}{P_{si}Y_{si}}\end{aligned} \tag{2.6}$$

According to Equations 2.6, the three wedges are derived from data on gross output, input costs, and the elasticities α_s, β_s, and σ.

As common in this framewor, I distinguish between total factor quantity productivity (TFPQ) and total factor revenue productivity (TFPR) as follows:

$$TFPQ_{si} = A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s}L_{si}^{\beta_s}X_{si}^{1-\alpha_s-\beta_s}} \tag{2.7}$$

$$TFPR_{si} = P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}L_{si}^{\beta_s}X_{si}^{1-\alpha_s-\beta_s}} \tag{2.8}$$

Equation 2.7 indicates how many units of output a firm can obtain by using one unit of the mix of production factors, while Equation 2.8 shows how much revenue can be generated from the same quantity of factors.

It can also be shown that:

$$TFPR_{si} = \frac{\sigma}{\sigma - 1} \Psi_s \frac{(1+\tau^l_{si})^{\beta_s} (1+\tau^x_{si})^{1-\alpha_s-\beta_s}}{(1-\tau^y_{si})}, \tag{2.9}$$

where

$$\Psi_s = \left[\frac{R_s}{\alpha_s}\right]^{\alpha_s} \left[\frac{W_s}{\beta_s}\right]^{\beta_s} \left[\frac{m_s}{(1-\alpha_s-\beta_s)}\right]^{1-\alpha_s-\beta_s}$$

Equation 2.9 tells that TFPR is a weighted measure that depends on the wedges, and, more importantly, only varies among firms within the same industry when they face some type of distortion. In the presence of such distortions, a high (low) TFPR indicates that the firm faces barriers (benefits) that increase (decrease) its marginal product, making the firm smaller (larger) than optimal. In the absence of distortions, more capital and labor should be allocated to firms with higher TFPQ to the point where their higher output results in a lower price and the exact same TFPR as that of smaller firms.

To compute the aggregate gains in gross output when the distortions are removed -i.e., TFPR_{si} are equalized across firms within the industry- I proceed as follows. Firstly, I calculate the gains at the industry level using the following expression:

$$\frac{Y_s^*}{Y_s} = \left[\frac{\sum_{i=1}^{M_s} A_{si}^{\sigma-1}}{\sum_{i=1}^{M_s} \left(A_{si} \frac{TFPR_s^*}{TFPR_{si}}\right)^{\sigma-1}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{2.10}$$

where Y_s^*, Y_s, and TFPR_s^* denote the industry's efficient real output, the industry's observed real output, and the industry's efficient TFPR, respectively.

Secondly, using the Cobb-Douglas aggregator given by Equation 2.1, I obtain the aggregate gains in gross output:

$$\frac{Y^*}{Y} = \prod_{s=1}^{S} \left\{ \left[\frac{\sum_{i=1}^{M_s} A_{si}^{\sigma-1}}{\sum_{i=1}^{M_s} \left(A_{si} \frac{\text{TFPR}_s^*}{\text{TFPR}_{si}}\right)^{\sigma-1}}\right]^{\frac{\sigma}{\sigma-1}} \right\}^{\theta_s} \tag{2.11}$$

In the exercise, it is assumed that the total amount of inputs is fixed. Therefore, I note that potential gross output gains will coincide with potential TFP gains in Equation 2.11.

Given that value-added gains are more closely related to welfare gains, I convert gross output gains into value-added gains using a basic national accounting identity, as follows:

$$\frac{V^*}{V} = \frac{PY\left(\frac{Y^*}{Y}\right) - mX}{PY - mX} = \frac{\left(\frac{Y^*}{Y}\right) - q}{1-q}, \tag{2.12}$$

where V^* represents the efficient value added, V denotes the observed value added, $\frac{Y^*}{Y}$ is determined by Equation (2.11), and q refers to the aggregate share of intermediate inputs, $\frac{mX}{PY}$.

2.4 Data

I use firm-level data from manufacturing companies provided by the National Office of Statistics and Information (ONEI) for the period 2004–2017. Data come from annual censuses and integrate two primary sources: administrative records and financial statements. Specifically, it includes selected variables from Form 5901 (for the subperiod 2004–2013) and Form 5903 (2004–2015), as well as variables extracted from firms' balance sheets and income statements for the subperiod 2014–2017.

In the final sample, after removing missing values, outliers, and other recording errors, there are a total of 8,478 observations that span the years 2004 to 2017. These observations correspond to more than 800 state-owned enterprises (SOEs) at the beginning of the period and a small group of approximately 30 foreign firms -primarily joint ventures - included only for the years 2005–2006. A notable characteristic of the data is sample attrition, with a visible 58% reduction in the number of companies between the beginning and end of the period. However, this decline does not compromise the representativeness of data, as it reflects a contraction in the population of companies rather than selective exclusion.[6]

The information I use from the census includes gross output at market prices and value added. These are computed directly by the ONEI using original data from balance sheets and applying the National Accounts identities. I also use consumption of intermediate inputs (materials, energy, and fuel), labor compensation (including wages, bonuses, and benefits), employment (average number of employees), and capital stock. Capital stock is defined as the average book value of the company's machinery, equipment, and structures at the beginning and end of the year, net of depreciation.

The reported values represent annual totals and are categorized by economic activity at the four-digit level. For my computations, I assign industry shares based on those of the corresponding U.S. manufacturing industry. Correspon-

[6] In 2011, an institutional reform brought changes to the governance of state-owned enterprises. This led to a restructuring process that saw many companies merge or close down. For example, according to data from the Statistical Yearbooks (ONEI, 2022b), there were 887 companies in the manufacturing industry in 2006 (the first year in the series), but by 2017, that number had dropped to 423, more than a 50% decrease.

dence tables were used to establish the match, resulting in a total of 90 different industries. Furthermore, I observed other firm characteristics during different subperiods of the sample mergers and the firm's age. I will use these variables later to correlate them with TFPR.

2.5 Empirical Results

In this section, I empirically examine the effects of resource misallocation in Cuba's manufacturing industry. To do so, I extract key parameters from the data, including industry output shares, labor shares, material shares and firm-specific distortions. In my baseline computations I present the results for gross-output and value added reallocation gains using the sample for the period 2004-2017.

A key parameter in this methodology is σ, the elasticity of substitution between firms' gross outputs. Since the potential gains from liberalization rise with σ, it is standard in these studies to set $\sigma = 3$ for baseline computations—a conservative choice—and then perform sensitivity tests with higher values. I follow this convention in our analysis.

To measure labor input, I use the company's wage bill rather than its employment, L_{si}. Implicitly, this assumes that wages per worker adjust for differences across firms—a reasonable assumption given the growing importance of variable pay in total compensation.[7] This implies that $L_{si} = w_{si} L_{si}$ and $W_s = 1$, where L_{si} represents employment and w_{si} is the firm-specific average wage rate.

I make a similar assumption for intermediate inputs, setting the price of intermediate products, Z_s, equal to 1. This ensures that expenditure on intermediate inputs reflects not only their quantity but also their quality. On the other hand, I set the rental price of capital to $R = 0.10$, reflecting a 5% real interest rate and a 5% depreciation rate.

I do not observe each company's real gross output (Y_{si}), but rather its nominal gross output ($P_{si} Y_{si}$). To infer quantities from prices, I adopt the convention that firms with higher real output must have lower prices, explaining why buyers would demand the larger output. In line with this, I use nominal gross output and an assumed elasticity of demand to derive quantities. From the model, I obtain:

[7] In 2017, variable pay across the economy accounted for approximately two-thirds of total compensation. This type of remuneration is known in Cuba as *Pago por Resultados*, a bonus received by workers based on meeting certain targets.

$$Y_{si} = (P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}} \tag{2.13}$$

I use equation 2.13 to compute the company's real gross output. From the estimates of real output, I can derive firm-level total factor productivity ($TFPQ_{si} = A_{si}$).

Before calculating the gains from the hypothetical liberalization, I trim the 1% tails of $\log(TFPR_{si}/TFPR_s^*)$ and $\log(A_{si}M_s^{\frac{1}{\sigma-1}}/TFP_s^*)$ across industries, with $TFP_s^* = (\sum_{i=1}^{M_s} A_{si}^{\sigma-1})$. Specifically, I pool all industries and trim the top and bottom 1% of companies within each pool. I then recalculate the relevant industry-level aggregates, such as L_s, K_s, Y_s, P_sY_s, $TFPR_s^*$, TFP_s^*, and $\theta_s = \frac{P_sY_s}{Y}$.

2.5.1 Mapping Productivity Gaps

Figure 2.3 illustrates the distribution of TFPQ $\log(A_{si}M_s^{\frac{1}{\sigma-1}}/TFP_s^*)$ for the first and last years of the sample. The distribution shows a pronounced left skew in both 2004 and 2017, suggesting the existence of policies and institutions that enable the survival of inefficient firms (i.e., those with relatively low TFPQ). This quantitative evidence is far from surprising according to my narrative that Cuban SOEs operate under soft budget constraints, and it aligns with a vast body of literature focused on analyzing the characteristics of the business environment and how these undermine productivity growth at the micro level (del Castillo, 2017; Fernandez and Torres, 2020; Diaz et al., 2024).

However, the left tail becomes thinner over time, and the overall distribution appears noticeably less dispersed in 2017 compared to 2004, indicating that firm productivity converged over the period. In Table 2.1, I show that this pattern is consistent across years and several measures of TFPQ dispersion. The results suggest that the standard deviation of scaled TFPQ remained constant between 2004 and 2009, although the interdecile range declined from 1.62 in 2004 to 1.56 in 2009 (a 3.7 percent reduction). In contrast, between 2009 and 2017, the reduction is much more pronounced across all three dispersion measures: the standard deviation, interquartile range, and interdecile range decreased by 10.1 percent, 34.1 percent, and 7 percent, respectively. This significant convergence in TFPQ since 2011 could be a consequence of structural transformations implemented as part of the reform, an issue I will revisit later.

I additionally highlight the fact that the ratio of the 75th to the 25th percentiles of TFPQ is low compared to international evidence. For instance, in 2009 (the highest value), this ratio reached 2.2 (the exponential of the corre-

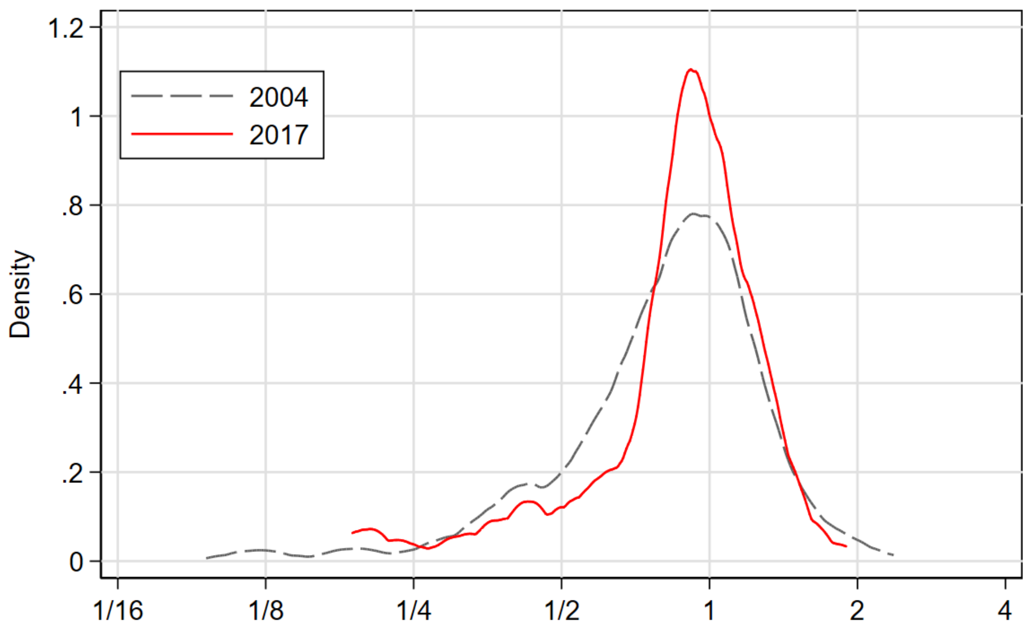


Figure 2.3: Scaled TFPQ

sponding number in Table 2.1). To provide context, this figure implies that the productivity of the firm at the 75th percentile of the distribution is twice as high as that of the firm at the 25th percentile. Hsieh and Klenow (2009) found differences of 5, 3.6, and 3.2 for India, China, and the United States, respectively. I attribute (at least part of) this result to the fact that productivity measurement is based on a gross-output methodology rather than a value-added approach. Gandhi et al. (2017) point out that productivity differences become substantially smaller and can even reverse when data are analyzed using gross output instead of value added. For example, they document a decline in the 90/10 productivity ratio in Chile from 9 to 2 when moving from value-added to gross-output productivity estimation. On the other hand, the outcome documented here is not exclusive to Cuba. For example, Kabiraj (2020) and Dias et al. (2016) found similar results for India and Portugal, respectively. The latter, which constitutes my closest methodological reference, reported a 75th-to-25th percentile ratio of 2.15, a figure close to mine.

Figure 2.4 illustrates the distribution of scaled TFPR (log $TFPR_{si}/TFPR_s^*$) for the years 2004 and 2017. Additionally, Table 2.2 provides dispersion statistics for the year 2009. Under an efficient allocation of resources, marginal returns are equalized across firms, resulting in a TFPR dispersion of zero. Conversely, greater dispersion of TFPR indicates a worse (mis)allocation of re-

Table 2.1: Dispersion of TFPQ

	2004	2009	2017
SD	0.69	0.69	0.62
75-25	0.75	0.79	0.52
90-10	1.62	1.56	1.45

Note: Statistics are for $\log\left(A_{si} M_s^{\frac{1}{\sigma-1}} / TFP_s^*\right)$. SD represents the standard deviation. The 75-25 and 90-10 rows show the interquartile range and the decile range, respectively. Industries are weighted by their gross-output shares.

sources among firms, driven by distortions that differently affect firms within each industry. In such cases, greater aggregate productivity gains could be achieved by reallocating inputs from firms with low marginal returns to those with higher marginal returns.

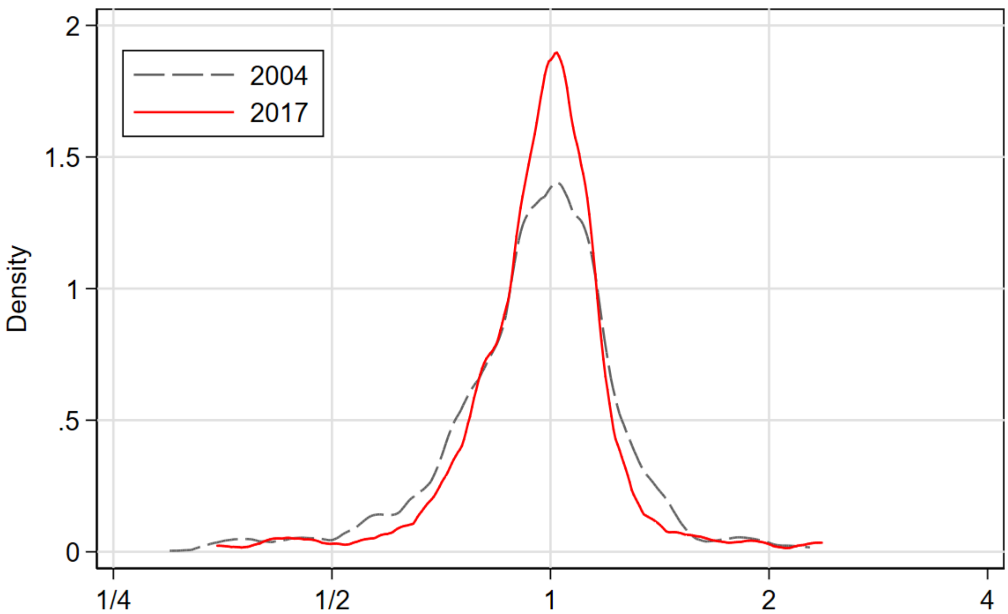


Figure 2.4: Scaled TFPR

From Figure 2.4 and Table 2.2, I conclude that resource misallocation declined over the studied period. All dispersion measures decreased, although, once again, the reduction from 2009 to 2017 was more pronounced than that from 2004 to 2009. Although the assessment of the long-term impacts of the reform is still a work in progress, I find suggestive evidence indicating that it may have contributed to productivity improvement through the misallocation channel—at least during the first half of the decade. Reforms represented a significant shift from the previous economic model, gradually evolving from 2011 onward. Although Cuba's reforms did not reach the depth of those implemented in China or Vietnam, the resulting economic model became considerably more market-oriented, with increased private sector participation, greater decentralization of price and wage setting, and more flexible corporate governance. I suggest that these changes may have contributed to the observed reduction in resource misallocation.

Table 2.2: Dispersion of TFPR

	2004	**2009**	**2017**
SD	0.38	0.36	0.33
75-25	0.40	0.39	0.31
90-10	0.85	0.82	0.63

Note: Statistics are for $\log(TFPR_{si}/TFPR^*_s)$. The 75-25 and 90-10 rows show the interquartile range and the decile range, respectively. Industries are weighted by their gross-output shares.

To better understand the potential sources of misallocation, I analyze the correlation between TFPQ and TFPR. A positive correlation suggests that more (less) physically productive firms face higher (lower) implicit distortions. This scenario is more damaging to aggregate productivity than one where TFPR is uncorrelated with TFPQ, as highlighted by Restuccia and Rogerson (2008) and Dias et al. (2016). Without such a correlation, dispersion in TFPR may not result in significant productivity losses, meaning the potential efficiency gains are likely to be small.

Table 2.3: Correlation between TFPR and TFPQ

	Average 2004-2017	**2004**	**2009**	**2017**
Correlation	0.69	0.73	0.65	0.69
N	8221	813	733	349

Note: Entries are pairwise correlation coefficients between $\log\left(A_{si}M_s^{\frac{1}{\sigma-1}}/TFP^*_s\right)$ and $\log(TFPR_{si}/TFPR^*_s)$. All coefficients are significant at the 1% level. N: number of observations. Industries are weighted by gross output shares.

Table 2.3 shows that the correlation between TFPQ and TFPR is 0.69, with statistical significance at the 1% level. This suggests that more productive firms face greater constraints and, as a result, underproduce, while less productive firms tend to overproduce. However, the correlation weakens slightly over time, with the decline being more pronounced in 2009 than in 2017 relative to 2004.

In the next subsection, I focus on comparing "efficient" and actual output levels to examine the welfare cost of misallocation.

2.5.2 Simulating Efficiency Gains

I simulate efficiency gains by assuming marginal products are equalized across companies within each industry. Aggregate gross output gains are calculated using Equation (2.11) and then converted into value-added gains using Equation (2.12). Computation is performed for each year in the sample, although I only report the values for 2004, 2009, and 2017.

Table 2.4 presents the results. I documented a noticeable decrease in both gross-output and value-added gains when comparing the first and last year of the analysis. For instance, between 2004 and 2017, gross-output gains decreased from 26% to 17%, translating into cumulative efficiency improvements of 7.7% over time (1.26/1.17). The overall decline reflects the reduction in TFPR dispersion mentioned in the previous subsection. As noted earlier, lower dispersion is indicative of a better allocation of resources.

Secondly, although efficiency gains are modest when measured in terms of gross output, they become much more significant when measured in terms of value added. For instance, in 2004, the gains were 26.8% for gross output compared to 61.5% for value added. This is a common outcome in applications similar to mine, partly driven by methodological factors. According to Dias et al. (2016), it stems from the assumption of constant intermediate inputs in the computation of gross-output gains. Equation (2.12) suggests that even small gains in gross output can translate into much larger value-added gains, especially in industries where intermediate inputs constitute a large share of gross output.

Table 2.4: Output gains from equalizing TFPR within industries, %

Year	Gross-Output	Value-Added
2004	26.8	61.5
2009	27.0	76.9
2017	17.4	44.5

Note: Gross-output gains and value-added gains are calculated using (2.11) and (2.12), respectively. Industries are weighted according to their gross-output shares.

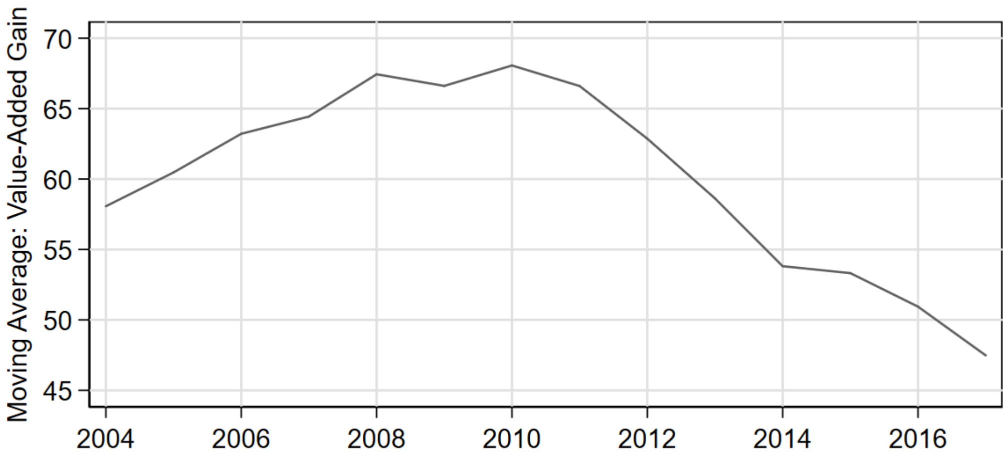


Figure 2.5: Trend in Value-Added Gains, 2004-2017

Note: The plot shows a moving average filter applied to the estimates from Equation (2.12).

I also examine how value-added gains from reallocation align with actual value-added growth, using figures from the statistical yearbook as a reference. The decline in gains from 61% to 44% between 2004 and 2017, as shown in Table 2.4, translates into cumulative efficiency growth of 11.8% over the period (or 0.9% per year). Given that real manufacturing value-added grew at an average annual rate of 2.6% during the same period (ONEI, 2022a), this implies that around one-third of this growth may be linked to improved resource allocation—a noteworthy proportion. However, the fact that this ratio reflects an annual average does not allow us to precisely capture that the efficiency improvement is primarily driven by the dynamics of the indicator after 2011. Figure 2.5 shows that value-added gains initially increased, peaking around 2010, before beginning to decline in 2011. This substantial improvement in allocative efficiency toward the end of the period—once again consistent with my narrative of structural reforms—appears to account for the ratio's value. Before 2010, efficiency showed no improvement.

Figure 2.6 plots the efficient vs. actual size distribution of companies in 2017, with size measured as company gross output. The fact that the hypothetical efficient distribution is more dispersed than the actual one suggests that, in the absence of misallocation, there would be fewer mid-sized companies and more small and large companies. I provide further insight into this finding in Table 2.5. It illustrates how the size of initially large and small companies would change if TFPR were equalized across companies within the same industry.

The entries represent unweighted shares of companies in 2004 and 2017. Rows correspond to initial (actual) company size quartiles, while columns categorize efficient company size relative to actual size: 0%–50% (the company should shrink by half or more), 50%–100%, 100%–200%, and 200%+ (the company should at least double in size).

The column with the highest concentration of companies is 50%–100% across all initial size quartiles, indicating that 48% of SOEs of all sizes would shrink in 2004 and 54% in 2017. This means that despite the substantial increase in average output, approximately half of the state-owned companies would face downsizing. Business "gigantism" is a widespread phenomenon, widely recognized in Cuba. It is commonly attributed to companies' tendency toward backward vertical integration, driven by supply chain vulnerabilities (Alvarez, 2014). This strategy seeks to mitigate risks associated with structural shortages throughout the supply chain and inefficiencies from external providers.

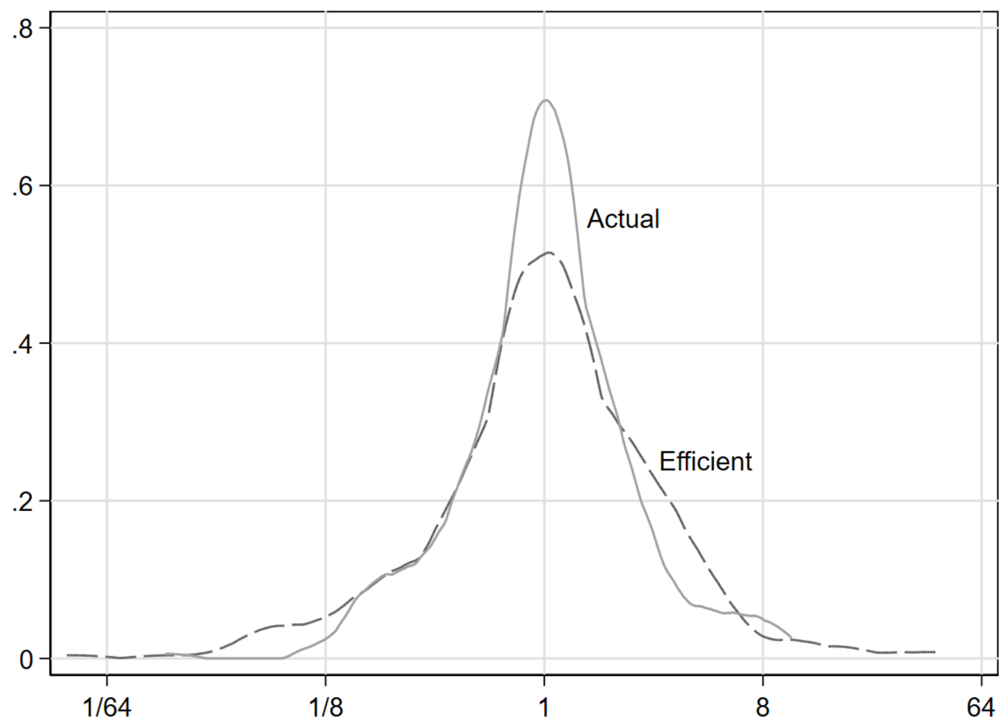


Figure 2.6: Distribution of Company Size

Note: 'Efficient' refers to the hypothetical distribution of gross output (as a measure of size) if TFPRs were equalized across firms within the same industry. In contrast, 'Actual' refers to the observed distribution, that is, without such equalization.

Here, I propose a complementary theory aligned with my framework. Some

Table 2.5: Percentage of Companies: Actual Size Compared to Efficient Size

Year	**0-50**	**50-100**	**100-200**	**200+**
2004				
Top Size Quartile	3.6	11.3	9.0	1.2
2nd Quartile	2.3	11.4	9.6	1.6
3rd Quartile	1.5	12.5	10.3	0.6
Bottom Quartile	0.7	12.2	11.7	0.4
Total	**8.1**	**47.5**	**40.6**	**3.8**
2017				
Top Quartile	1.1	12.3	10.0	1.7
2nd Quartile	1.1	13.5	9.5	0.9
3rd Quartile	1.4	14.3	8.6	0.6
Bottom Quartile	0.0	13.8	10.6	0.6
Total	**3.7**	**53.9**	**38.7**	**3.7**

Note: Companies are categorized into quartiles based on their actual gross output, ensuring an equal number of companies in each quartile. The hypothetically efficient gross output for each company is then calculated under the assumption that distortions are eliminated and TFPR levels are equalized across industries. The table above displays the percentage of companies with efficient-to-actual output ratios falling into four categories: 0%–50% (efficient output is less than half of actual output), 50%–100%, 100%–200%, and 200%+ (efficient output is more than double the actual output). Each row sums to 25%, and the rows and columns together total 100%.

Cuban companies expand as a result of distortions created by government policies. These distortions may stem from government support for strategic industries through mechanisms such as subsidies, favorable regulations, access to special financing, and other advantages. Without these benefits, their scale of operations would be significantly smaller—in other words, they would not be able to sustain the same level of resource access. Such distortions are characteristic of developing economies (Busso et al., 2013). In Cuba, evidence indicates that most SOEs are disproportionately large relative to what they would be under an efficient allocation.

2.5.3 Robustness Checks

I conducted some robustness checks to evaluate how sensitive results are to changes in the model's key parameters. The analysis begins with the elasticity of substitution. The model predicts that the gains from liberalization increase with σ. In my baseline estimates, I set $\sigma = 3$, following the methodological standard used in similar applied studies. This value is widely seen as conservative, representing the lower bound of empirical estimates observed internationally. However, I acknowledge that it may not be as conservative in the Cuban context due to the unique characteristics of its markets. Since there are no prior estimates for the Cuban case to anchor this parameter, I arbitrarily increased σ to 5 to analyze how the results change.

Estimates indicate moderate sensitivity to this elasticity. For instance, hypothetical value-added gains in 2009 increase from 77% under $\sigma = 3$ to 105% with $\sigma = 5$. Meanwhile, in 2017, these gains rise from 45% to 64%. Intuitively, this means that TFPR gaps close more slowly in response to the reallocation of inputs from low to high TFPR firms. However, I also observe that these gains are slightly lower compared to similar studies, such as those by Hsieh and Klenow (2009); Dias et al. (2016). There is no clear reason to which this difference can be attributed, a matter that will require further investigation.

On the other hand, in baseline calculations, I used the wage bill as a proxy for labor input to account for differences in worker skills and hours worked. However, this approach could potentially underestimate the differences in TFPR, and consequently the estimated reallocation gains—due to rent-sharing mechanisms, where more profitable firms pay higher wages. To address this concern, I compared baseline estimates with those derived using the average number of workers, expecting to observe an increase in the estimated gains.

Findings reveal minimal differences (albeit slightly lower) between the two measures. In 2004, the reallocation gains are 61.5% when labor input is measured by the wage bill and 60.6% when measured by employment. Similarly, in 2017, these figures are 44.5% and 43.7%, respectively. Therefore, my results are not particularly sensitive to how employment is measured—a finding that is unsurprising given the rigidity of wages in Cuba, which are determined by administrative criteria and state controls.

I also investigated whether the efficiency gains are the result of measurement errors. While more robust methods are available to address this issue (e.g., Bils et al., 2021), I chose a simpler approach by examining the impact of extreme outliers. In baseline estimates, I trimmed the 1% tails of TFPR and TFPQ in each year, amounting to 3% of the total observations. When I trim the 2% tails (6% of observations), value-added gains fall from 61.5% to 58.4% in 2004, and from 44.5% to 38.7% in 2017. Thus, measurement error in the remaining 1% tails could well be important, but it does not come close to accounting for the substantial gains from equalizing TFPR. Additionally, when I repeated this exercise using alternative thresholds, I did not find significant differences in the results.

Despite this finding, which supports the hypothesis on misallocation in Cuba, I cannot completely rule out the possibility of some type of measurement error in the data—a matter that will require further investigation in the future.

2.6 Potential Mechanisms

I have suggested that TFPR dispersion in Cuba is more than just the simple result of measurement errors. In this section, I take a further step and examine whether TFPR variation across firms could additionally be linked to explicit government policies, focusing on three specific examples (Ownership, Age and Mergers). I am aware that my analysis, based on correlations, does not imply causality. However, the findings align with extensive anecdotal evidence, which are consistent with the quantitative evidence provided.

2.6.1 Ownership

First, I examine potential differences in TFPR and TFPQ between state-owned and foreign-owned firms. Unfortunately, ownership data is only available for two years in sample—2005 and 2006—limiting the analysis to this restricted timeframe. I acknowledge that the subsample size is not ideal, comprising 30 foreign-owned companies, which represent just over 50% of the population of foreign-owned manufacturing firms during 2005–2006.

Table 2.6: TFPQ and TFPR Regression by Ownership

2005-2006	**TFPR**	**TFPQ**
Foreign	0.156*	0.345**
	(0.091)	(0.146)
N	1554	1554

Note: The dependent variable is the deviation of log TFPR or log TFPQ from the industry's efficient mean. *Foreign* is a dummy variable equal to 1 if the firm is a foreign joint venture. Entries are coefficients from simple regressions with clustered standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. N: number of observations. All regressions include time and industry dummies. Results are pooled for all years.

In Table 2.6, I documented approximately 15% higher TFPR and 35% higher TFPQ at foreign companies compared to state-owned companies in 2005–2006. This result means that foreign firms are, on average, more productive than state-owned firms but face greater implicit barriers that ultimately limit their growth. That foreign firms outperform state-owned ones is hardly surprising. They generally enjoy greater decision-making autonomy, better access to international supply chains, advanced technologies, and a more skilled and motivated workforce, among other advantages. Additionally, the sample includes export joint ventures of internationally recognized brands in industries such as beverages and tobacco, as well as other food companies with a prominent presence in the Cuban market.

However, they also encounter significant operational and institutional restrictions (referred to as 'internal obstacles' by Travieso-Diaz and Trumbull, 2003, 919-938). For instance, they face bureaucratic impediments, labor contracting frictions—including higher currency exchange taxes—or experience significant challenges in repatriating profits to their parent companies (Travieso-Diaz and Trumbull, 2003; Triana, 2021; Castro and Saenz, 2021). This often makes them reluctant to scale production, invest further, or procure additional raw materials. In theory, these implicit barriers would explain why, despite their higher productivity, foreign firms do not expand, and thus under-produce relative to their potential level. As (Feinberg, 2013, 18) noted, FDI employs less than 1% of the active labor force in Cuba, and its contribution to capital accumulation has lagged behind other economies with similar characteristics. To what extent many of these restrictions—and therefore their quantitative expression—were mitigated by the reforms introduced after 2011 is a question I leave unanswered due to the unavailability of data. Nevertheless, many of the issues described persist in one form or another (see, e.g., Triana, 2021).

2.6.2 Age

Next, I examine whether a firm's age, defined as the time elapsed since its founding, correlates with TFPR. In conducting this analysis, I use firm age as a proxy for its relative importance or relevance within the economy. This approach is rooted in the hypothesis that firms with longer operational histories often enjoy greater social recognition. Such recognition may arise from two potential sources. First, due to genuine competitiveness, i.e., older firms might objectively outperform others because of higher productivity, as is often the case with exporters. Second, due to strategic government support, where certain industries or firms may be prioritized by the state for reasons unrelated to their actual productivity. This latter could reflect industrial policy objectives or political considerations.

In both cases, this relatively superior status increases the likelihood of receiving preferential treatment, such as subsidies or favorable regulations. My objective is to evaluate whether this relative position translates into tangible economic benefits. Specifically, if such benefits are present, I would anticipate observing a negative correlation between firm age and TFPR.

In dataset, *Age* is a discrete variable that captures the firm's age in years, ranging from zero (indicating the year of establishment) to forty. In Figure 2.7, I present the correlation between scaled TFPR and scaled TFPQ with firm age. The results indicate that age is associated with higher productivity in quantities (bottom panel) and lower productivity in revenues (top panel). Complementary regression analysis confirmed that the coefficients are significant at the 1% level.[8]

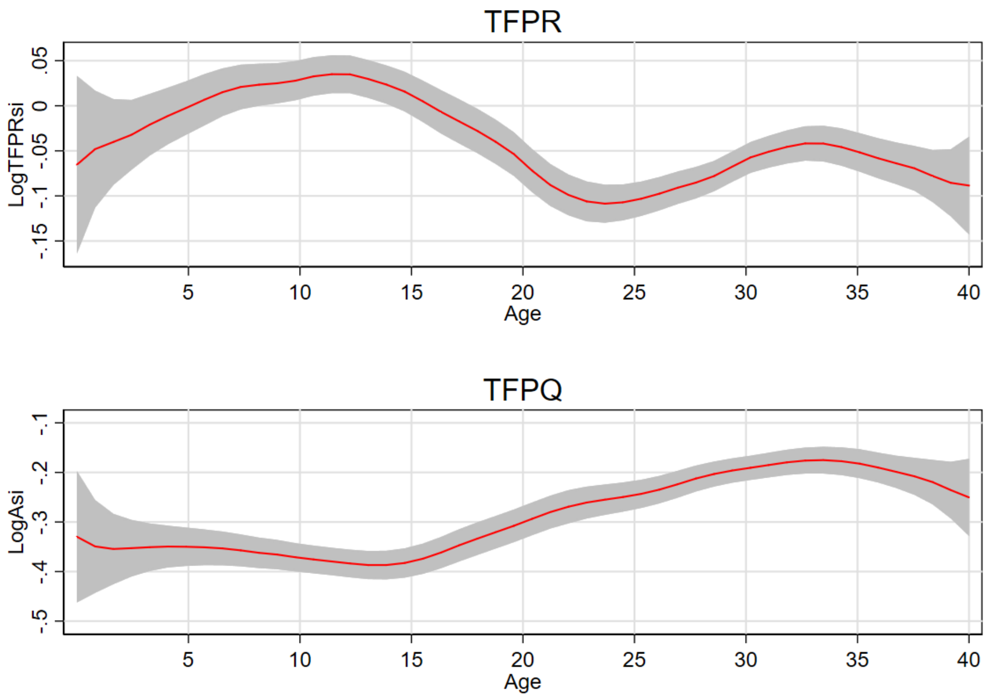


Figure 2.7: Correlation Between TFP and Age

Note: The red line represents the Kernel-weighted local polynomial smooth. The gray area is the 95% confidence interval. Data pooled.

From these results, I conclude that the government favors older firms, which, coincidentally, are also the most productive. The fact that older firms are also the most productive suggests the presence of a selection mechanism –defined in a broad sense– that favors the continued operation of companies with relatively higher performance. At the same time, the observation that older firms exhibit, on average, a lower TFPR means that they receive greater implicit benefits,

[8] Results available upon request.

which are likely part of this mechanism. However, it is also worth noting that the relationship between benefits and firm age is non-monotonic. For example, firms aged 23–33 face, on average, increasing barriers, even though they are also more productive.

On the other hand, the fact that, on average, scaled TFPQ is relatively lower for young firms surprised me, as it runs counter to the expectation that entrants are generally more productive than incumbent firms. Similarly, the finding that less productive firms in terms of quantity are also the ones facing more barriers challenges the soft budget constraints hypothesis discussed in Section 2.2. This hypothesis suggests that the paternalistic state tends to "shield" insolvent firms from bankruptcy.

A possible explanation for the apparent rejection of this hypothesis may lie in the merger policy implemented by the Cuban government as part of the 2011 reform. Gonzalez and Torres (2024) suggests that, rather than directly subsidizing insolvent assets, the state chose to merge them with more profitable or productive firms. This strategy was considered a socially less costly alternative to providing subsidies or permanently liquidating the entities. It could be the explanatory factor –for instance– behind the productivity convergence observed in Figure 2.3. I explore this further in the next subsection.

2.6.3 Mergers

I investigate whether the merger policy implemented since 2011 correlates with misallocation indicator, drawing on the literature that explores the impact of mergers on productivity (see, e.g., Braguinsky et al., 2015). Acquiring firms experienced substantial growth in both resources and output levels following the policy's implementation. For instance, pooling data from 2011 to 2017 reveals that the average gross output ratio between acquiring and non-acquiring firms was 2.9. Similarly, the corresponding ratios for capital, labor, and raw materials were 4.0, 2.7, and 2.6, respectively, underscoring the disparities in resource allocation between these two groups. In other words, the mergers represented a significant reallocation of factors and outputs for the SOEs involved. Mergers also contributed to a significant reduction in the number of SOEs, decreasing by approximately one-third after 2011 (ONEI, 2022b).

Theoretically, more productive firms are expected to "acquire" less productive ones, with no subsequent decline in productivity for either the acquiring or the acquired firms post-merger. While testing these hypotheses lies beyond the scope of my paper, and the currently available data do not permit a thorough exploration, I provide some suggestive evidence supporting this notion. Given

Table 2.7: TFPQ and TFPR Regression by Mergers

2011-2017	**TFPR**	**TFPQ**
Merge	-0.088**	0.223***
	(0.038)	(0.064)
N	2932	2932

Note: The dependent variable is the deviation of log TFPR or log TFPQ from the industry's efficient mean. *Merge* is a dummy variable equal to 1 if the firm acquired other firms in a merger process. The variable *Age* is added as a control. Entries are coefficients from regressions with clustered standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. N: number of observations. All regressions include time and industry dummies. Results are pooled for all years.

the administrative nature of the merger process in Cuba, I suspect that the exogenous allocation of resources to acquiring firms contributed to increased dispersion in TFPR, suggesting that mergers acted as a form of "benefit" for the acquiring firms.

When comparing TFPR and TFPQ between merged and non-merged firms within the same industry—where the variable *Merge* is binary and equals one if the firm participated as an acquirer in such a process—I find evidence supporting this suspicion. In Table 2.7, I show that acquiring firms were 22% more productive in quantities but 8% less productive in revenues compared to non-merged firms. This is a clean effect, free from the potential influence of firm age, which I included as a control variable in the regression. If the hypothesis from the previous subsection had been correct, it might have been the case that the government merged less productive firms with older ones, which, as I recall, are also the most productive. By controlling for firm age, I account for this potential correlation.

I interpret the regression results as follows: the government reallocated resources to firms that were, on average, 22% more productive. However, I clarify that I do not distinguish whether there was a change in the SOEs' TFPQ before and after the acquisition. On the other hand, the 8%-lower TFPR indicates that these SOEs have an "excess" of resources, meaning they overproduce relative to their actual scale. Once again, I cannot determine whether this result is due to a change in TFPR before and after the acquisition. Given the static nature of this analysis, I cannot attribute this outcome directly to the merger policy. A question that remains open for future papers.

2.7 Final Remarks

This paper presents the first quantitative assessment of the potential effects of misallocation on TFP in Cuba. Using the methodology developed by (Hsieh and Klenow, 2009), I analyze the degree of firm-level productivity heterogeneity and sector-specific distortions. Findings reveal significant dispersion, indicating substantial potential gains from transitioning toward a more efficient allocation of resources within the manufacturing sector. Specifically, I estimate these gains to range between 45% and 75%.

These pronounced productivity disparities and the extent of resource misallocation highlight critical opportunities to boost aggregate productivity—an essential priority for Cuba, given its severely constrained short-term growth drivers. Furthermore, my analysis documents an improvement in allocative efficiency following the 2011 reforms. This underscores the pivotal role of these reforms in addressing distortions inherent in a centrally planned economy.

By shedding light on the scale of misallocation and its implications, this study emphasizes the urgency of implementing targeted reforms to unlock Cuba's productivity potential.

3 Currency Misalignment and Resource Misallocation

3.1 Introduction

Resource misallocation has been identified as one of the main determinants of income differences between countries (Jones, 2016). In a nuthsell, when production factors are inefficiently allocated among heterogeneous producers, observed aggregate output falls below its potential level, leading to overall welfare losses. The causes of misallocation are diverse. As highlighted in the survey by Restuccia and Rogerson (2017), they include regulations, weak property rights, trade protectionism, limited competition, financial market imperfections (e.g., credit constraints), corruption, and other distortionary policies.

In this paper, I investigate the relationship between misallocation and currency misalignment. Misalignment refers to a situation where the real exchange rate deviates significantly from its equilibrium value (Hinkle and Monteil, 1999). The latter is typically defined as the rate that ensures both internal and external balance, given the underlying fundamentals of the economy. Although the causes of misalignment are varied (Eden and Nguyen, 2012), this paper focuses on a very specific context: currency overvaluations that arise under fixed exchange rate regimes. The prolonged use of such regimes—or even crawling pegs—is usually associated with a higher likelihood of current account imbalances, depletion of international reserves, real exchange rate overvaluations, losses in external competitiveness, and an increased risk of a currency crisis (Coudert and Couharde, 2009).

Within this context, I examine the effects on misallocation in the extreme case where monetary authorities have exhausted their international reserves due to the relative cheapening of foreign currency and are, thus, unable to sustain the convertibility of the domestic currency at the official nominal rate. At that point, some governments opt for an unconventional strategy based on the introduction of foreign exchange controls instead of acting conventionally by devaluing the economy's nominal exchange rate (Gray 2021). The policy of exchange controls represents an attempt to administratively restrict excess demand for foreign currency. It is often justified by the belief that it will help control inflation, direct scarce foreign exchange resources toward national priorities, and promote income distribution by safeguarding vulnerable social sectors.

Foreign exchange controls are, in themselves, a sufficient condition for the emergence of parallel currency markets and multiple exchange rates (see, for example, Hinkle and Monteil 1999 and Gray 2021 for a detailed discussion). Broadly defined, multiple exchange rate systems refer to exchange rate regimes in which two or more exchange rates are applied to transactions involving the same currency (Marion 1994, 213). Monetary authorities typically set a fixed official exchange rate (e hereafter), where the demand for foreign currency—managed through the aforementioned controls—is not fully satisfied. Meanwhile, market forces—often in a black market, although occasionally legalized by governments—determine a floating exchange rate on the parallel foreign exchange market (z hereafter). The exchange rate gap—generally positive, i.e., $z - e > 0$—forms the basis of multiple distortions, as summarized by Gray 2021 (pp. 8–10).

One of the numerous consequences of multiple exchange rate regimes is the creation of a complex system of currency transfers (taxes and subsidies) among individual producers, which distorts relative prices, incentives, and resource allocation. Some privileged agents, who have preferential access to foreign currency at the official exchange rate, benefit from a subsidy per unit of foreign currency at a rate of $s = \frac{z-e}{z}$. Conversely, disadvantaged agents, particularly those who are net suppliers of foreign currency at the official exchange rate (such as exporters), effectively bear a tax per unit of foreign currency at the same rate, $t = \frac{z-e}{z}$.

These transfers impact individual producers heterogeneously, depending on their degree of connection to foreign markets. International trade models underscore that decisions to export are intricately tied to firm-specific characteristics, even among firms within the same industry (Bernard et al., 2007; Melitz, 2003). Meanwhile, the discretionary nature of foreign exchange controls inherently prioritizes certain agents over others, exacerbating the unequal distribution of these transfers.

In this paper, I develop a model to quantitatively assess the effects of misalignment on resource misallocation. My model is a simple extension of Hsieh and Klenow (2009) for open economies, in which I incorporate the exchange rate based on a series of simplifying assumptions (essentially, it is Hsieh and Klenow (2009) with the exchange rate included). The baseline model suggests that Hsieh and Klenow (2009) represents a particular solution when the real exchange rate is in equilibrium. Using this model, I conduct a counterfactual analysis to reveal the welfare consequences associated with multiple exchange rates.

I then apply this model to the Cuban case, which represents an interesting and underexplored example of misallocation in the literature. The interest in Cuba is driven not only by its historical past as a centrally planned economy (now transitioning to a mixed model) but—above all—by its highly peculiar monetary system. Between 1995 and 2021, Cuba operated a dual monetary system based on two domestic currencies and two legal exchange rates: an official rate and a parallel rate. To the best of my knowledge, the spread between these two rates –reaching 2400%– is unprecedented in the literature. This figure represents the tax (subsidy) paid (received) by each agent selling (buying) one US dollar at the official exchange rate.

I focus my analysis on the year 2015 in the manufacturing industry, a period during which this dual monetary system and the resulting distortions were still in effect. I also chose 2015 because it was a "good year": real GDP grew by 4.4% (ONEI, 2016, Table 5.1), annual inflation stood at 2.8% (ibid.), the unemployment rate was just 2.4% (ibid.), and relations with the United States had improved significantly following an unprecedented rapprochement under the Obama administration.

The results, however, are striking. The counterfactual exercise suggests that when the exchange rate gap and the resulting transfer system are internalized, the variance of marginal revenue products (MRP) between firms within the same industry increases significantly compared to the baseline model. The greater dispersion implies that, when simulating efficiency gains through the equalization of MRPs, the TFP gains rise from 54% in the baseline model to 106% in the counterfactual model. The difference of more than 50 percentage points is attributed to the effect of the 2400% exchange rate gap.

Related Literature and Contributions My work contributes to the literature by bridging two strands of research. On one hand, the literature linking the real exchange rate to long-term output—what Frenkel and Ros (2006) referred to as the "development channel." In general, this channel is based on the idea that RER devaluation improves the relative profitability of the tradable sector, which, in turn, affects long-term output through three mechanisms: (i) savings and investment, (ii) structural diversification, and (iii) factor reallocation[1] (see the survey by Demir and Razmi (2022)).

[1]This mechanism differs from the intra-industry reallocation emphasized in the misallo-

My work introduces a complementary mechanism based on a system of currency taxes and subsidies that is activated when the exchange rate becomes misaligned. In my model, changes in relative profitability are associated with this transfer system, and the final effect of a devaluation on real variables depend on the remaining non-exchange rate distortions. In the specific case of Cuba, I argue that a nominal exchange rate devaluation would have positive real effects on output because it would reduce the distortions caused by the transfer system, particularly by alleviating the exchange rate tax imposed on the export sector.

This paper also contributes to the literature on misallocation. Although misallocation has been studied in relation to other economic and financial variables, its connection to the exchange rate has received much less attention. Recent work by Weinberger (2020), based on the Variable Elasticity (VES) framework, shows that real exchange rate shocks can be an important driver of allocative efficiency due to the way firms pass through these shocks into markups. In this framework, changes in markups caused by exchange rate shocks play a central role in explaining variations in misallocation over time.

My paper differs in that exchange rate misalignment is absorbed by a "monetary wedge," (in the counterfactual exercise) which is interpreted as distortions similar to those described by Hsieh and Klenow (2009). These monetary wedges represent the aforementioned taxes or subsidies stemming from exchange rate policies, designed to ensure that markups remain constant despite the misalignment. By focusing on the role of monetary wedges, this paper highlights a different mechanism through which exchange rate misalignment affects allocative efficiency, complementing the insights from the VES framework.

Finally, I emphasize that this working paper is not (only) about Cuba. While multiple exchange rate systems are less common today than they were in the past, the types of unconventional policies that lead to exchange rate misalignment remain highly relevant. For instance, a recent article by J.P. Farah-Yacoub, Nada Hamadeh, and Carmen Reinhart highlights that nearly one-fifth of emerging and developing economies still have active parallel markets (Yacoub et al., 2022). On the other hand, although the model is designed to capture distortions in contexts of multiple exchange rates, its application

cation literature (Restuccia and Rogerson, 2017), which focuses on the existence of distortions that prevent the equalization of marginal products across firms within the same industry (Hsieh and Klenow, 2009). In contrast, it relies on inter-industry allocations favoring the tradable sector, which is often associated with higher average productivity, learning-by-doing, technological innovation, and a greater presence of economies of scale.

could be extended to other, less extreme cases of exchange rate misalignment. Therefore, the conclusions reached in my paper are broadly applicable to any economy grappling with this type of monetary challenge.

The paper is structured as follows. Following this brief introduction, Section 3.2 introduces the model used to capture foreign exchange transfers and their impact on welfare, which is essentially an extension of Hsieh and Klenow 2009. Section 3.3 examines the Cuban monetary model, characterized by two domestic currencies, partial dollarization, multiple exchange rates, and exchange controls, and also describes the data used for the estimates. In Section 3.4, I present and discuss the results. Finally, Section 3.5 concludes.

3.2 The Model

In this section, I develop a model to establish a theoretical channel linking exchange rate misalignment to resource misallocation. The baseline model is a simple generalization of Hsieh and Klenow (2009) (hereafter HK), where the exchange rate is incorporated into HK's seminal framework. It is shown that HK represents a particular solution when the exchange rate clears the foreign exchange market and other simplifying assumptions hold. Next, I build a counterfactual framework to quantitatively assess the impact of currency misalignment on TFP. The results show that the HK wedges are amplified compared to the baseline model.

3.2.1 The Baseline Model with no-Misalignment

A representative firm produces a single final good Y under the assumption of a small open economy. One part of Y is sold within the national borders and the other part is exported. The share of the quantities sold inside is denoted as Y^p, and the share traded outside is denoted as Y^d. The suffixes p and d represent pesos and dollars, respectively, and indicate the currency at which Y is sold, with $Y = Y^p + Y^d$. The representative firm combines the output Y_s from S manufacturing industries using a Cobb-Douglas production technology:

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s}, \tag{3.1}$$

where $\sum_{s=1}^{S} \theta_s = 1$ and $Y_s = Y_s^p + Y_s^d$.

The industry output price within national borders is P_s^p, denominated in pesos, while the industry output price outside national borders is P_s^d, denominated in dollars. Dollar-denominated prices are frictionlessly converted into pesos using the nominal exchange rate, e, under the assumption of the Law of One Price (LOOP), which implies that $P_s^p = P_s^d e$. The nominal exchange rate, defined as the number of pesos per dollar, is treated as an exogenous variable determined in the foreign exchange market, which is assumed to be in equilibrium. Any movement in e automatically translates into prices, with an effect equivalent to a direct variation in the prices themselves.

In this setting Y_s is sold at the weighted average price $\overline{P}_s^e$, defined as:

$$\overline{P}_s^e = P_s^p \frac{Y_s^p}{Y_s} + P_s^d \frac{Y_s^d}{Y_s} e \tag{3.2}$$

Note that, due to LOOP—where prices are equalized across both markets—$\overline{P}_s^e$ is, in fact, equivalent to the industry price in a closed economy as described in the HK model. Cost minimization implies:

$$\theta_s = \frac{\overline{P}_s^e Y_s}{\overline{P}^e Y}, \tag{3.3}$$

where $\overline{P}^e = \prod_{s=1}^{S} (\overline{P}_s^e / \theta_s)^{\theta_s}$ is the price of the final good which is assumed to be numeraire and therefore $P = 1$. Industry output, Y_s, is a CES aggregate of differentiated products M_s:

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si} \right)^{\frac{\sigma}{\sigma - 1}}, \tag{3.4}$$

where $Y_{si} = Y_{si}^p + Y_{si}^d$. σ stands for the elasticity of substitution between varieties of differentiated goods. Assuming the same demand function for both domestic and international trade, the (inverse) individual demand equation can be written as:

$$\overline{p}_{si}^e = Y_s^{\frac{1}{\sigma}} \overline{P}_s^e (Y_{si})^{\frac{-1}{\sigma}} \tag{3.5}$$

In Equation (3.5), the term $Y_s^{\frac{1}{\sigma}} \overline{P}_s^e$ is unobservable. A common practice is to normalize this term to 1 for each of the S industries, as this adjustment does not affect relative productivities or intra-industry reallocation gains—one of the primary objectives of this framework.

At the firm level, unit i from industry s produces gross output Y using a Cobb-Douglas technology:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{\beta_s} X_{si}^{1-\alpha_s-\beta_s}, \tag{3.6}$$

where K, L, and X represent capital stock, labor, and material input factors, respectively. A denotes total factor productivity. The output elasticities with respect to each factor of production, α, β, and $(1-\alpha-\beta)$, vary across sectors but remain constant within each sector.

I assume that there is no free international mobility of capital and labor. However, some of the intermediate material inputs are imported, while others are produced within the national borders, so that $X = X^p + X^d$. For imported materials, firms pay prices in dollars, m^d, while for locally produced materials, they pay prices in pesos, m^p. Assuming perfect competition in the market for intermediate inputs among firms within the same industry, the average price of intermediate material inputs is given by:

$$\overline{m}_s^e = m_s^p \frac{X_{si}^p}{X_{si}} + m_s^d \frac{X_{si}^d}{X_{si}} e \tag{3.7}$$

In addition, all three factors of production are subject to distortions—or wedges—*à la* Hsieh and Klenow (2009). As is usual under this conceptual framework, output distortions, τ_{Ysi}, denote any distortions that affect the marginal product of capital, labor, and intermediate inputs by the same proportion. In turn, τ_{Lsi} represents distortions that increase the marginal product of labor relative to capital and material inputs, while τ_{Xsi} captures distortions that raise the marginal product of material inputs relative to capital and labor (see Hsieh and Klenow 2009, 1407 for further details). Therefore, after-tax profits are given by:

$$\pi_{si}^e = (1-\tau_{Ysi})\overline{p}_{si}^e Y_{si} - R_s K_{si} - (1+\tau_{Lsi}) W_s L_{si} - (1+\tau_{Xsi}) \overline{m}_s^e X_{si} \tag{3.8}$$

where R_s and W_s represent the user cost of capital and the wage, respectively. In Equation (3.8), τ_{Ysi}, τ_{Lsi}, and τ_{Xsi} are modeled as taxes on income, labor, and material costs, respectively.

This model differs from HK in that output and material input prices are averages of dollar and peso prices, connected to each other through the nominal exchange rate, e. However, this difference does not introduce any variation in the HK methodological framework as long as LOOP and the other assumptions used are held. Any exogenous change in the nominal exchange rate will affect firms through output and material input prices, just as any other temporary change in relative prices would in the closed economy model of HK. When the exchange rate varies, it affects revenues and costs simultaneously, and the final effect on profitability will depend on which of the two effects prevails.

To illustrate this, I derive the following two expressions from total revenues and material input expenses:[2]

$$\phi^p_{si} + \phi^d_{si} e = 1 \tag{3.9}$$
$$\psi^p_{si} + \psi^d_{si} e = 1 \quad \forall e, \tag{3.10}$$

where $\phi^d_{si} = \frac{p^d_{si} Y^d_{si}}{\overline{p}^e_{si} Y_{si}}$ and $\psi^d_{si} = \frac{m^d_s X^d_{si}}{\overline{m}^e_s X_{si}}$, respectively.

The parameters ϕ^d_{si} and ψ^d_{si} are key elements of the model and are treated as firm-specific characteristics. They represent the share of revenues and material expenses in dollars, expressed as a proportion of total revenues and material expenses in both pesos and dollars, respectively.

Given the long-term nature of the model, when e varies —while holding all else constant— prices in pesos adjust automatically due to LOOP to maintain parity across both markets. However, the relative impact on firm profitability will differ depending on the relationship between ϕ^d and ψ^d. Specifically, if e increases, firms with $\phi^d > \psi^d$ will experience a rise in profitability as a result of the exchange rate devaluation. Conversely, firms where $\phi^d < \psi^d$ will see a decline in profitability. It is important to emphasize that such variations in the nominal exchange rate are not, by themselves, a source of resource misallocation. Firms will adjust their optimal choice of output and inputs accordingly, and no real effect would arise. Resource misallocation is inherently a real sector problem, and therefore, no significant real effects should be expected from changes in the nominal exchange rate.

To make this point clearer, Equation (3.8) can be rewritten as follows:

$$\begin{aligned}\pi^e_{si} = {} & (1-\tau_{Ysi})(\phi^p_{si} + \phi^d_{si} e)\overline{p}^e_{si} Y_{si} - R_s K_{si} - (1+\tau_{Lsi}) L_s W_{si} \\ & - (1+\tau_{Xsi})(\psi^p_{si} + \psi^d_{si} e)\overline{m}^e_s X_{si}. \end{aligned} \tag{3.11}$$

Since both $\phi^p_{si} + \phi^d_{si} e$ and $\psi^p_{si} + \psi^d_{si} e$ are equal to 1 $\forall e$, there is no additional effect on the wedges $(1-\tau_{Ysi})$ and $(1+\tau_{Xsi})$, respectively. This implies that there will be no further impact on resource misallocation. In other words, ϕ and ψ represent redundant information when e is in equilibrium and the assumptions are satisfied. The baseline model can, therefore, be solved as

[2] Total revenues are given by $\overline{p}^e_{si} Y_{si} = p^p_{si} Y^p_{si} + p^d_{si} Y^d_{si} e$, while material input expenses are given by $\overline{m}^e_s X_{si} = m^p_s X^p_{si} + m^d_s X^d_{si} e$.

in HK, yielding results equivalent to those of a closed economy model. In fact, it should be emphasized that the approach followed in this paper is fully equivalent to HK's closed economy model when $Y^d = X^d = 0$, that is, in the absence of international transactions.

3.2.2 The Counterfactual Model

Let us now examine the case of a permanent deviation in the equilibrium exchange rate caused by exchange control policies.[3] Once controls are imposed on the official foreign exchange market, a parallel currency market emerges. In addition to the official exchange rate, e, a parallel exchange rate, z, is set for some international transactions, with $z > e$. Unlike the situation described in the previous subsection, this is a real sector issue stemming from the inconvertibility of the domestic currency at the official rate.

With excess demand for dollars at e, foreign currency allocation is managed administratively, resulting in heterogeneous impacts on economic agents. Privileged net buyers of dollars at the official rate, typically importers, benefit from a currency subsidy, while disadvantaged net sellers, usually exporters, are burdened with a currency tax. Moreover, exchange rate misalignment incentivizes improper recording of transactions on corporate balance sheets, particularly when the parallel exchange rate is illegal. Although both issues are sources of misallocation, I will address only the first one at a theoretical level, while addressing the second one in the next section as a practical problem, using an institutional peculiarity of the case study (Cuba).

Unveiling the system of currency transfers is particularly challenging due to the fact that the economy's real exchange rate is typically an unobservable variable. However, the parallel nominal exchange rate, z, is generally considered more aligned with the real exchange rate than the official rate, e. Therefore, the assumption I use to proceed is based on this alignment. I will estimate the value of currency subsidies and taxes using z as a counterfactual state, implicitly assuming that the exchange rate z aligns the real exchange rate with its equilibrium level.

In this effort, the parameters ϕ^d and ψ^d—-recall, they represent the dollar proportion of firm revenues and material expenses—-will provide valuable information to quantitatively characterize the total magnitude of these taxes

[3] These policies often arise from a government's reluctance to devalue the exchange rate, driven by fears of fueling inflation. For example, when the exchange rate is fixed and foreign currency is inexpensive, controls are typically introduced after international reserves are depleted. Instead of devaluing, the government manages the excess demand for dollars through such measures.

and subsidies. In the previous subsection, it was stated that these parameters contained redundant information about misallocation under the assumption that LOOP and other conditions held. However, the introduction of exchange controls is a sufficient reason for LOOP violations and the activation of the system of exchange rate transfers. Now, understanding the dollar proportion of firms' expenses and revenues becomes crucial to estimating the level of currency subsidies and taxes they face.

To this end, I proceed as follows. First, I define the output and material input prices in the counterfactual state, i.e., the prices that firms would charge and pay, respectively, if the exchange rate were z instead of e.[4] Using the definition of average prices, it follows that: $\overline{p}^z_{si} = p^p_{si}\left(Y^p_{si}/Y_{si}\right) + p^d_{si}\left(Y^d_{si}/Y_{si}\right)z$ and $\overline{m}^z_s = m^p_s\left(X^p_{si}/X_{si}\right) + m^d_s\left(X^d_{si}/X_{si}\right)z$.

Then, I differentiate each of these prices with respect to $\overline{p}^e_{si}$ and $\overline{m}^e_s$, respectively, and multiply the results by Y_{si} and X_{si}, which yields:

$$\overline{p}^e_{si}Y_{si} = (1-\tau^z_{Ysi})\overline{p}^z_{si}Y_{si} \tag{3.12}$$

$$\overline{m}^e_s X_{si} = (1-\tau^z_{Xsi})\overline{m}^z_s X_{si}, \tag{3.13}$$

where

$$\tau^z_{Ysi} = \frac{\phi^d_{si}(z-e)}{1+\phi^d_{si}(z-e)}, \quad \tau^z_{Xsi} = \frac{\psi^d_{si}(z-e)}{1+\psi^d_{si}(z-e)}$$

Equations (3.12) and (3.13) reveal that a firm's revenues and material costs, valued at the official exchange rate e, represent a fraction $(1-\tau^z_{Ysi})$ and $(1-\tau^z_{Xsi})$ of the values they would have had if the exchange rate were z. In other words, τ^z_{Ysi} and τ^z_{Xsi} correspond to the currency tax rate and currency subsidy rate, respectively, induced by the misalignment. It is worth noting that these rates depend on firm-specific parameters ϕ^d and ψ^d, as well as the exchange rate spread. Both rates become zero if the exchange rate gap closes $(z-e=0)$ or if foreign trade disappears $(Y^d = X^d = 0)$. Although τ^z_{Ysi} and τ^z_{Xsi} are modeled as "taxes and subsidies" (typical of this framework), what they truly reveal is a more profound structural problem linked to the transfer system created by the exchange rate misalignment—a real economy issue.

Plugging Equations (3.12) and (3.13) into Equation (3.8), I obtain:

[4] Think of the counterfactual state as the immediate moment when policymakers, without prior notice, depreciate the exchange rate from e to z without allowing domestic prices to adjust.

$$\pi^z_{si} = (1-\tau'_{Ysi})\overline{p}^z_{si}Y_{si} - R_sK_{si} - (1+\tau_{Lsi})W_sL_{si} - (1+\tau'_{Xsi})\overline{m}^z_sX_{si} \tag{3.14}$$

where:

$$1-\tau'_{Ysi} = (1-\tau_{Ysi})(1-\tau^z_{Ysi}) \tag{3.15}$$
$$1+\tau'_{Xsi} = (1+\tau_{Xsi})(1-\tau^z_{Xsi}) \tag{3.16}$$

Equation (3.14) requires some explanations. First, π^z_{si} is now a counterfactual profit function, i.e., the profits the firm would make if the prices of output and material inputs were $\overline{p}^z_{si}$ and $\overline{m}^z_s$, respectively. Of course, neither of these prices are observable, an issue to which I will return later. Second, the output wedge is now τ'_{Ysi} (Equation 3.15) instead of τ_{Ysi} (in Equation 3.8). This is a counterfactual wedge that encompasses the initial HK wedge (a wedge I will refer to as the structural or HK wedge) and τ^z_{Ysi}, which I will define as the monetary wedge. In other words, τ'_{Ysi} represents a corrected estimate of the "true" output wedge faced by the firm in economies subject to currency misalignment. Similarly, τ'_{Xsi} is the counterfactual material input wedge and, as before, it contains the joint effect of τ_{Xsi} and τ^z_{Xsi}. Finally, (again) in the absence of misalignment, or equivalently in a closed economy, τ^z_{Ysi} and τ^z_{Xsi} would be equal to zero, and, thus, the resulting wedges would strictly equal the HK wedges. This result should not be surprising, as without currency misalignment, misallocation is fully captured by such wedges.

Let's now turn to discuss how to infer $1-\tau'_{Ysi}$ and $1-\tau'_{Xsi}$. These are wedges that correspond to counterfactual revenues, $\overline{p}^z_{si}Y_{si}$, and counterfactual material costs, $\overline{m}^z_sX_{si}$; which in turn depend on unobservable prices $\overline{p}^z_{si}$ and $\overline{m}^z_s$, respectively. The strategy I follow here to infer such wedges is to simulate $\overline{p}^z_{si}Y_{si}$ and $\overline{m}^z_sX_{si}$ (not $\overline{p}^z_{si}$ and $\overline{m}^z_s$!) using Y, X, ϕ, and ψ as constants and allowing the exchange rate to vary from e to z. More specifically:

$$\overline{p}^z_{si}Y_{si} = 1+\phi^d_{si}(z-e)\overline{p}^e_{si}Y_{si} \tag{3.17}$$
$$\overline{m}^z_sX_{si} = 1+\psi^d_{si}(z-e)\overline{m}^e_sX_{si} \tag{3.18}$$

Once $\overline{p}^z_{si}Y_{si}$ and $\overline{m}^z_sX_{si}$ are simulated through Equations (3.17) and (3.18), the rest of the strategy consists of solving the model using the counterfactual equations in the same way as HK. In other words, the whole identification strategy simply consists of applying HK to simulated data that reflects the effect of exchange rate misalignment on misallocation.

The solution to the HK model is widely known. To solve the counterfactual model, I followed Dias et al. (2016), who solves the HK model by adding materials, a key input in my model. The firm's counterfactual output price is derived from profit maximization as a fixed markup over marginal cost:

$$\overline{p}_{si}^{z} = \frac{\sigma}{\sigma-1}\Psi_s^z \frac{(1+\tau_{Lsi})^{\beta_s}(1+\tau_{Xsi}^{'})^{1-\alpha_s-\beta_s}}{(1-\tau_{Ysi}^{'})A_{si}^{z}}, \tag{3.19}$$

where

$$\Psi_s^z = \left[\frac{r_s}{\alpha_s}\right]^{\alpha_s}\left[\frac{w_s}{\beta_s}\right]^{\beta_s}\left[\frac{\overline{m}_s^z}{(1-\alpha_s-\beta_s)}\right]^{1-\alpha_s-\beta_s}$$

The superscript z indicates that the variable depends on the simulated value at exchange rate z.

Counterfactual wedges are given by:

$$(1+\tau_{Xsi}^{'}) = \frac{1-\alpha_s-\beta_s}{\alpha_s}\frac{R_s K_{si}}{\overline{m}_s^z X_{si}} \tag{3.20}$$

$$(1+\tau_{Lsi}) = \frac{\beta_s}{\alpha_s}\frac{R_s K_{si}}{W_s L_{si}} \tag{3.21}$$

$$(1-\tau_{Ysi}^{'}) = \frac{\sigma}{\sigma-1}\frac{1}{\alpha_s}\frac{R_s K_{si}}{\overline{p}_{si}^z Y_{si}} \tag{3.22}$$

As is standard in this framework, the presence of distortions is inferred through Equations (3.20) to (3.22), based on information about gross output, input costs, and elasticities. Simulated values of material costs and revenues allow the inference of $(1+\tau_{Xsi}^{'})$ and $(1-\tau_{Ysi}^{'})$. However, it is important to note that the labor distortion, $(1+\tau_{Lsi})$, remains unchanged in this counterfactual exercise compared to the benchmark model. This is because $(1+\tau_{Lsi})$ depends on the capital-to-labor share, which is entirely valued in pesos due to the assumption of no free mobility of capital and labor.

Total factor quantity productivity (TFPQ) is distinguished from total factor revenue productivity (TFPR) as follows:[5]

$$TFPQ_{si}^z = A_{si}^z = \frac{Y_{si}^z}{K_{si}^{\alpha_s} L_{si}^{\beta_s} X_{si}^{(1-\alpha_s-\beta_s)}}, \tag{3.23}$$

[5] TFPQ (TFPR) is defined as the units of output (revenue) obtained by a firm when using one unit of a mix of factors of production.

While, total factor revenue productivity is given by:

$$\begin{aligned} TFPR^z_{si} = \overline{p}^z_{si} A^z_{si} &= \frac{\overline{p}^z_{si} Y_{si}}{K^{\alpha_s}_{si} L^{\beta_s}_{si} X^{1-\alpha_s-\beta_s}_{si}} \\ &= \frac{\sigma}{\sigma - 1} \Psi^z_s \frac{(1+\tau_{Lsi})^{\beta_s}(1+\tau^{'}_{Xsi})^{1-\alpha_s-\beta_s}}{1-\tau^{'}_{Ysi}} \end{aligned} \tag{3.24}$$

Equation (3.24) demonstrates that TFPR varies across firms within the same industry only when firms face distortions, which is a standard outcome in this conceptual framework. However, two distinct situations must be considered.

Similarly to the HK framework, a high $TFPR^z$ indicates that the firm is facing barriers, while a low $TFPR^z$ suggests that it benefits from subsidies (these barriers and benefits include structural and monetary factors). However, the most interesting interpretation arises when comparing $TFPR^z$ with $TFPR$ in the benchmark model. When $TFPR^z$ is greater than $TFPR$, it means that the firm is effectively paying a "currency tax" for holding dollars due to the misaligned exchange rate. If this tax did not exist (i.e., if the exchange rate gap were zero), the firm's marginal revenues would be higher. Exporting firms are a clear example. Conversely, when $TFPR^z$ is lower than $TFPR$, it means that the firm has benefited from currency subsidies. In other words, if the exchange rate misalignment did not exist, the firm would not have had access to the same imported inputs enabled by an artificially cheap dollar. Importing firms serve as a prime example of this situation.

To determine the welfare cost of misalignment on TFP, in the following way. Let $TFPR^{z*}_s$ denote the efficient level of total factor revenue productivity common to all firms in industry S. The definition of $TFPR^{z*}_s$ that I adopt here is the one suggested by Dias et al. (2016), which is the $TFPR$ that would be obtained if all firms faced the same average wedge.[6] Thus, it can be shown that:

$$TFPR^{z*}_s = \left(\frac{\sum^{M_s}_{i=1} (A^z_{si})^{\sigma-1}}{K^{\alpha_s}_s L^{\beta_s}_s X^{1-\alpha_s-\beta_s}_s} \right)^{\frac{1}{\sigma}} \tag{3.25}$$

Equation (3.25) is then used to obtain the real gross-output gains in each industry as follows:

[6] Doing so ensures that factor demand in each industry remains the same before and after the hypothetical factor reallocation exercise.

$$\frac{Y_s^{z*}}{Y_s^z} = \left[\frac{\sum_{i=1}^{M_s} (A_{si}^z)^{\sigma-1}}{\sum_{s=1}^{M_s} (A_{si}^z \frac{TFPR_s^{z*}}{TFPR_{si}^z})^{\sigma-1}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{3.26}$$

which in turn is used to get an aggregate effect through the Cobb-Douglas aggregator given by Equation (3.1), and thus:

$$\frac{Y^{z*}}{Y^z} = \prod_{s=1}^{S} \left(\frac{Y_s^{z*}}{Y_s^z}\right)^{\theta_s^z} \tag{3.27}$$

In this setting, the gross-output potential gains—given by Equation 3.27—coincide with the potential TFP gains. However, because value added is a more closely related measure of welfare, the gross-output gains translate into value-added gains as follows:

$$\frac{V^{z*}}{V^z} = \frac{\left(\frac{Y^{z*}}{Y^z}\right) - q^z}{1 - q^z}, \tag{3.28}$$

where $q^z = \frac{\overline{m}^z X}{\overline{P}^z Y}$, which represents the total economy intermediate input share.

Finally, having a counterfactual model opens up the possibility of using HK as a "direct approach," a potential contribution of this paper. In general, HK is classified as an "indirect approach" in the sense that it allows the measurement of misallocation without any association to an underlying specific cause of it (Restuccia and Rogerson 2017, 155). However, the difference between the value-added gains in the counterfactual model and the baseline model is attributable to the misalignment in exchange rates, resulting in:

$$\delta^V = \frac{V^{z*}}{V^z} - \frac{V^{e*}}{V^e}, \tag{3.29}$$

where δ^V represents the welfare loss caused by the misalignment. Equation (3.29) allows us to decompose value-added gains into two components: one that I call "structural real" and another that I call "structural monetary." The former captures HK's value-added gains, while the latter captures the gains caused by exchange rate misalignment.

3.3 Why Cuba? Background and Data

For those interested in studying resource misallocation, Cuba offers a compelling case. In the past, it was a typical centrally planned economy, characterized by large state-owned enterprises (SOEs) operating in an environment riddled with incentive problems. Today, it is a mixed economy with an emerging private sector, yet it remains marked by underdeveloped market institutions, excessive government intervention, and a significant presence of SOEs. Additionally, it is a country subject to U.S. sanctions that hinder productive and commercial specialization. More importantly, for the purposes of this study, Cuba faces severe macroeconomic imbalances—where until very recently, it operated a highly peculiar monetary system with two domestic currencies, administrative controls, and multiple exchange rates.

To understand the foundations of this system, it is necessary to go back to the 1990s, when Cuba experienced a profound economic crisis (1990–1993) triggered by the collapse of the Soviet Union and Eastern Europe. In the context of isolation from international financial markets, the loss of Cuba's main trading and financial partner led to a severe external financing crisis. Official figures indicate that between 1990 and 1993, Cuba's bilateral trade with COMECON declined by 68 percentage points, resulting in a 45% deterioration in the terms of trade (CEPAL 1999, 150). Fiscal revenues plummeted, mainly due to the sizable losses of SOEs, many of which virtually ceased operations. However, fiscal expenditures were not adjusted accordingly to prevent massive unemployment and an extremely unequal distribution of the crisis. GDP contracted by 34.8%, and the fiscal deficit averaged 24.9% from 1990 to 1993 (Vidal 2012, 40). With no clear source of funding, the fiscal deficit was ultimately monetized.

The resulting monetary imbalances from the described situation were not fully reflected in (official) inflation figures due to widespread price controls that kept prices at pre-crisis levels (CEPAL 1999, 133). Output and material input prices, the nominal exchange rate, interest rates, and wages remained largely unchanged—made possible by the predominance of SOEs across all sectors of the economy, including production, commerce, and finance. Since nominal variables were not adjusted, shortages became widespread. Rationing extended across the economy, and black markets for goods and services flourished, reflected in an inflation rate of 1,550% in these markets by 1993 (CEPAL 1999, 134).

Soaring inflation led to a process of partial dollarization, while the very limited availability of foreign currency was increasingly managed through exchange controls. With the exchange rate remaining fixed, an unprecedented expansion of the parallel foreign exchange market took place. The official exchange rate maintained parity with the U.S. dollar, while the parallel rate gradually soared, reaching a peak of 130 CUP per U.S. dollar (CEPAL 1999, 130). This resulted in a dual monetary system, with two currencies—the U.S. dollar (illegal at the time) and the Cuban peso—and two exchange rates, the official and the parallel (also illegal).

Segmentation between the official market and the black market—characterized by a highly divergent price structure—became the new normal. SOEs operated with regulated prices anchored to pre-crisis levels, which failed to reflect relative scarcities and, as a result, led to severe shortages of raw materials, fuel, and energy. Meanwhile, a thriving black market emerged, with prices that more accurately reflected the initial shock and the monetary imbalances stemming from the failure to adjust fiscal expenditures.

In 1994, the government responded to this new reality by implementing structural reforms across various areas. The country opened up to foreign investment, new private cooperatives emerged in the agricultural sector, private self-employment expanded significantly, and SOEs were granted much greater autonomy in decision-making (see CEPAL 1999 for further details). A macroeconomic stabilization plan was also introduced. As part of this plan, the government institutionalized the parallel foreign exchange market and legalized the circulation of U.S. dollars, while simultaneously introducing a new domestic currency, the Cuban Convertible Peso (CUC), which was pegged one-to-one to the U.S. dollar. This monetary model became known as the Dual Monetary System (DMS hereinafter).

As Doimeadios and Hidalgo, 2011, 67-73 point out, the DMS was built upon three key components: (1) partial dollarization, (2) multiple exchange rates, and (3) market segmentation.[7] Regarding the first component, partial dollarization, many state entities were authorized to operate in both currencies, US dollars (or CUC) and Cuban pesos (CUP). However, transactions carried out in dollars (or CUC) were converted to CUP at the official one-to-one exchange rate. Eventually, all economic agents became part of this mechanism, including cooperatives, foreign companies,[8] and all entities of the public sector.[9] Foreign companies operated directly in U.S. dollars until at least 2004, after which they transitioned to using CUC, which gradually began to lose its convertibility to the dollar. Meanwhile, SOEs and cooperatives relied on

[7] In this research, I primarily focus on explaining the micro foundations of the system, with particular emphasis on price segmentation. However, the literature provides a comprehensive explanation of its macroeconomic fundamentals, its evolution over time, and how the system successfully managed to ensure macroeconomic stability (see, particularly: **?**, 67-107).

[8] Foreign companies operated in USD; however, when paying wages or other expenses in CUP, the conversion was carried out at the official exchange rate.

[9] Public sector entities transacted in CUP; however, when they needed to import goods, they exchanged CUP for USD at the official exchange rate.

foreign exchange allocations from the central planner, which were required to be exchanged for Cuban pesos at the official rate.

Regarding the second component, the system was characterized by two exchange rates with a significant spread between them. (Doimeadios and Hidalgo, 2011, 67-73; Vidal 2012; de la Torre and Ize 2014, 8-9; Lage 2016, 49-50; Hernández-Catá et al. 2017; Di Bella et al. 2020). On the one hand, as previously explained, all dollar (or CUC) transactions within the corporate sector took place at the official exchange rate. At this overvalued rate, the CUP was not convertible, and the resulting excess demand for dollars was managed through exchange controls. At the same time, the newly introduced currency, the CUC, was pegged to the U.S. dollar at a 1:1 rate, ensuring convertibility—at least until 2008.

On the other hand, any dollar (CUC) transactions by households took place at the parallel exchange rate, which was initially legalized under a dirty float regime. This exchange rate was commonly known as CADECA's exchange rate (the Spanish acronym for Foreign Exchange Houses). Although, in the beginning, the CADECA exchange rate fluctuated above 100 CUP per US dollar, over time, its value stabilized, and in 2005, the government decided to fix it at 24 CUP per US dollar. In addition to households, the emerging private sector (self-employed workers) also operated at this exchange rate.

Segmentation was a key component of the system as it effectively prevented the convergence of the two exchange rates. Companies—whether state or foreign-owned—were prohibited from accessing CADECA's foreign exchange market to buy or sell foreign currency. Similarly, private self-employed individuals were not allowed to buy or sell US dollars at the official rate. This segmentation was strictly enforced through the state-controlled financial system. I will use segmentation as the cornerstone of my identification strategy. By preventing the official corporate sector—comprising SOEs, foreign private companies, and cooperatives—from transacting at the parallel market rate, I can ensure that firms' balance sheets are not distorted by arbitrage operations between the parallel and official foreign exchange markets.

Based on these three characteristics, a key conclusion can be drawn. Unlike other cases where multiple exchange rates were formally implemented to separate current and capital account transactions (see, for example, Marion, 1994), in Cuba, the multiple exchange rate system was used to segment dollar transactions between producers and consumers. Table (3.1) illustrates this conclusion.

The literature has consistently suggested that the most prominent consequence of the DMS was the distortion it created in relative prices. Due to

Table 3.1: Multiple exchange rates under the Dual Monetary System

	USD to CUC [CUC per US. dollar]	**CUC to CUP** [CUP per CUC]
Producers	1	1
Consumers	1	24

Note: The table omits a number of details: i) there was a difference between the buying and selling rates, ii) exchange rate was fixed at 24 CUP:CUC for consumers only after March 2005; before there existed a dirty floating rate, iii) additional 10% tax was levied on the purchase of CUC and CUP with U.S. dollars in the form of cash. None of those implied loss of generality to current research.

segmentation, the exchange rate spread between producers and consumers ultimately translated into a gap between producer and consumer prices.

To better understand this, consider the producer price in dollars converted to pesos as $p_p^d = pe$, while the consumer price in dollars (also converted to pesos) as $p_c^d = pz$, where e and z represent the official and parallel market exchange rates, respectively. It follows that, for the market to clear, $p_p^d = (z-e)p = p_c^d$, meaning that consumers paid a significantly higher price for goods in dollars than what producers actually received. In essence, the exchange rate spread functioned as a tax on corporate income. However, this is only part of the story. Exchange rate segmentation also led to considerable price distortions in the input market. Let $m_p^d = me$ denote the price in dollars converted to pesos that producers paid for raw materials, and $m_s^d = mz$ the price charged by international suppliers. Following the same logic, $m_p^d = -(z-e)m = m_s^d$.[10] In this case, the exchange rate spread effectively acted as a foreign exchange subsidy for producers' inputs.

These taxes and subsidies on companies remain hidden behind the overvalued currency. Corporate balance sheets were reported in Cuban pesos at the official exchange rate (1 CUP: 1 USD), artificially making each Cuban peso of revenue or cost equal in value to each U.S. dollar of revenue or cost. To some extent, companies were able to conceal their tendency to import raw materials in dollars—made cheaper by the overvalued exchange rate—by generating revenues in pesos, which were also overvalued for the same reason.

These exchange rate distortions created heterogeneous transfers that depended on the specific characteristics of each company and, in particular, on

[10] I am assuming that the CADECA rate reflects the economy's equilibrium rate, which may not necessarily hold true due to segmentation. It should be noted that producer segments operated under exchange controls. Nonetheless, this assumption does not affect the generality of the argument, as the applicable rate would always be higher than the official rate.

their access to foreign currency. However, given the substantial 24:1 gap between the official and parallel exchange rates, one would expect significant distortions in relative prices and profitability, leading to a considerable misallocation of resources. To the best of my knowledge, Cuba's exchange rate differential is among the highest—if not the highest—globally. Table 3.2 provides supporting evidence in this regard.

Table 3.2: Spread Between Formal and Informal Exchange Rates (in Percent)

Country and Time Period	**Mean**	**Min**	**Max**
Belgium, 1971-1974	0.04	-2.05	1.27
France, 1971-1974	-0.46	-5.54	5.03
Italy, 1973-1974	3.90	0.76	7.48
Argentina, 1981	44.3	32.8	57.2
Bolivia, 1982	289.8	84.7	524
Dominican Republic, 1982	18.8	-4.59	76.5
Guatemala, 1984-1988	80.5	0.4	280
Mexico, 1982-1988	15.1	-1.26	108.8
Cuba, 2005-2021	2,300	2,300	2,300

Note: Spread is defined as $\left(\frac{P-O}{O}\right) \times 100$, where P is the parallel rate and O is the official rate. All spread values except those for Cuba were taken from Marion, 1994, 219.

3.3.1 Data

The National Bureau of Statistics and Information (ONEI, by its Spanish acronym) provided access to non-public data for Cuban manufacturing firms in 2015. The data are collected annually by the bureau through a series of administrative registers and cover all enterprises. These registers contain detailed accounting and financial information derived from corporate balance sheets. In this research, I have utilized three different registers: *Modelo-5903*, titled *"Cumplimiento del Plan Económico"* (see ONE 2009, 6 for an overview), *Modelo-5920*, titled *"Proforma EFE 5920 – 02 Estado de Situación"* (ONEI, 2017, 4-8), and the *"REEUP – Registro Estatal de Empresas y Unidades Presupuestadas"* (ONE, 2009, 4).

Consistent with the fact that the Cuban manufacturing industry has been predominantly composed of large state-owned enterprises, the statistical units in the data are mainly state-owned entities (i.e., SOEs and other types of public companies). Private firms, more specifically, small and medium-sized enterprises that are commonly found today, did not exist in 2015.[11] On the other hand, the existing private self-employment was primarily concentrated in service activities rather than manufacturing. The industry also included, albeit on a much smaller scale, private non-agricultural cooperatives and foreign companies. According to official figures, there were 49 private cooperatives in manufacturing in 2015 (ONEI, 2016, Table 4.3); however, they were not included in the database provided by the bureau.

In contrast, 42 foreign companies were included in the database; however, they had to be removed due to missing data on the number of employees and wages. Although the statistical yearbooks do not provide official public figures on the number of foreign companies, the REEUP database identifies 56 firms in the sector, most of which were joint ventures. Additionally, entities whose primary activity was administrative rather than productive were excluded from the dataset.

The information extracted from administrative records includes the firm's age (based on its date of creation), net sales, total expenses, gross output, value added, consumption of intermediate inputs (such as expenses on raw materials, fuels, and energy), labor costs (primarily wages), employment (measured as the average number of employees), and capital stock (book value of fixed capital net of depreciation). Most of the accounting items are denominated in Cuban pesos (CUP) at the official exchange rate (1 CUP:1 USD). However, for sales and total expenses, it is possible to distinguish the portion valued in Cuban Convertible Pesos (CUC). Therefore, hereinafter, I will refer to dollars and Cuban convertible pesos interchangeably unless otherwise indicated.

The dataset also includes information on the primary industry of operation of the firms, based on the Cuban Nomenclator of Economic Activities (NAE). This classifier is essentially Cuba's adapted version of ISIC, Rev. 3. According to the 4-digit classification provided by this nomenclator, the sector comprises 114 different industries. A common practice in this framework is to impute factor shares at the industry level using the average factor shares from the corresponding U.S. industries. The U.S. data are sourced from the 2015 NBER-CES Manufacturing Industry Database, from which I obtained the parameters α_s, β_s, and $1-\alpha_s-\beta_s$. To match the Cuban manufacturing industries with their U.S. counterparts, I utilized correspondence tables and excluded Cuban industries that did not have a close equivalent in the U.S. industry.

According to official aggregate figures in the Yearbooks, there were 368 state-owned entities in the Cuban manufacturing industry in 2015 (ONEI, 2016, Ta-

[11] These firms were only established in Cuba after September 2021.

ble 4.3). In general, manufacturing firms (not plants) in Cuba are large, with approximately 1,000 workers each.[12] After merging the three administrative registers into a single database and trimming 1% from both sides of the productivity distribution to ensure robustness against outliers, the final dataset contained 317 SOEs, representing 86% of the total number of state-owned entities. On average, each firm has 904 workers,[13] with the number of employees ranging from 18 to 9,005 and a standard deviation of 1,124.

The 317 companies were grouped into 72 four-digit industry NAE codes. One drawback of the data is that there is only one firm in 32 of these industries. Since the Hsieh and Klenow (HK) methodology relies on intra-industry variability to estimate misallocation effects, it will not be possible to obtain estimates for these 32 industries. An alternative approach would have been to aggregate firms into three-digit NAE industries; however, this would increase estimation bias due to technological heterogeneity. Nonetheless, even if it is not possible to obtain a baseline estimate for misallocation in these industries, the counterfactual model will at least provide insights into the direction in which exchange rate misalignment affects firms within these industries.

3.4 Empirical Results

In this section, I present the results on the reallocation gains for gross output and value added, derived from both the baseline model and the counterfactual model. The results for the baseline model correspond to those obtained by applying HK to data valued at the economy's official exchange rate. These results are denoted with the suffix e. Conversely, the results for the counterfactual model are based on applying HK to simulated data generated using Equations 3.17 and 3.18, which incorporate exchange rate distortions. Specifically, I set $z = 24$ and $e = 1$, representing the parallel and official exchange rates, respectively, with the suffix z referring to these values.

From the data, I imputed the parameters ϕ^d and ψ^d by calculating two distinct ratios. First, ϕ^d was estimated by dividing U.S. dollar sales by the sum of U.S. dollar sales and Cuban peso sales at the official exchange rate, i.e., $\phi^d = \frac{\text{USD sales}}{\text{USD sales+CUP sales}}$. I then multiplied this ratio by the gross output to sep-

[12] This number was imputed by dividing the total number of employees in manufacturing ONEI 2016, Table 7.4 by the total number of companies, which exceeds the number of state-owned entities as it includes all types of organizations ONEI 2016, Table 4.3.

[13] This figure is not strictly comparable to the previous one, as the database records the variable as "average headcount" instead of the total number of employees reported in the yearbook.

arately impute the portion sold in U.S. dollars and the portion sold in Cuban pesos. Similarly, ψ^d was estimated by dividing total expenditure in U.S. dollars by the sum of total expenditure in U.S. dollars and total expenditure in Cuban pesos at the official rate, i.e., $\psi^d = \frac{\text{USD total expenditure}}{\text{USD total expenditure+CUP total expenditure}}$. Assuming this ratio is proportional to the ratio of intermediate consumption in U.S. dollars to intermediate consumption in Cuban pesos, I then multiplied ψ^d by intermediate consumption in CUP to derive its U.S. dollar and Cuban peso components.

The elasticity of substitution between firms' gross output, σ, is a key parameter in this framework. Consistent with other studies, I set $\sigma = 3$ for both the baseline and counterfactual models. The value of this parameter has significant implications for the magnitude of the results. Specifically, liberalization gains increase with σ, making this a conservative choice.

Also consistent with other studies (see, for example, Dias et al. 2016, 18), I use the wage bill paid by firms to measure labor input. This approach implicitly assumes that wages per worker are adjusted for differences in both hours worked per worker and workers' skills. Consequently, this implies that $L_{si} = w_{si}H_{si}$ and $w_s = 1$, where H_{si} represents employment and w_{si} denotes the firm-specific average wage rate.

For the rental price of capital, I used the same value across all industries, setting $R_s = R = 10\%$. Regarding intermediate inputs, I adopted a similar approach to that used for labor input, assuming that the price of intermediate products, $\overline{m}_s^j$ (with $j = e, z$), equals 1. This ensures that intermediate input expenditure reflects not only the quantity of inputs but also their quality.

While nominal output $(P_{si}Y_{si})$ is observed in the data, firm-specific output prices are not. Therefore, to calculate the real gross output for each firm, I rely on the relationship between nominal output and real output derived from the model. From Equation 3.5, and assuming $Y_s^{\frac{1}{\sigma}}\overline{P}_s^e = 1$, it follows that:

$$Y_{si}^j = \left(\overline{p}_{si}^j Y_{si}^j\right)^{\frac{\sigma}{\sigma-1}}, \text{with}\, j = z, e \tag{3.30}$$

As is standard in this framework, Equation 3.30 is used to calculate real gross output. In other words, I infer prices relative to quantities from nominal gross output and an assumed elasticity of demand. Once real gross output has been obtained, I estimate total factor productivity at the firm level using Equation 3.23.

3.4.1 Descriptive Statistics

Figures 3.1 and 3.2 describe some statistical characteristics of ϕ^d and ψ^d in the data. First, the mean values are both low ($\overline{\phi^d} = 0.11$ and $\overline{\psi^d} = 0.13$), but their standard deviations are relatively higher ($sd(\phi^d) = 0.17$ and $sd(\psi^d) = 0.18$). This indicates that dollar transactions in the manufacturing industry generally constitute a small proportion of total transactions, but they exhibit considerable variation across firms.

Moreover, Figure 3.2 suggests a strong positive correlation between firms' dollar revenues and expenditures. I interpret this result as evidence that, in dollarized contexts, the parameters ϕ and ψ play a key role in firms' decision-making processes. Although firms could exploit overvalued CUP revenues to mask a higher propensity to import at cheaper dollar rates without affecting their accounting profits, they instead choose to maintain some balance between these two flows. This behavior can be explained by the understanding that domestic currency cannot be freely converted into foreign currency. As a result, firms protect themselves against currency instability by aligning their dollar revenues and expenses whenever possible.

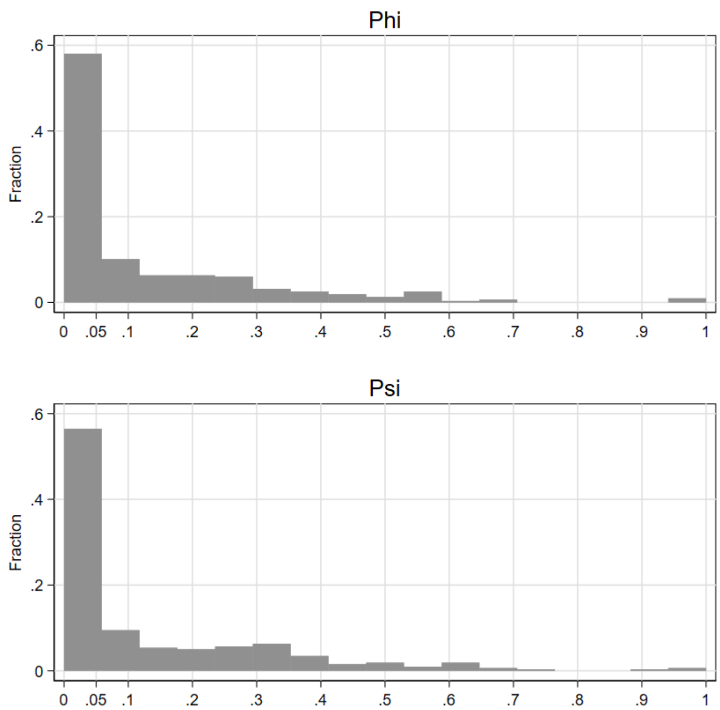
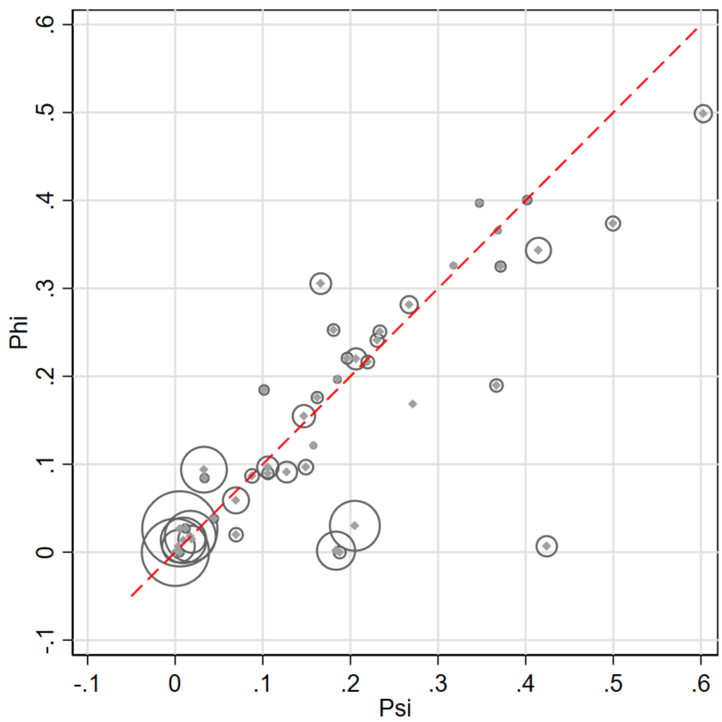


Figure 3.1: Empirical Distribution of ϕ and ψ (Histogram)

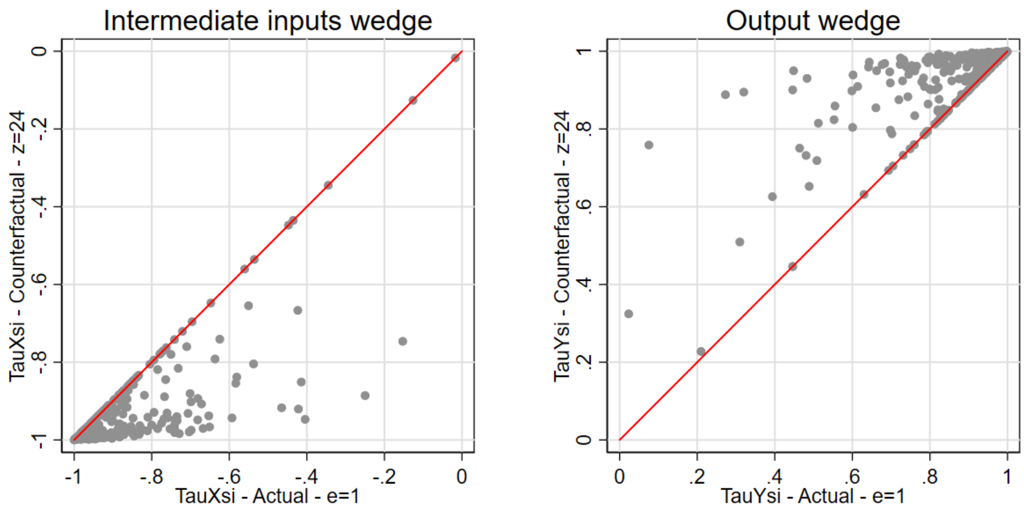


Notes: Dash line is the 45 degree line. Industries are classified according to 4-digit NAE codes. Only industries with more than one firm are shown in the graph. The size of the bubble represents the relative weight of gross output in the industry aggregate.

Figure 3.2: Mean of ϕ and ψ, by industry

On the other hand, Figure 3.3 compares the wedges of the baseline model and the counterfactual model, i.e., $(1-\tau_{si})$ and $(1-\tau_{si}^{'})$, respectively, for both output and inputs. The results are unequivocal: both wedges increase heterogeneously in the counterfactual model compared to the baseline model. As expected, the income tax rises (right panel), while the subsidy on intermediate inputs also increases (left panel).

This outcome is explained by the fact that the counterfactual model reveals τ^z for both output and inputs, which emerge from the implicit transfer system initially obscured by the overvaluation of the official exchange rate. It is also worth noting that for many firms, there is no change in wedges, as their values lie directly on the 45-degree line. This occurs because ϕ and/or ψ are equal to zero in the data, meaning that these firms do not record dollar transactions in their financial statements. I emphasize that even if the specific wedge of a given unit does not change, its relative position within the industry average does, once the wedges of other units experience a shift in their values.



Notes: The diagonal line is the 45 degree line. Each point represents the wedge of the baseline model (X-axis) and that of the counterfactual model (Y-axis). τ_Y greater than zero means a tax, while τ_X less than zero means a subsidy.

Figure 3.3: Actual Vs. counterfactual wedges

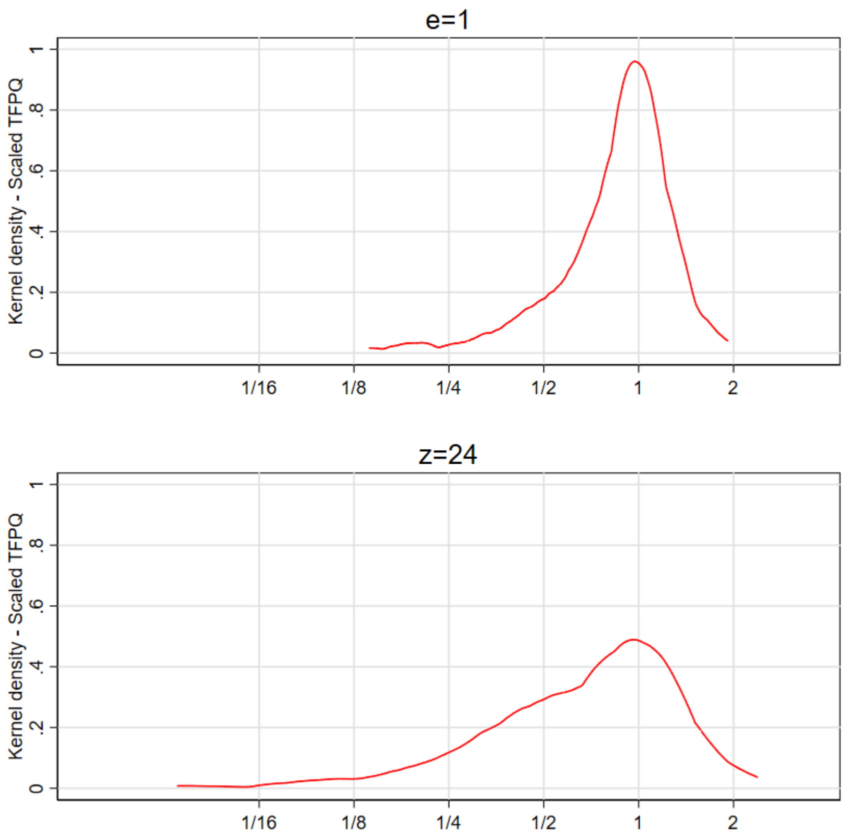
3.4.2 Productivity Distortions

I analyze the distribution of physical productivity (TFPQ) and revenue productivity (TFPR) in 2015. Figure 3.4 displays the distribution of scaled TFPQ, expressed as $\log\left(A_{si}M_s^{\frac{1}{\sigma-1}}/TFP_s^*\right)$, for the baseline and counterfactual models. Scaled TFPQ measures the deviation of firm i's TFPQ from the efficient TFP of industry s, where "efficient TFP" represents a TFP without misallocation.

Both distributions exhibit clear leftward skewness, consistent with policies that favor the survival of inefficient firms. However, the left tail of the distribution becomes noticeably thicker in the counterfactual model. This may provide evidence that foreign exchange transfers could falsely make many firms appear more productive in the original data than they actually are—or at least more productive than they would be at the official exchange rate.

The reasoning behind this interpretation is that, by not paying the real value for their dollar-denominated transactions, firms might be hiring production factors (in quantities) beyond the level their intrinsic profitability would realistically allow. In other words, an overvalued CUP enables firms to import intermediate goods at a reduced dollar cost, permitting them to acquire inputs (again, in quantities) beyond their “true” affordable level.

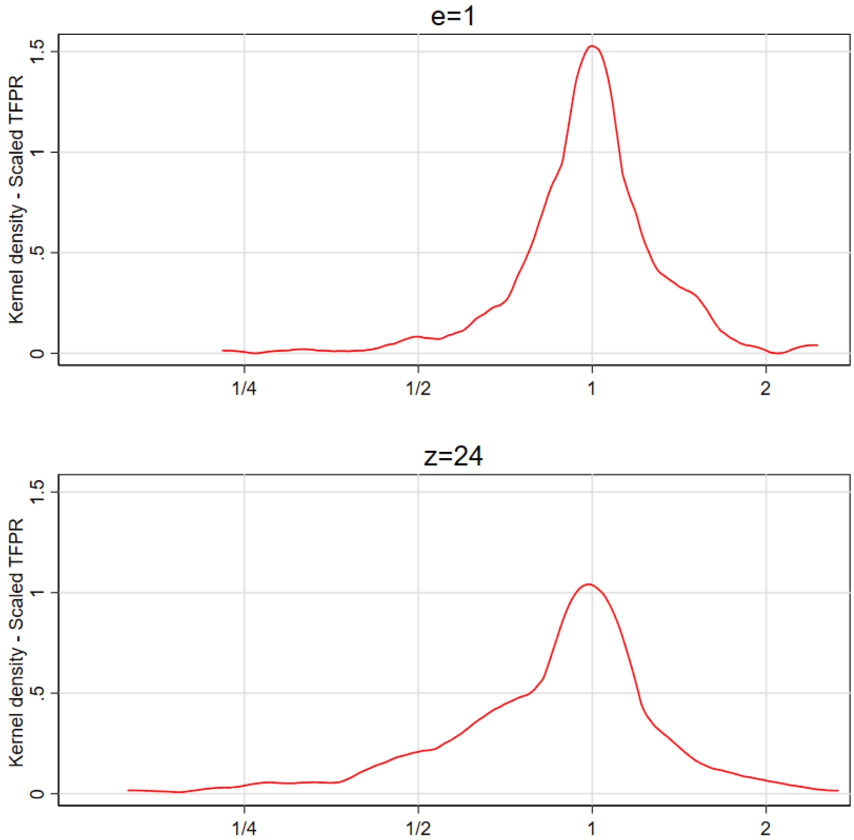
The problem is that it is unfeasible to predict exactly what that level is at the official exchange rate because, unlike other distortions, currency transfers do not appear directly in firms’ balance sheets and, consequently, are not reflected in their marginal revenue products. By correcting input and output prices in the counterfactual model, using ϕ^d and ψ^d, one can reveal how much higher that level truly is. The fact that TFPQ decreases for many firms indicates that the overvaluation of the official exchange rate grants them access to inputs they would otherwise be unable to afford if the exchange rate were higher.



Notes: Scaled TFPQ is computed as $\log\left(A^j_{si}M_s^{1/(\sigma-1)}/TFP_s^{j*}\right)$ with $j = e, z$, and then a monotonic transformation 2^x was applied. TFP_s^{j*} is defined as the efficient TFP at the industry level, given by $TFP_s^{j*} = \sum_{i=1}^{M_s}(A^j_{si})^{\sigma-1}$.

Figure 3.4: Density of TFPQ

Figure 3.5 illustrates the distribution of the scaled TFPR, expressed as $\log\left(\frac{TFPR_{si}}{TFPR_s^{j*}}\right)$. In an efficient allocation, marginal returns are equalized across firms; consequently, both the dispersion of marginal returns and TFPR would be zero. Therefore, greater TFPR dispersion serves as evidence of increased resource misallocation across firms, stemming from distortions that affect firms differently within each sector. Furthermore, a high TFPR dispersion underscores the potential productivity gains that could be realized by reallocating factors of production from firms with low marginal revenue products to those with higher marginal revenue products.

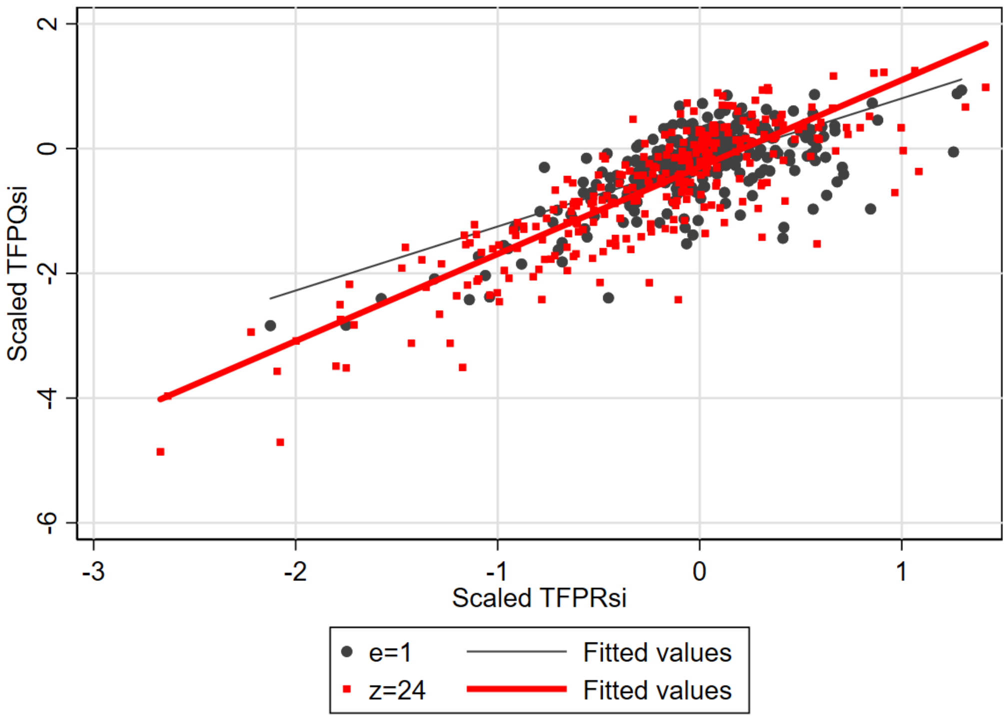


Notes: Notes: Scaled TFPR is computed as $\log\left(TFPR_{si}/TFPR_s^{j*}\right)$ with $j = e, z$, and then a monotonic transformation 2^x was applied. $TFPR_s^{j*}$ comes from Equation (3.25).

Figure 3.5: Density of TFPR

From Figure 3.5, it can be concluded that TFPR dispersion is high in both models. However, the distribution is notably wider in the counterfactual model, and, once again, the left-skewness of the distribution becomes more pronounced. I interpret this increased left-skewness as further evidence that currency subsidies outweigh currency taxes; in other words, the current over-valuation of the official exchange rate, relative to z, functions as a net subsidy for most manufacturing firms. Consequently, if an exchange rate depreciation were to take place in an effort to unify the official and parallel rates, a significant number of firms would face the removal of this subsidy—an outcome with important policy implications.

In addition, to investigate the potential consequences of resource misallocation, the correlation between TFPQ and TFPR is calculated. The literature suggests that a higher correlation implies greater losses in aggregate productivity (Dias et al., 2016). Figure 3.6 shows that TFPQ and TFPR are positively correlated in both models; however, the correlation is significantly stronger in the counterfactual model, as evidenced by the steeper slope of the fitted values (red line). This finding indicates not only that more productive firms tend to face larger HK distortions, but also that they encounter greater monetary distortions— a factor not captured in the baseline model. This implicitly suggests that less productive firms tend to overproduce, while more productive firms tend to underproduce.



Notes: Scaled TFPR is computed as $\log\left(TFPR_{si}/TFPR_s^{j*}\right)$, while scaled TFPQ is computed as $\log\left(A_{si}^{j}M_s^{1/(\sigma-1)}/TFP_s^{j*}\right)$, with $j = e, z$. Fitted values are from an OLS regression of scaled TFPQ on scaled TFPR.

Figure 3.6: Correlation Between TFPR and TFPQ: Actual vs. Counterfactual

3.4.3 Reallocation Gains

Table 3.3 presents the gross-output and value-added gains resulting from fully equalizing TFPR across firms within each industry, as per Equations 3.27 and 3.28. Two key findings emerge from Table 3.3. First, the potential gains from eliminating distortions appear modest in terms of gross output but are considerably more substantial in terms of value added, both in the baseline and counterfactual models. For instance, while gross output would increase by 19.5% in the baseline model, the corresponding value-added gains amount to 54.4%. This is a common finding in the literature, often attributed to the gross-output approach used to measure misallocation. For example, Dias et al. (2016) obtained a similar result for the Portuguese economy and specifically attributed these differences to the fact that gross output gains were computed under the assumption of constant intermediate inputs, as indicated by Equation 3.28.

Second, Table 3.3 also underscores the substantial efficiency gains in the counterfactual model. Notably, the difference compared to the baseline model is 50 percentage points higher. To put this into perspective, if HK distortions were completely eliminated—holding all else constant—manufacturing value added could increase by approximately 50%. Furthermore, if exchange rate distortions were also removed (by unifying the official and parallel exchange rates), these gains could roughly double, meaning manufacturing value added could increase by about 100%.

The results call for additional discussion. At first glance, the value-added gains associated with the counterfactual model may seem too large to be attributed to a single (monetary) factor, and even disproportionately large when compared to the gains from eliminating HK distortions. It should be recalled that, in the model, the impact of a nominal exchange rate devaluation on a real variable—such as value added—operates through the exchange rate transfer system. Therefore, aligning the official exchange rate with the parallel rate effectively removes the (heterogeneous) tax on net exporters and the subsidies granted to net importers.

On the one hand, I emphasize the significant scale of this transfer system in Cuba, where the role and scope of monetary distortions should by no means be underestimated. As shown in Table 3.2, the 2,300% exchange rate spread is among the largest—if not the largest—ever recorded in international experience. The sheer magnitude of this spread highlights the severe distortions in Cuba's monetary market and the heavy burden it imposes on economic agents. This could potentially explain why the real effect is so pronounced.

On the other hand, I acknowledge the possibility that HK distortions may be somewhat underestimated, as the quantitative exercise does not account for private firms. As mentioned in Chapter 2, the estimates should be considered a lower bound for this reason. Ultimately, resource misallocation is a relative measure, and the absence of data on private agents in the estimates

Table 3.3: Gross-Output and Value-Added Gains by Exchange Rate

Exchange Rate	Gross-Output	Value-Added
$e = 1$	19.5%	54.4%
$z = 24$	38.9%	106%

Notes: Gross-output gains are computed as $(Y^*/Y - 1) \times 100$, where Y^*/Y is derived from Equation (3.27). Value-added gains are computed as $(V^*/V - 1) \times 100$, where V^*/V is derived from Equation (3.28).

could manifest as lower observed variation compared to the true TFPR gaps across firms—a likely scenario in a context like Cuba's, characterized by market segmentation.

In any case, the absence of the private sector does not invalidate my results, which consistently highlight the pervasive role of misallocation in Cuba. In this chapter, specifically, they underscore the considerable distortions in its monetary markets.

Country and time period	Initial	Final
Argentina[1] (1997-2002)	52.2	60.0
Bolivia[1] (1988-2001)	52.5	60.6
El Salvador[1] (2004)	n.a.	55.1
Brazil[1] (2000-2005)	49.1	41.4
Chile[1] (1996-2006)	47.5	53.7
Mexico[1] (1999-2004)	140.1	109.5
United States[2] (1977-1997)	36.1	42.9
China[2] (1998-2005)	115.1	86.6
Vietnam[3] (2015)	n.a.	81.2
Cuba[4] (2015)	n.a.	54/106

Notes: The entries correspond to the TFP gains from a hypothetical exercise in which all distortions preventing the equalization of TFPR across firms are removed. In all cases, the distortions refer to HK distortions, except for Cuba, where monetary distortion is also included. [1]Busso et al. 2013; [2]Hsieh and Klenow 2009; [3]Nguyen and Nguyen 2020; [4] Table 3.3. Results not fully comparable due to methodological differences (e.g., gross output vs. value-added approach) and data differences (size of companies involved, type of ownership, etc.), among others.

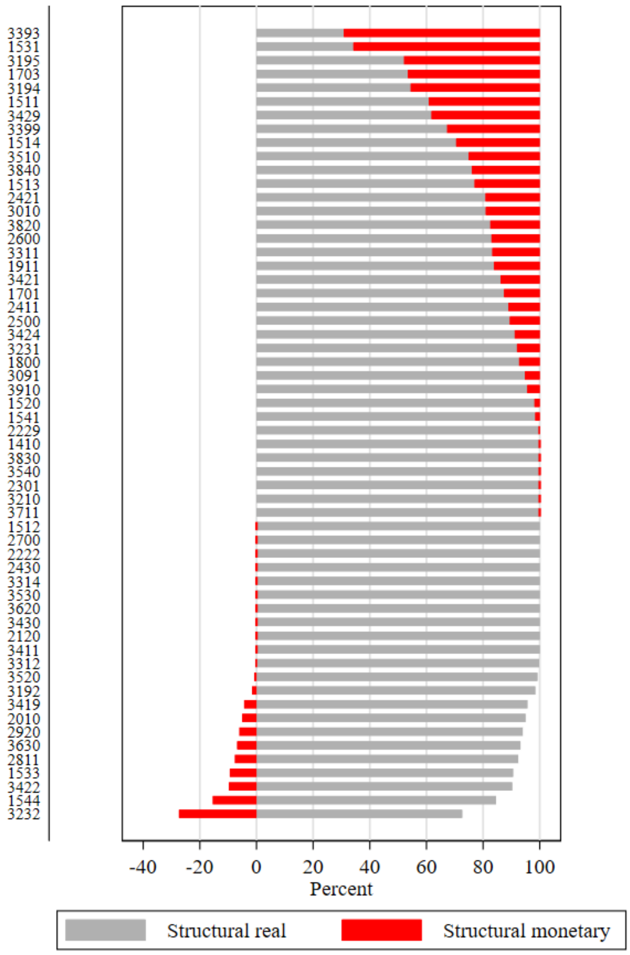
Table 3.4: Comparative TFP gains across countries (Structural wedge only)

In order to put these results into context, Table 3.4.3 presents a cross-country comparison of the TFP gains that result from equalizing TFPR across firms. In the table, TFP gains are computed for each country by removing the typical Hsieh and Klenow distortions, which I have distinguished throughout this work from monetary distortions. For the Cuban case, I included the gains obtained from the counterfactual model reported in Table 3.3.

I acknowledge that the multiple methodological differences among the referenced studies in the table clearly limit the precision of the analysis. However, the results of the comparison suggest that they should not be overlooked. According to the table, Cuba's TFP gains are relatively low (54%) compared to those of other transition economies, such as Vietnam (81%) and China (86%-115%), which suggests that misallocation imposes a lower cost in terms of welfare. This finding is difficult to justify given the overwhelming amount of suggestive and anecdotal evidence pointing to the contrary (see Chapter 1). In line with the hypothesis of this study, at least part of this discrepancy could be explained by the underestimation of the monetary component. Notably, when the exchange rate gap is taken into account for Cuba, the TFP gains become more consistent with expectations (106%) and, at the same time, are not disproportionately distant from international benchmarks. Of course, this evidence is only suggestive, and further research will be needed to reach more conclusive results.

Finally, in Subsection 3.2.2, I suggested the possibility of using the Hsieh and Klenow (2009) framework as a "direct approach " to measuring misallocation, given the nature of the problem addressed in my work. In Figure 3.7, I estimate Equation 3.29 to put into context the significance of value-added gains by industry according to each of the two distortion components (HK distortions and monetary distortions). This analysis has clear policy implications, as it helps identify which of the two drivers of misallocation is most significant in each industry. In particular, it allows us to anticipate which industries would be most severely affected by a potential depreciation of the official exchange rate.

The results suggest that the impact across industries would be heterogeneous. This implies that while the overall effect of unifying the official and parallel exchange rates would be positive—as indicated by the TFP gains reported in Table 3.3—in some industries, particularly those at the top of Figure 3.7, the consequences could be devastating. For instance, many state-owned enterprises (SOEs) could face bankruptcy and lay off workers—both of which are expected consequences of such a significant adjustment in relative prices. This effect would occur because exchange rate unification would eliminate the system of transfers created by the misalignment. Managing these effects will be crucial for the government to ensure the political viability of this measure and to prevent, among other risks, potential policy reversals.



Notes: Each bar shows the value-added gains (unweighted) under the baseline model and the counterfactual model as a percentage of the total. Each one is computed from Equation (3.29). It is negative if the gains under the counterfactual model are lower than those under the baseline model. The codes on the vertical axis represent the 4-digit NAE industry codes. Industries where the ϕ and ψ parameters were equal to zero were excluded.

Figure 3.7: Value Added by Industry and Component

3.5 Concluding Remarks

In this paper, I have used firm-level Cuban data—generated under an idiosyncratic institutional monetary arrangement—to investigate the welfare effects of exchange rate misalignment and multiple exchange rates.

I document significant potential efficiency gains associated with exchange rate unification in Cuba. In particular, when MRPs are adjusted to internalize unobserved FX transfers, revealed through a counterfactual exercise, I find that manufacturing value added doubles. This result is 50 percentage points higher than the gains obtained when MRPs are not adjusted. From this, it can be inferred that the results of studies applying the Hsieh and Klenow (HK) methodology in contexts of currency misalignment should be revisited. This raises the possibility that such studies may be underestimating the true efficiency gains derived from improving resource misallocation.

Another conclusion drawn from this paper is that the parameters ϕ^d and ψ^d contain valuable information for fully characterizing resource misallocation in contexts of exchange rate misalignment.

4 Identifying Redundant Labor in Emerging Economies

4.1 Introduction

Misallocation of resources has become an increasingly relevant framework for explaining productivity differences across countries (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2017). In market economies, misallocation typically appears as a wedge in the marginal revenue product. In these contexts, firms optimally choose the amount of capital and labor, but allocative inefficiencies arise due to heterogeneous distortions captured by such wedges (e.g., taxes, financial frictions, and others).

In contrast, in transition economies where the government plays a more dominant role, firms face a different type of distortion: the direct allocation of inputs by the state, which limits their autonomy in determining the optimal demand for factors. For instance, driven by political motivations to maintain low unemployment levels, a government may impose minimum hiring requirements on state owned enterprises (SOEs). This leads to redundant workers in companies—an example of a broader phenomenon known in the transition literature as 'policy burden' (Lin et al., 1998; Lin and Tan, 1999; Lixing, 2008; Liao et al., 2023). In a nutshell, SOEs are compelled to comply with political mandates, which impose additional costs and ultimately reduce profitability and competitiveness. Thus, beyond the standard labor wedge emphasized in the labor misallocation literature (Petrin and Sivadasan, 2013; Alpysbayeva and Vanormelingen, 2022), economies with strong state intervention face a political burden that can result in labor redundancy and greater deviations from the efficient allocation of resources.

Another characteristic of developing economies is the low quality of their financial institutions, which, among other things, implies that firms face credit constraints (Buera et al., 2015). Taken together, credit constraints and policy burdens can lead to a suboptimal choice of other inputs used by the firm. For example, to meet the state-mandated employment level, firms may be forced to reduce the quantity of non-labor inputs used in production, such as materials or capital investment.

The goal of this paper is to quantify the magnitude of labor redundancy in a transition economy where the state maintains a strong presence and credit constraints are relevant. To undertake this exercise, I employ a short-term partial equilibrium model developed by Dong and Putterman (2003), in which the state exogenously sets an employment threshold for firms. This creates a distortion (constraint) in the firm's working capital, leading to the endogenization of wages and raw materials. The main implication of the model is that firms facing redundant workers respond not only by reducing wages but also by limiting their demand for non-labor inputs or immobilizing more short-term financial resources. Furthermore, the model allows to estimate the magnitude of redundant employment and correlate it with various explanatory factors.

The model is applied to state-owned firms in Cuba's manufacturing industry from 2009 to 2013. The choice of the Cuban case was intentional, as certain institutional aspects of its economic model align with the theoretical framework of redundant employment.

On the one hand, although Cuba was once a typical centrally planned economy, it has been gradually transitioning into a mixed economy since 2011, when a significant reform process was launched to expand the private sector. However, despite these reforms, the government has maintained a strong interventionist role in the economy. For instance, while the 2011 reforms granted companies some degree of autonomy compared to the past, managers were still not allowed to adjust labor demand according to market conditions, as the government continued to regulate this decision.

On the other hand, Cuba has long been suspected of harboring redundant workers due to the persistent combination of low unemployment rates and high levels of state employment—factors that were further reinforced by the severe restrictions on private ownership and entrepreneurship prior to 2011. To illustrate, between 2005 and 2009, the average annual unemployment rate stood at just 1.8% (ONEI, 2023), while in 2009, the public sector employed 82.4% of the country's total labor force (ONEI, 2023). Recognizing this imbalance, one of the government's first objectives at the end of 2010 was to reduce state employment by 500,000 workers throughout 2011 (Granma, 2010), a figure that accounted for 10% of the total labor force.

Finally, many of the outcomes documented in the literature on overhiring are clearly evident in Cuba, including low wages, stagnant labor productivity, widespread shortages of raw materials, idle capacity within firms, the existence of loss-making enterprises, substantial budget subsidies, and chains of unpaid debts. However, this does not imply that overstaffing is the sole issue in this transitional context, where incentive misalignments, soft budget constraints, regulated prices, and other structural challenges also persist.

The results of this study show that in 2009, 93% of manufacturing SOEs in Cuba were affected by labor redundancy, with an estimated 66% of total state employment in this industry considered redundant. These figures significantly exceed those documented in similar studies conducted in other transition economies, where labor redundancy rates have been estimated to range from 20% to 50% (World Bank, 1996; Sheehan et al., 2000; Belser and Rama, 2001). However, by 2013, a significant reduction of 15 percentage points in the rate of redundant employment was observed, a change attributed to a combination of rising prices, increased availability of raw materials, and productivity growth during the period. Additionally, the easing of financial constraints also contributed, albeit to a lesser extent. The findings suggest that labor redundancy remains a persistent challenge for the Cuban economy, but market-oriented reforms aimed at enhancing productivity and strengthening financial institutions could help mitigate it.

Related literature This paper relates to several strands of the literature. First, it draws on the literature that examines the economic and financial distortions created by political burdens in general, and the excess of workers in SOEs in particular, as discussed in studies such as Dong and Putterman (2003); Lin and Li (2008); Liao et al. (2023); Talavera et al. (2024). My work is most closely related to Dong and Putterman (2003), who analyze data from Chinese state-owned manufacturing firms to investigate the causes behind the rise in labor redundancy in the early 1990s. They conclude that binding constraints on downward employment adjustment, increased competition among firms, and growing unavailability of financial resources were the main drivers of this rise. There are two key differences between my study and theirs. Theoretically, I introduce a productivity term to account for the influence of this variable on the rate of redundant employment. Empirically, I employ a summary measure of working capital to capture all possible financial strategies that firms might use in the short term to mitigate the effects of excess labor.

Secondly, this paper contributes to the relatively underexplored case of Cuba within the literature that examines the effects of redundant employment in different transition experiences, such as those of China (Sheehan et al., 2000; World Bank, 1996; Dong and Xu, 2009), Vietnam (Belser and Rama, 2001), and other countries (Svejnar and Terrell, 1991). The Cuban case is particularly interesting not only because it represents a late transition experience but also because, unlike the aforementioned examples, it followed a much more moderate path of economic opening, with far less penetration of market institutions. Specifically, despite the notable expansion of the private sector, SOEs remained largely shielded from competition, there was no unification of the price system, privatizations did not take place, and enterprise autonomy was far more restricted compared to similar transition experiences.

The paper is structured as follows. After this brief introduction, Section 4.2 describes the Cuban context and the main changes that have taken place since 2011. Section 4.3 introduces a simple model of redundant employment, while Section 4.4 explains the empirical strategy and the data used for its measurement. Section 4.5 discusses the results, and finally, Section 4.6 provides the conclusion.

4.2 The Cuban Economy: Context and Institutional Setting

In this section, I provide a brief overview of the key institutional features of Cuba's economic model. The aim is to show why Cuba is an ideal case for studying labor redundancy, especially after the 2011 reforms, which led to a mixed or transitional economy where a growing private sector coexists with a State that still plays a major role in the economy.

For many years, Cuba adopted the typical institutions of a centrally planned system, dominated by state ownership, with limited market participation in resource allocation. In 2009, for instance, 82.4% of the total workforce was employed in state entities ONEI, 2023. The small amount of private investment was mostly concentrated in agriculture and certain service activities through self-employment. Private activity remained heavily regulated, subject to operational restrictions that significantly limited its ability to expand and grow. On the other hand, SOEs had very little autonomy in decision-making. Key decisions, such as production levels, the use of inputs like labor, capital, and raw materials, and even the setting of output prices and wages, were tightly controlled by the State through central planning.

The consequences of this institutional production model are well-known. It resulted in a highly centralized business system weighed down by a range of structural issues: widespread shortages of goods and inputs, low productivity, profitability, and external competitiveness; insufficient savings and capital accumulation; underdeveloped infrastructure; outdated technology; overemployment; and extremely low wages. SOEs operated in an environment where rules were lax and constantly changing, the risk of default was high, budget subsidies were substantial, and government interventions on companies very frequent.

The resulting economic stagnation prompted the government in 2011 to launch a reform program known as the "Update of the Cuban Economic Model." The goal was to overhaul many of the institutional foundations of the centralized system, largely inspired by market-oriented reforms in China and Vietnam, with the aim of revitalizing the economy. These changes led to the emergence of a mixed economy, with a greater role for the private sector and market mechanisms, while the State remained influential in the broader economic system and, in particular, over its state-owned enterprises, which it did not privatize but sought to reform.

For example, entry barriers to private investment were removed in many markets, while the range of activities open to foreign investment was significantly expanded. Additionally, many of the legal regulations that had previously restricted private business operations were eased. At the same time, SOEs were granted more autonomy, particularly in setting wages, prices, and, for export companies, accessing foreign currency. Their corporate mandates were also broadened, allowing them to offer secondary products at less regulated prices, and the cap on decentralized investments — those outside the central plan — was raised. The government also initiated an institutional reform process for state entities, which led to the merger or dissolution of many SOEs. As a result, between 2009 and 2013, the number of state entities in the manufacturing sector was reduced by a third (ONEI, 2022b).

Additionally, one of the first announcements made as part of the "Update" process was the reduction of staff in state-owned enterprises and public administration, a policy that had been promoted since 2009 -see, for example, Hernández (2018). The workers' union acknowledged that "the surplus of positions exceeds one million people in both the public and business sectors" and agreed to cut state employment by 500,000 workers in 2011 (Granma, 2010), representing 10% of the total labor force. In fact, the parallel expansion of the private sector was seen as a way to create jobs in response to the anticipated rise in unemployment within the state sector. Figure 4.1 shows the trend in public employment at an aggregate level from 2001 to 2022, indicating that during the study period (2009-2013), state employment fell by nearly 20 percentage points.

Although it is undeniable that the "updating" of the model brought about significant structural changes compared to the pre-reform centralized system, much of the old institutional framework remained intact. For instance, State assets were neither privatized nor allowed to fail, even when companies re-

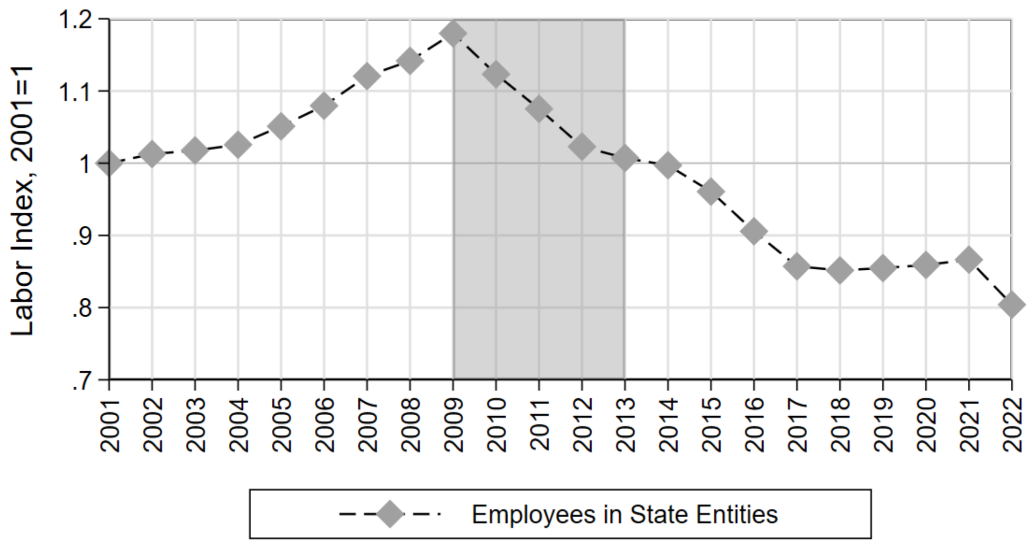


Figure 4.1: Employees in State-Owned Entities, Cuba 2001-2022
Source: Authors' calculations based on Table 4.2, ONEI, 2022

ported losses. In line with this, public budget transfers (subsidies) to SOEs steadily increased throughout the decade, rising from 12% of GDP in 2009 to 24% in 2018, the highest pre-pandemic level (ONEI, 2022c). Furthermore, the allocation of critical resources, such as imported raw materials and fuel, continued to be regulated by the central plan, and the government retained its influence over employment levels in SOEs. Finally, although price setting was decentralized across different regulatory bodies, including corporate entities (MFP, 2013b), strict oversight kept prices and wages from aligning with market levels.

Finally, in the first half of the decade—the period covered by this study—the economy showed some signs of recovery. For example, GDP grew at an average annual rate of 2.8% between 2010 and 2015 (ONEI, 2022a), while wages increased at an average rate of 8.9% (ONEI, 2022d). To what extent these results can be attributed to the reforms is an ongoing debate that goes beyond the scope of my work. However, this specific context is relevant to my study of redundant labor. For example, Figure 4.2 shows that the manufacturing sector performed notably well, with physical output growing at an average rate of 7% between 2010 and 2016, a factor that may have contributed to reducing levels of redundant labor in this industry.

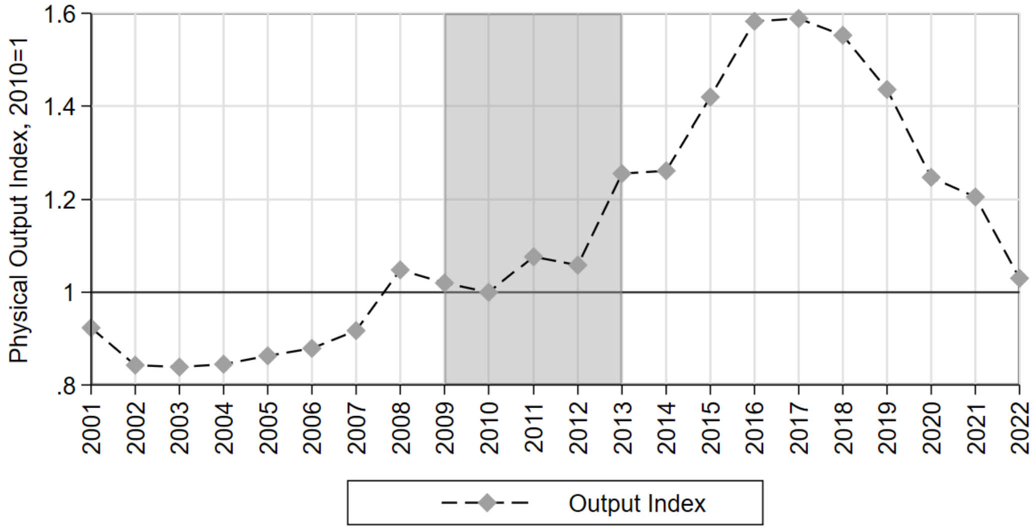


Figure 4.2: Industrial Output, Cuba 2001–2022
Source: Authors' calculations based on Table 11.1, ONEI, 2022

4.3 Theoretical Framework

In this section, I outline the model for identifying redundant labor. The model closely follows that of Dong and Putterman (2003), which is based on a two-stage framework for the firm's decision-making process, as developed in the labor literature (Manning, 1987). The model is short-term and partial equilibrium, although capital stock is controlled for in the empirical estimation. Optimal resource demand is derived from profit maximization, subject to two constraints. First, the SOE operates under an exogenously imposed binding labor constraint. Second, given this employment level, the firm must cover its expenses for non-labor inputs and wages.

In other words, the decision-making process occurs in two stages: in the first stage, employment is determined, and in the second stage, the firm selects material inputs and wages, internalizing the exogenous labor constraint and the financial constraint. Finally, optimal labor demand is determined using backward induction, based on the solutions for materials and wages in the second stage.

The inverse demand function is isoelastic and is represented by the following equation:

$$P = \sigma Y^{-\epsilon_P} \tag{4.1}$$

where P represents prices, Y the quantity, σ is a parameter capturing exogenous shifts in demand, and ϵ_P is an elasticity assumed to be constant for simplicity. I assume that firms have some market power, and consequently, their production choices inversely affect prices.

The production technology follows a Cobb-Douglas form, more specifically, an efficiency-wages augmented production function:

$$Y = AK^{\alpha_K}(EL)^{\alpha_L}M^{\alpha_M}, \tag{4.2}$$

where Y represents the physical output produced by the firm, and K, L, and M denote capital stock, labor, and materials, respectively. The parameters α_K, α_L, and α_M represent the output elasticities of capital, labor, and materials. Materials include intermediate inputs used in production, such as raw materials, fuel, and energy. The production process is influenced by two productivity factors: Hicks-neutral productivity, denoted by A, which represents Total Factor Productivity (TFP), and labor-augmenting productivity, denoted by E, which captures workers' effort and enhances the effective labor input.

Effort is endogenous and follows the function $E = (W - W^0)^\beta$, where W and W^0 are the total wage and the base wage, respectively. The base wage $W^0 > 0$ is set exogenously by the government based on a technical-professional scale, compensating for both workers' differentiated educational levels and a given (basic) effort level in the economy, $E_0 = 1$. The term $(W - W^0)$ is a flexible wage premium paid to workers to incentivize effort beyond E_0, and its value primarily depends on the firm's performance.[1] Finally, $\beta < 1$ represents the elasticity of effort with respect to the wage premium.

In the first stage, firms choose L, given the optimal selection of materials, wages, and an employment constraint, $L \geq \bar{L}$. $\bar{L}$ is the minimum number of workers set exogenously by a central, local, or sectoral government authority. Assuming that the employment constraint is binding, the firm employs $\bar{L}$ workers, resulting in redundant labor.

With the employment level determined exogenously, firms then choose M^* and W^* in the second stage by maximizing the following profit function, subject to a liquidity constraint:

$$\pi = P(Y)Y(A, E, \bar{L}, M, K) - W\bar{L} - P_M M - rK \tag{4.3}$$

[1] I choose this particular specification for simplicity, specifically to obtain linear factor demand functions. However, the quantitative result is similar if efficiency wages are modeled as a ratio between the firm's wage and the industry average, as is commonly done in the literature.

subject to the liquidity constraint:

$$W\bar{L} + P_M M = R \tag{4.4}$$

where P_M represents the price of material inputs, and r is the cost of capital. Equation 4.4 is a binding liquidity constraint, justified by the presence of redundant labor, with R specifically referring to working capital, meaning the firm's ability to cover its variable costs in the short term.

Before moving forward, I provide further motivation for the inclusion of the liquidity constraint. In a market economy, a firm automatically balances its budget constraint by selecting input levels so that the marginal product of each input equals its marginal cost. As a result, the firm's profit-maximizing decisions are not typically subject to financial constraints. However, if the firm is required to maintain an employment level, $\bar{L}$, where the marginal product of labor falls below its marginal cost, it will struggle to balance its budget unless it generates sufficient monopoly profits or receives government subsidies or preferential loans from the financial system.

Additionally, while Equation (4.4) functions as a constraint in the theoretical model, in practice, it can be eased by third-party transfers. This concept is tied to the idea of soft budget constraints, as developed by Kornai (1980). The third party could be a state bank preventing business failure through preferential loans, the central government through subsidies, or, more broadly, a lenient legal environment that allows firms to avoid paying their debts.

The firm chooses M^* and W^* by solving the following program:

$$L = P(Y)Y(A, E, \bar{L}, M, K) - W\bar{L} - P_M M - rK + \lambda(R - W\bar{L} - P_M M) \tag{4.5}$$

Equation (4.5) is the Lagrangian, and λ is the Lagrange multiplier, which measures the impact of a slight relaxation of the liquidity constraint on the firm's profits. The first-order conditions are as follows:

$$\frac{\partial L}{\partial W} = (1 - \epsilon_P)P\frac{\partial Y}{\partial E}\beta(W - W^0)^{\beta - 1} - (1 + \lambda)\bar{L} = 0 \tag{4.6}$$

$$\frac{\partial L}{\partial M} = (1 - \epsilon_P)P\frac{\partial Y}{\partial M} - (1 + \lambda)P_M = 0 \tag{4.7}$$

$$\frac{\partial L}{\partial \lambda} = R - W\bar{L} - P_M M \tag{4.8}$$

Equations (4.6) and (4.7) show that if firms face a working capital constraint, meaning $\lambda > 0$, their choice of wages and materials will fall below the optimal level where marginal revenue equals marginal cost.

By solving for M^* and W^* under working capital constraints, I find that:

$$W^* = \frac{1}{\gamma+1}\frac{R}{\bar{L}} + \frac{\gamma}{\gamma+1}W^0 \tag{4.9}$$

$$M^* = \frac{1}{P_M}\frac{\gamma}{\gamma+1}(R - W^0\bar{L}) \tag{4.10}$$

Where $\gamma = \frac{\alpha_M}{\beta\alpha_L}$ is the elasticity ratio of materials to the elasticity of effort. Equations (4.9) and (4.10) show that W^* increases with R and W^0 and decreases with $\bar{L}$, while M^* decreases with P_M, W^0, and $\bar{L}$, but increases with R.

Once W^* and M^* are determined in the second stage, I use backward induction to derive the optimal employment level, L^*, and consequently, the level of redundant employment. Since I assume that current employment is exogenously set and constrained at $\bar{L}$, the backward induction impacts employment redundancy through its effect on L^*. Given the short-term nature of my model, the optimal number of workers is derived from the following optimization problem:

$$\max \pi = P(Y)Y(A, E, \bar{L}, M^*, K) - W^*L \tag{4.11}$$

The first-order condition is:

$$(1-\epsilon_P)P\frac{\partial Y}{\partial L} = W^* \tag{4.12}$$

The left-hand side of Equation (4.12) represents the marginal revenue product of labor (MRPL). According to the isoelastic demand function, ϵ_P — the inverse of the price elasticity of demand — regulates the firm's market power. All else being equal, a decrease in ϵ_P leads to a horizontal expansion of the marginal revenue product of labor for any given value of L, and thus, an increase in L^*. Likewise, higher prices, technological change, or greater availability of raw materials will expand L^*. Conversely, an increase in wages, W^*, reduces L^*.

The redundant employment rate is defined as:

$$LRR = \ln \bar{L} - \ln L^*$$

where ln stands for the natural logarithm. Given that the production technology follows a Cobb-Douglas form, I get:

$$\begin{aligned} \ln L^* = \frac{1}{1-\alpha_L}\Big(& \ln \alpha_L + \ln P(\sigma, \epsilon_P) + \ln A \\ & + \beta\alpha_L \ln(W^* - W^0) + \alpha_M \ln M^* + \alpha_K \ln K - \ln W^* \Big) \end{aligned} \tag{4.13}$$

Additionally, changes in the LRR in response to variations in the exogenous variables are derived from the following structural model:

$$\begin{aligned} LRR = \beta_0 + \beta_1 \ln \bar{L} + \beta_2 \ln P + \beta_3 \ln A + \beta_4 \ln E \\ + \beta_5 \ln M + \beta_6 \ln K + \beta_7 \ln W \end{aligned} \tag{4.14}$$

Finally, to assess the impact of working capital on the redundant employment rate, I use the wage equation (Equation (4.9)) while holding M constant. Substituting into Equation (4.14), I arrive at the following reduced-form model:

$$\begin{aligned} LRR = \pi_0 + \pi_1 \ln \bar{L} + \pi_2 \ln P + \pi_3 \ln A + \pi_4 \ln E \\ + \pi_5 \ln M + \pi_6 \ln K + \pi_7 \ln W^0 + \pi_8 \ln R \end{aligned} \tag{4.15}$$

The effect of a change in working capital, R, on the LRR is ambiguous because, while R is positively correlated with both W^* and M^*, it increases with W^* but decreases with M^*.

4.4 Indentification and Data

I have sketched a model to identify redundant labor and its explanatory factors. In this section, I discuss the empirical strategy used for its quantitative assessment, placing particular emphasis on the Cuban data used in the regressions. I begin by estimating the efficiency-wage augmented production function from Equation (4.2), which provides the elasticity α_L needed to calculate the LRR. Next, I estimate Equations (4.14) and (4.15), accounting for potential endogeneity issues in both cases.

4.4.1 Identification

The literature documents various methodological challenges that arise when estimating production functions, which can lead to bias and inconsistencies in many estimators (Griliches and Mairesse, 1995; Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Blundell and Bond, 2000). Moreover, the inclusion of the efficiency-wage hypothesis introduces an additional issue related to the potential simultaneity between wages and productivity, with the latter being embedded in the error term of the production function. One approach often taken in the literature to address this problem is the use of instrumental variables. Financial variables are typically used to instrument wages, based on the assumption that variables reflecting the firm's financial health do not influence total factor productivity (Wadhwani and Wall, 1991).

In my case, I base the instrumental variable strategy on the wage equation (Equation (4.9)), specifically using the base wage and firm liquidity as instruments, assuming that neither of these variables is related to productivity. In this strategy, I absorb fixed effects at different levels. I also explored the literature on the use of lagged instruments and lagged first differences (e.g., Blundell and Bond (2000)); however, unlike Gonzalez and Cribeiro (2018), I found no conclusive evidence that firms' current input choices were influenced by autoregressive productivity shocks.

On the other hand, while the theoretical model is based on the unidirectional influence of working capital on redundant employment, I acknowledge the possibility that, at the empirical level, this relationship could be bidirectional. In other words, redundant employment might create liquidity distortions within firms. To address this simultaneity issue, I implement an instrumental variable strategy, treating working capital as an endogenous regressor. Financial variables are employed as exogenous instruments for working capital. Specifically, the difference between accounts receivable and accounts payable, labeled as *Trade Credit*, is used as an instrument. An increase in *Trade Credit* is associated with improvements in liquidity, reflecting how firms manage their cash flow cycle with customers and suppliers, and is unrelated to excess labor. To further strengthen the strategy, I include another instrument, *Leverage*. The key assumption in the identification strategy is that these financial variables influence redundant employment only through their effect on working capital.

4.4.2 Data on the Cuban Manufacturing Industry

For the estimation of the main equations in my model, I use a panel dataset containing information on manufacturing SOEs from 2009 to 2013. For reference, I treat 2009 and 2010 as baseline years, prior to the start of the reforms. The data come from administrative records of Cuba's National Office of Statistics and Information, specifically from forms 5901-08 and 5903-07. The values reported by the firms represent annual totals and are classified by economic activity. A detailed description of each variable, based on the methodological notes from the administrative records, is provided in Appendix 4.6.

Equation (4.2) is estimated using the logarithmic values of gross output (Y), capital stock (K), employment ($\bar{L}$), materials (M), total wages (W), and base wage (W^0).[2] Monetary variables were deflated using price indices based on 2009, constructed from the Statistical Yearbook. The administrative records do not directly report W and W^0, but instead provide the firm's total wage bill (WL) and base wage expenditure (W^0L). To derive W and W^0, I divide these by the number of employees. In estimating the elasticity of effort, I use the difference between the firm's wage and the industry average.

The variables used to calculate the redundant employment rate, LRR, are defined as follows. With $L^* = \alpha_L \frac{(1-\epsilon_P)PY}{W^*}$, I take the labor elasticity of output, α_L, from the estimation of (4.2). Since there are no demand elasticity estimates for the Cuban economy, I approximate ϵ_P using a Lerner index (see variable definition in Appendix 4.6). In the data, the Lerner index is calculated for each firm and year in the sample. Finally, PY and W^* represent the nominal values of gross output and wages, respectively. LRR is calculated as the difference in logarithms between the exogenously set number of employees, $\bar{L}$, and the estimated optimal employment, L^*.

To measure the explanatory factors of the LRR not yet described, as well as other control variables added to the model, I proceed as follows. Prices, $P(Y, \sigma, \epsilon_P)$, are approximated using the firm's market share (FMS). FMS is commonly linked to both market power (Bresnahan, 1989; Shy, 1996) and demand characteristics that are not easily observable, such as product quality differences (Jin and Leslie, 2003; Kirmani and Rao, 2000). Alternatively, it may also be associated with greater efficiency, enabling a firm to sell products of equivalent quality at a better price than its competitors, thus capturing a larger share of market demand (Conner, 1991; Posner, 1978). All else being equal, I associate an increase in FMS with higher firm margins through any of these channels, leading to an outward shift of the MRPL curve and, consequently, a reduction in the LRR.

Productivity, A, is measured as average labor productivity, defined as real value added divided by the number of exogenously set workers. I associate increases in productivity with shifts in the marginal product of labor, and thus with reductions in the redundant employment rate. Additionally, I constructed a set of financial variables (Appendix 4.6). For working capital, R, I use the

[2] I use the inverse hyperbolic sine, defined as $\ln(x + \sqrt{x^2+1})$, for variables with negative values, such as financial variables, to avoid the loss of observations.

difference between current assets and current liabilities. An increase in R indicates that the firm has more liquid financial resources to meet short-term obligations, particularly wages and raw materials. As instruments for R, I use the net balance between short-term accounts receivable and accounts payable, which I label as *Trade Credit*. To further strengthen the instrumental variable strategy, I also measure long-term financial leverage, defined as the ratio of long-term liabilities to total assets. The strategic use of long-term debt can enhance a company's liquidity by allowing it to maintain more current assets to cover operational needs, potentially increasing working capital.

As control variables, I include three ratios: the subsidy-to-sales ratio (S), the export-to-sales ratio (X), and the dollar expenditure-to-total expenditure ratio (U), the latter serving as a proxy for the firm's reliance on imports. Additionally, to capture heterogeneity in the model, I add a dummy variable, *loss*, equal to one if the firm reported losses in the current period, and another dummy variable, *merge*, if the SOE was an acquirer during any merger process within the study period.

4.4.3 Descriptive Statistics

Table 4.1 presents the descriptive statistics. The sample consists of 751 state-owned manufacturing entities from 2009, which, according to the Office of National Statistics, represent 88% of the total companies in the sector (ONEI, 2022b). However, there is some attrition over time, in line with the institutional restructuring process that began in 2011, which led to the merger and closure of many SOEs (see Section 4.2). The table shows that the number of SOEs declined by just over a third between 2009 and 2013, matching the trend estimated using aggregated figures. Despite this attrition, the sample consistently represents 88% or more of the total number of companies each year, supporting the reliability of the data. A key feature of the sample is that the number of missing values is low, accounting for less than 1% of total observations, while variation from the mean remains high.

Variables	Statistics	2009	2010	2011	2012	2013
SOEs in Sample	N	751	731	556	510	483
Sample Coverage	%	88.8	92.1	90.7	91.1	88.0
SOEs Receiving Subsidies	%	18.5	16.0	18.3	20.0	21.3
Loss-Making SOEs	%	17.3	16.4	9.7	9.6	7.5
Merging Acquirer SOEs	%	-	-	6.5	8.2	9.7
Gross Output	Mean	1,812	1,749	2,461	2,835	2,983
(in 10,000 CUP)	S.D.	3,242	3,158	4,293	4,697	4,982
	Min	3.3	1.5	1.5	1.9	20.3
	Max	27,985	27,112	31,009	33,256	35,764
	Miss	4	3	1	1	1
Employees	Mean	522	506	606	617	632
(in Units)	S.D.	734	732	980	900	905
	Min	23	22	25	8	24
	Max	8,400	8,496	8,374	7,116	7,102
	Miss	3	3	2	0	0
Capital Stock	Mean	582	573	788	872	851
(in 10,000 CUP)	S.D.	1,008	983	1,724	1,830	1,836
	Min	0.3	0.8	0.8	1.6	0.2
	Max	12,230	11,010	18,173	16,153	15,775
	Miss	3	5	3	2	2
Materials	Mean	1,158	1,212	1,548	1,778	2,303
(in 10,000 CUP)	S.D.	2,532	2,669	3,366	3,860	4,829
	Min	2	0.2	1.5	1.3	0.7
	Max	23,523	24,260	24,525	34,295	33,950
	Miss	1	3	0	0	1
Wages	Mean	438	443	453	480	505
(in CUP)	S.D.	123	127	142	139	148
	Min	161	126	149	139	248
	Max	1,106	1,186	1,445	1,448	1,501
	Miss	4	3	2	0	1
Ratio Fixed Wage to Total Wage	Mean	0.68	0.68	0.65	0.63	0.60
(CUP/CUP)	S.D.	0.193	0.187	0.182	0.171	0.161
	Min	0.079	0.054	0.054	0.096	0.098
	Max	1	1	1	1	1
	Miss	7	7	1	1	0
Labor Productivity	Mean	1.267	1.279	1.539	1.705	1.698
(in 10,000 CUP)	S.D.	1.44	1.44	1.59	1.47	1.58
	Min	-1.77	-1.31	-2.49	-6.28	-6.2
	Max	28.41	22.17	14.71	12.61	16.47
	Miss	7	6	3	1	3
Ratio Subsidy to Sales	Mean	0.037	0.031	0.04	0.043	0.048
(CUP/CUP)	S.D.	0.107	0.101	0.116	0.122	0.131
	Min	0	0	0	0	0
	Max	0.717	0.641	0.672	0.796	0.709
	Miss	6	7	3	2	1
Margin	Mean	0.02	0.02	0.079	0.085	0.084
(CUP/CUP)	S.D.	0.329	0.391	0.277	0.211	0.475
	Min	-5.973	-6.499	-4.752	-2.127	-8.912
	Max	0.8	0.824	0.684	0.602	0.633
	Miss	4	6	3	4	1
Herfindahl-Hirschman Index	Mean	1,888.5	1,989.2	2,251.8	2,070.8	2,169.8
(in Units)	S.D.	2,071.2	2,095.8	2,310.6	2,339.6	2,457.8
	Min	219.2	252.2	332.1	323	349.2
	Max	10,000	10,000	10,000	10,000	10,000
	Miss	0	0	0	0	0
Return on Assets	Mean	0.051	0.066	0.09	0.078	0.093
(CUP/CUP)	S.D.	0.163	0.169	0.161	0.14	0.149
	Min	-0.942	-0.947	-0.754	-0.514	-0.323
	Max	0.784	0.681	0.72	0.786	0.915
	Miss	3	2	1	2	1
Financial Leverage	Mean	0.361	0.381	0.376	0.401	0.437
(CUP/CUP)	S.D.	0.327	0.332	0.329	0.337	0.346
	Min	0	0	0	0	0
	Max	1.895	1.574	0.977	1.043	1.135
	Miss	2	0	2	2	1
Trade Credit	Mean	-17.28	-12.6	-35.29	-40.51	-58.58
(in 10,000 CUP)	S.D.	455.84	440.73	662.29	501.73	760.99
	Min	-2,953.1	-1,797.6	-8,897.5	-4,395.1	-8,567.7
	Max	6,617.9	7,278.9	4,443.7	1,983.5	7,455.9
	Miss	1	1	1	0	0
Working Capital	Mean	142	144.7	319.8	437.6	563
(in 10,000 CUP)	S.D.	1032	1056	986	1109	1376
	Min	-13,000	-9,772	-4,976	-8,149	-3,990
	Max	17,174	14,482	15,008	14,425	13,996
	Miss	0	0	0	0	0

Note: Margin is defined as the ratio of profit to total revenue. Return on Assets (ROA) is calculated as profit divided by total assets. Financial Leverage is measured as long-term liabilities relative to total assets. Trade Credit is the difference between short-term accounts receivable and short-term accounts payable. Working Capital is defined as current assets minus current liabilities.

Table 4.1: Descriptive Statistics, 2009-2013

The observed increase in the number of exogenous workers, $\bar{L}$, from 522 to 632 over the study period, warrants further clarification, as it appears to contradict the narrative of this paper. In fact, if the values reported by firms were aggregated at the industry level, it would observe an average annual decline of nearly 6% in the number of workers between 2009 and 2013, which is consistent with the facts discussed in Section 4.2. The reason $\bar{L}$ increases, on average, is due to a reduction of more than 30% in the number of state-owned entities, resulting from the institutional restructuring process previously described. The combined effect of these two trends led to a smaller manufacturing sector overall, but one composed of larger firms on average.

To further clarify these findings, I leverage the potential of the data to identify the acquiring SOEs involved in the merger process. What I observed was that, compared to non-merged firms, the scale of the approximately 40 acquiring firms is disproportionately larger—not only in terms of the number of workers but also with respect to other inputs and production levels. In addition to this, the increase in the Herfindahl-Hirschman Index (HHI) of market concentration, as shown in the table, could be attributed to the reduction in the number of firms per industry and the significant expansion in scale of a select few. Finally, this statistical result aligns with anecdotal evidence from official sources, which suggests that the relative deregulation and increased autonomy granted to SOEs under the 2011 reforms were exploited by these firms to raise prices excessively, enabling them to earn extraordinary profits (Acosta and Riverí, 2017; Gutiérrez, 2017).

Regarding the explanatory factors of the LRR, I highlight the following facts. The average total wage, as shown in Table 4.1, increased from 438 CUP in 2009 to 505 CUP in 2013, figures that closely match the industry average reported in the Yearbooks (ONEI, 2022d), supporting the representativeness of data. It is worth noting the low wage levels when converted to U.S. dollars using the official parallel exchange rate —approximately 20 U.S. dollars per month. These low wages are statistically consistent with the theoretical framework of redundant workers. In other words, one way firms manage excess labor is by paying each worker less. However, wages increased by 3.6% over the period, driven by growth in the variable wage component. This increase led to a reduction in the share of the fixed wage, W^0, in the total wage structure by 8 percentage points. According to the model, an increase in wages pushes the LRR upward.

I also note that wage variation across firms, relative to the central value, is low compared to the variation observed in other variables presented in the table. I interpret this distinct characteristic as a direct result of the wage-setting regulations in Cuba, where state wages are tightly controlled by the economic authorities. These regulations even influence the internal composition of wages, with evidence showing lower variation in fixed wages compared to variable wages. For example, the coefficient of variation for fixed wages—0.322—is less than half that of variable wages, 0.752. The lower relative variation in fixed wages arises from the legally mandated pay scale, which is applied uniformly across all firms. In contrast, variable wages are tied to company performance, which is far more heterogeneous in the data, leading to greater variability.

Meanwhile, labor productivity increased by 7.6% on average across the industry, driven by faster growth in value added relative to the number of workers. In the model, higher productivity is associated with reductions in the LRR. In addition to this rise in productivity, the share of loss-making firms dropped significantly, from 17% in 2009 to 7% in 2013, while profitability, measured as return on assets (ROA), nearly doubled. However, the percentage of subsidized firms in the sample increased, and this expansion was accompanied by a 1 percentage point increase in the subsidy-to-sales ratio. It is important to clarify that these subsidies are not provided to loss-making firms but are compensations from the public budget for companies whose selling prices are regulated by the government.

From a financial perspective, I highlight the following key facts. Overall, the data show a significant increase in liquidity, averaging 39.5% annually. This growth may have been driven by several factors. First, improved conditions in the cash flow cycle, as reflected by the *Trade Credit* variable, which grew by an average of 8% during the period. Second, I observe that long-term *Financial Leverage* increased by 7 percentage points between 2009 and 2013, suggesting that firms may have used part of their permanent resources to finance current operations, contributing to the rise in working capital. Lastly, I document a 6.2 percentage point increase in the *Margin* variable, indicating that, on average, firms captured a larger share of the price, consistent with market concentration and the extraordinary profits mentioned earlier.

4.5 Results: Measuring Labor Redundancy and its Determinants

The descriptive exploration of the data suggests that manufacturing firms experienced a period of expansion between 2009 and 2013, with improvements in profitability and liquidity indicators. Production increased, raw materials became more readily available, wages rose, and the number of loss-making firms decreased. In this section, I focus on measuring the Labor Redundancy Rate (LRR) and conducting a quantitative assessment of its main explanatory factors.

4.5.1 Labor Redundancy Rate

Table 4.2 presents the regression results for the log-linearized, efficiency-wage-augmented production function. I compare estimates of Equation (4.2) using Ordinary Least Squares (OLS), Fixed Effects (FE), and Instrumental Variables (IV) to examine how the estimated coefficients for the production factors change under different assumptions about the correlation between the error term and the regressors. Two variants are estimated for each model, differing by the inclusion of controls. The added controls include the export-to-sales ratio, the dollar expenditure-to-total expenditure ratio, and the subsidy-to-sales ratio. In all cases, fixed effects at various levels were absorbed, as shown in the table.

All estimated elasticities are statistically significant at the 5% level or higher, except for the elasticity of output with respect to capital stock. This result is common when estimating production functions in Cuba, particularly in the manufacturing sector (see, for example, Doimeadios and Sanchez (2015); Gonzalez and Cribeiro (2018)). In addition to the well-known methodological challenges in estimating these functions, which have been studied in the Cuban context by Gonzalez and Cribeiro (2018), another possible explanation could be the existence of idle capacity in firms—consistent with the theoretical framework on overemployment—or the weakness of the investment process and the high depreciation of physical capital. In any case, this is a topic that warrants further research, which is beyond the scope of this paper.

The strategy of controlling for firm-level fixed effects was validated using the Hausman test, which rejected the null hypothesis at all levels of statistical significance (p-value = 0.000). It is generally accepted that unobservable firm-specific characteristics—such as management skills—impact the choice of inputs, thus breaking the independence between the error term and the inputs required for OLS to consistently estimate the parameters of the production function (Griliches and Mairesse, 1995). To further strengthen my estimate, I also examined whether the current choice of inputs was influenced by past realizations of the error term, following the work of (Blundell and Bond, 2000); however, I found no conclusive evidence to support this.

Dependent Variable: Gross Output

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Labor	0.421***	0.428***	0.666***	0.679***	0.647***	0.660***
	(0.024)	(0.029)	(0.089)	(0.087)	(0.093)	(0.091)
Capital	-0.020	-0.022	-0.001	-0.002	-0.000	-0.002
	(0.015)	(0.016)	(0.020)	(0.020)	(0.020)	(0.020)
Materials	0.550***	0.547***	0.414***	0.404***	0.426***	0.415***
	(0.036)	(0.039)	(0.042)	(0.040)	(0.046)	(0.045)
Effort	0.151***	0.152***	0.128***	0.128***	0.092**	0.092**
	(0.024)	(0.024)	(0.012)	(0.012)	(0.038)	(0.038)
Absorbed Fixed Effects						
Firm	No	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Loss	Yes	Yes	Yes	Yes	Yes	Yes
Merged	Yes	Yes	Yes	Yes	Yes	Yes
Subsidized	Yes	Yes	Yes	Yes	Yes	Yes
Controls						
Export		-0.075		0.117		0.127
		(0.178)		(0.167)		(0.163)
Import		0.096		0.220***		0.213***
		(0.115)		(0.061)		(0.062)
Subsidy		0.156		-0.172		-0.184
		(0.150)		(0.122)		(0.128)
Excluded Instruments						
w^0					0.420***	0.418***
					(0.055)	(0.056)
r					-0.065***	-0.064***
					(0.021)	(0.020)
Hausman Test †			0.000	0.000		
Instruments					9	12
Hansen-J †					0.524	0.536
Cragg-Donald Statistic					235.2	232.4
Kleibergen-Paap Statistic					35.2	34.6
Observations	2982	2982	2949	2949	2949	2949
R-squared	0.943	0.943	0.976	0.976	0.518	0.522

All variables in logs. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Export represents the ratio of exports to sales, Import refers to the ratio of material expenses valued in USD to total material expenses valued in both USD and CUP at the official exchange rate, and Subsidy is the ratio of subsidies to sales. w^0 denotes the normalized fixed wage, and r represents the liquidity ratio

Table 4.2: Regression of the Efficiency Wage-Augmented Production Function

The use of instrumental variables suggests that, in the Fixed Effects model, the elasticity of output with respect to effort may be overestimated, potentially due to the simultaneity between wages and productivity. Notably, the coefficient estimated by IV is nearly one-third lower than that estimated by Fixed Effects. The instruments used in the strategy —normalized fixed wage (w^0) and the liquidity ratio (r)— were individually significant at the 1% level in the first stage of the regression, and they easily passed the weak instrument tests of Cragg-Donald and Kleibergen-Paap. Finally, according to the J statistic from the Hansen test, I do not reject the hypothesis that all instruments identify the same parameter, indicating no evidence of overidentification. Consistent with these results, I use the labor elasticity, $\hat{\alpha}_L$, estimated by IV with additional controls (column 6) to calculate the redundant employment rate, LRR.

Before delving into the results of the rest of the model, I present the calculations for marginal products, total factor productivity (TFP), and the redundant employment rate in Table 4.3. I highlight three key findings. First, the geometric mean of the redundant employment rate in the base year, 2009, is significantly high; taking the exponential of 1.38 gives an average $\bar{L}$ to L^* ratio of about 4:1. While this result exceeds similar estimates from other transition economies, such as China and Vietnam (World Bank, 1996; Dong and Putterman, 2003; Belser and Rama, 2001), it is not surprising given the Cuban context described in Section 4.2 and the substantial difference between the average marginal product of labor (MPL) and total wages, reported in columns 3 and 4 of Table 4.3, respectively. Notably, in the base year, worker compensation is more than double the contribution of an additional unit of labor on average.

Additionally, it is striking that nearly 93% of state-owned enterprises had redundant employment in 2009 (column 6), indicating that overemployment is not an isolated issue but rather a widespread problem in the industry. Lastly, estimation suggests that the total number of redundant workers in the industry would be 236,000 in 2010, representing 47% of the redundant employment estimated by the government across all state-owned enterprises in that year (Granma, 2010).

The second notable finding relates to the temporal trend of redundant employment during the reform period. The negative sign of the LRR indicates that, compared to 2009, the rate has experienced a significant accumulated decline, bringing the $\bar{L}$ to L^* ratio down to 3:1 (still a high value, but undeniably lower than in the base year). This reduction coincides with increased contributions from materials and labor—particularly the latter—to product growth, along with a parallel expansion in TFP. The greater contribution of labor to output may be attributed to heightened effort, driven by rising variable wages, which is the component of total wages that saw the most substantial increase during this period (see Table 4.1). The decline in the LRR is also accompanied by a 2 percentage point decrease in the number of firms with redundant employment and a reduction of over 100,000 in the number of excess employees.

Third, I highlight that the average marginal product of materials (MPM) consistently exceeds 1 throughout the examined period. This indicates that SOEs are allocating fewer raw materials than would be necessary to equate the marginal return of the factor with its marginal cost. In other words, this finding aligns with the widespread scarcity of raw materials, energy, and fuels, which is both a well-documented empirical regularity in the Cuban context and compatible with the implications of excess labor within the theoretical framework.

4.5.2 Explanatory Factors of Redundant Employment

Table 4.4 presents the estimates for Equations 4.14 and 4.15 regarding redundant employment, representing the structural model and the reduced-form model, respectively. While the structural model is estimated using only fixed effects, the reduced-form model employs both fixed effects and instrumental variables to address the possibility of simultaneity between excess employees and working capital. Additionally, the equations are estimated using differenced variables rather than levels to provide better control over potential common trends among the model variables, reducing the risk of obtaining spurious regression results.

Columns 1 and 2 of Table 4.4 present the results of the structural model, both with and without controls, respectively. Apart from the capital stock, which was also found to be insignificant, all coefficients in the model have the expected signs. The growth rate of the FMS, which in the empirical model captures the effect of prices explained not only by market power but also by demand characteristics and firm efficiency, corresponds nearly one-to-one with the reduction in the growth rate of the LRR during the period. This finding suggests that the price increases—documented not only in this paper but also by Acosta and Riverí (2017) and Gutiérrez (2017)—were a key factor in financing excess labor in the study period, presumably by raising the marginal revenue product of labor in SOEs.

	LRR	MPM	MPL	Wages	LnTFP	% of SOE with RL	Excess, in 1000s (sum)	L, in 1000s (sum)	t-test	# of SOEs
2009 (Base year)	1.388***	1.202***	2.050***	5.274***	1.767***	0.928***	259.7	389.0	0.000	739
2010	-0.042**	-0.089*	-0.028	0.066**	-0.095***	0.002	236.2	366.8	0.000	716
2011	-0.173***	0.292**	0.391***	0.191***	0.092***	-0.016***	189.5	334.6	0.000	547
2012	-0.211***	0.322***	0.614***	0.489***	0.211***	-0.018**	170.3	313.9	0.000	503
2013	-0.248***	-0.014	0.646***	0.764***	0.070*	-0.020*	155.9	304.7	0.000	479

Notes: Table entries (columns 1 to 6) correspond to regressions with annual dummies, controlling for firm, industry, losses, mergers, and subsidies fixed effects. The 2009 base year coefficient represents the annual average, followed by the cumulative yearly deviation from that average. Standard errors (not reported) are robust. Excess is the sum of the absolute difference between L and L^*. The t-test refers to a paired t-test where $H_0 : \text{mean}(L - L^*) = 0$, with p-values reported. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.3: Average Marginal Products, TFP, Wages and Labor Redundancy Rate

Another key factor contributing to the observed reduction in the LRR during this period was the growth rate of materials (raw materials, fuels, and energy), which was statistically significant at the 1% level. Quantitatively, a 1% increase in the growth rate of raw materials results in a 0.27% and 0.32% reduction (depending on the model) in the growth rate of the redundant employment rate. It could be inferred, as a conjecture, that the combination of the rising marginal product of materials—identified in Section 4.5.1— along with the greater availability of this production factor during the study period (see Table 4.1), may have led to an increase in the marginal product of labor, potentially influencing the reduction in redundant employment through this channel.

On the other hand, a 1 percentage point increase in the growth rate of labor productivity is associated with reductions in the redundant employment rate of between 0.07 and 0.08 percentage points. The fixed effects estimation remains robust even with the introduction of both controls and working capital in the reduced-form model (Column 3). This result aligns with the TFP growth documented in Section 4.5.1; however, the impact of this factor on the reduction of the LRR is not as strong as that of the other two variables discussed earlier.

Consistent with the theoretical model, the Labor Redundancy Rate (LRR) was positively correlated with both the growth rate of wages and the growth rate of exogenous employment, across all levels of statistical significance. Specifically, the structural model shows that for each percentage point increase in the growth rates of wages (W) and average employment ($\bar{L}$), the growth rate of redundant employment increased by 0.57 and 0.36 percentage points, respectively. However, the combined effect of these two variables was insufficient to counterbalance the decline in the LRR, which was driven downward by the combined influences of prices, materials, and labor productivity. Moreover, I found no statistically significant evidence that effort contributed to the reduction of the LRR within the structural model. That said, I cannot draw definitive conclusions on this matter, as the result may be influenced by potential multicollinearity between total wages and effort, given that effort is proxied by variable wages. Further research is needed to clarify this finding.

Before delving into the results of the reduced-form model, I briefly discuss the impact of the control variables added to capture heterogeneity in the results and strengthen the estimates. On one hand, the ratio of dollar expenditures to total expenditures was not statistically significant in any specification. In contrast, the export-to-sales ratio showed a statistically significant negative correlation with the LRR. One possible explanation for this finding is that exposure to international trade may improve the relative productivity of these firms, resulting in a higher average employment level (L^*) in both static and

Dependent Variable: ΔLabor Redundancy Rate

	Fixed Effects			Instrumental Variables	
	(1)	(2)	(3)	(4)	(5)
ΔFirm Market Share	-0.986**	-0.940**	-0.946**	-0.966**	-0.920**
	(0.401)	(0.399)	(0.396)	(0.401)	(0.400)
Δ Labor	0.356***	0.355***	0.274***	0.254***	0.253***
	(0.063)	(0.060)	(0.064)	(0.065)	(0.063)
ΔCapital	0.019	0.019	0.025*	0.032**	0.031**
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
ΔMaterials	-0.317***	-0.309***	-0.269***	-0.275***	-0.267***
	(0.044)	(0.042)	(0.041)	(0.041)	(0.040)
ΔEffort	-0.009	-0.009	0.096***	0.070***	0.071***
	(0.015)	(0.015)	(0.022)	(0.021)	(0.021)
ΔLabor Productivity	-0.068***	-0.078***	-0.045***	-0.024	-0.034*
	(0.016)	(0.014)	(0.015)	(0.019)	(0.018)
ΔWage	0.566***	0.574***			
	(0.066)	(0.065)			
ΔFixed Wage			0.082***	0.061***	0.061***
			(0.016)	(0.022)	(0.022)
ΔWorking Capital			-0.006*	-0.030**	-0.026*
			(0.003)	(0.013)	(0.013)
Absorbed Fixed Effects					
Firm	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Loss	Yes	Yes	Yes	Yes	Yes
Merged	Yes	Yes	Yes	Yes	Yes
Subsidized	Yes	Yes	Yes	Yes	Yes
Controls					
Export		-0.683**	-0.604**		-0.574**
		(0.271)	(0.260)		(0.251)
Imports		0.069	0.071		0.034
		(0.076)	(0.079)		(0.081)
Subsidy		0.446**	0.475**		0.513**
		(0.207)	(0.204)		(0.197)
Excluded Instruments					
ΔTrade Credit				0.048***	0.046***
				(0.014)	(0.014)
ΔLeverage				1.191***	1.185***
				(0.324)	(0.321)
Total Number of Instruments				12	15
Hansen-J †				0.801	0.851
Cragg-Donald Wald F statistic				70.4	67.96
Kleibergen-Paap rk Wald F statistic				14.5	12.99
Observations	1996	1996	1996	1991	1991
R-squared	0.483	0.498	0.447	0.283	0.306

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All variables are in logarithms. X represents the export-to-sales ratio, U represents the ratio of expenses in USD to total expenses, and S represents the subsidy-to-sales ratio. w^0 is the normalized fixed wage, and r is the current ratio. † denotes that the p-value of the statistic is reported. Uncentered R-squared reported in IV estimates. The differences in the number of observations are due to the exclusion of singletons in the estimation process.

Table 4.4: Labor Redundancy Regressions: Structural and Reduced Form Models

dynamic terms. Alternatively, it could be argued that the government takes a firm's international exposure into account when exogenously setting average employment ($\bar{L}$), leading to a more restrictive employment policy for this specific group of companies.

In any case, a precise explanation for this phenomenon requires further research beyond the scope of this paper. Lastly, the results suggest that receiving subsidies from the state budget contributes to the rate of redundant employment in firms. Given the nature of these subsidies—compensations provided to firms with regulated prices—it can be inferred that, unlike firms with greater pricing autonomy, regulated firms have less flexibility to use pricing strategies to finance excess labor. Alternatively, subsidies may act as a disincentive for productivity, making it more difficult for these firms to manage redundant employment, which could explain this outcome. As with the export case, further research is needed to clarify this issue.

Columns 3 to 5 of Table 4.4 present the estimates from the reduced-form model, which allows to analyze the impact of financial constraints (working capital) on the rate of redundant employment. In all three variants, working capital is significant at the 5% level (t-statistic ≈ 2). The negative coefficient suggests that short-term improvements in firms' liquidity contributed to reducing the redundant employment rate. Moreover, the instrumental variable strategy (columns 4-5) increased the estimated coefficient by five to six times compared to the fixed effects estimation. The instruments used —variations in *Trade Credit* and *Leverage*— were individually significant at the 1% level in the first stage of the estimation and passed the Cragg-Donald and Kleibergen-Paap weak instrument tests. Additionally, I did not reject the null hypothesis in the Hansen test, indicating no evidence of overidentification.

While the magnitude of the working capital coefficient is modest compared to other variables analyzed, it is reasonable within the context of this study. Unlike the reference work of Dong and Putterman (2003), where Chinese firms operated in an adverse environment marked by significant declines in profitability, increased competitive pressure, tighter financial constraints, and shortages of productive inputs, the situation in Cuba is quite different. Without delving into the specific causes (whether reforms or exogenous factors), my results indicate that between 2011 and 2013, the performance of state-owned enterprises improved significantly compared to the two preceding years. This suggests that these firms had to rely less on the relaxation of financial constraints to manage the effects of redundant employment, as was observed in the Chinese case. However, it is important to clarify that this improvement does not imply that redundant employment is no longer a burden; rather, it indicates that, despite its persistence, the economy was able to manage its consequences more effectively.

4.6 Conclusions

In this paper, I examine the distortionary effects that arise when firms lose the ability to freely choose their production factors, particularly labor. Using a two-stage optimal choice model, I find that redundant labor not only distorts firms' profitability but also negatively impacts their short-term financial health. I apply the model to the Cuban case from 2009 to 2013 to measure and evaluate the factors contributing to over-hiring. Findings show that in 2009, 93% of state-owned manufacturing firms in Cuba were affected by this issue, with an estimated 66% of state employment in this sector being redundant. By 2013, I observe a substantial reduction of 15 percentage points in the rate of redundant employment. I attribute this decline to a combination of rising prices, increased availability of raw materials, productivity growth, and improved financial performance during the period.

Appendix 4A: Detailed Description of Variables

Variable	Methodological Notes	Source
Gross Output	Consists of net production sales, adjustments in the variation of finished products valued at selling price, and changes in finished goods inventory, also valued at selling price, for reasons other than production and delivery.	Formulario 5903-07, Fila-100
Net Sales	Includes the delivery of finished products to customers, completed work, services rendered, and goods purchased for resale. It also accounts for state budget subsidies to be received, less sales taxes, returns, and discounts.	Formulario 5903-07, Fila-200
Subsidy	Total amount of subsidies received by the company from the state budget.	Formulario 5901-08, Fila-23
Exports	Sales of goods and services destined for export.	Formulario 5903-07, Fila-601
Labor	Average number of employees. It represents the annual average, calculated from the monthly employee count.	Formulario 5903-07, Fila-1100
Capital Stock	Book value of tangible net fixed assets, calculated as the average of values at the start and end of the year, after accounting for depreciation.	Formulario 5901-08, Fila-
Materials	Total company expenditure on raw materials, fuel, and energy.	Formulario 5903-07, Fila-
Wage Bill	Total annual wage expenditure of the company.	Formulario 5903-07, Fila-900
Fixed Wage	Total wage bill determined according to the Wage Scale Law, which sets the fixed or basic compensation for each occupational category (Resolution 30/2005 of the Ministry of Labor and Social Security).	Formulario 5903-07, Fila-910
Value Added	Calculated by subtracting the cost of materials and purchased services from the gross output of goods and services.	Formulario 5903-07, Fila-1600
Profits	Profit or loss represents the difference between total revenue and total expenses before taxes. This aligns with the balance reported on the company's Income Statement.	Formulario 5903-07, Fila-700
Assets	Total value of all assets and claims against third parties, including current assets, long-term assets, fixed assets, deferred assets, and other assets.	Formulario 5903-07, Fila-1901
Current Assets	Goods and rights expected to be converted into cash, sold, or consumed within the company's normal operating cycle, typically one year. Includes cash, accounts receivable, inventories, and other assets that can be liquidated in the short term to meet short-term obligations.	Formulario 5903-07, Fila-2000
Short-term Notes and Accounts Receivable	Outstanding credit documents received and awaiting collection, including bills of exchange and promissory notes accepted by customers. Includes sales of products and merchandise, as well as completed work, excluding discounted receivables and the allowance for doubtful accounts.	Formulario 5903-07, Fila-2300
Liabilities	Total value of the company's obligations to third parties, including both short-term and long-term liabilities, as well as deferred and other liabilities.	Formulario 5903-07, Fila-1903
Current Liabilities	Short-term financial obligations due within one year or within the company's normal operating cycle, including accounts payable, short-term loans, accrued expenses, and other near-term debts.	Formulario 5903-07, Fila-2100
Short-term Notes and Accounts Payable	Outstanding credit documents awaiting payment, including bills of exchange and promissory notes received from suppliers. This includes amounts due to suppliers for routine transactions, whether payment occurs before or after receipt of goods, materials, or services.	Formulario 5903-07, Fila-2400
Lerner Index	Pre-tax profits divided by the firm's total revenue. For firms with negative profits, we assume a very high price elasticity of demand, preventing them from exercising market power. In such cases, the Lerner index is set to zero.	
Firm Market Share (Revenue)	Share of the total gross output in the industry produced by SOE i in year t. Industries are classified according to Cuba's 4-digit Economic Activity Code.	
Herfindahl-Hirschman Index	For each industry-year, the sum of the squares of each firm's market share, multiplied by 100.	

Table 4.5: Summary of Variables and Data Sources

5 Concluding Remarks

This thesis has explored various dimensions of misallocation in transition economies, using Cuba as a case study. It presents the first quantitative assessment of the potential welfare costs of misallocation—measured in terms of TFP—in this particular transition experience. The potential drivers of misallocation in Cuba have been analyzed through a conceptual and theoretical framework, supported by statistical evidence. Specifically, the study examines idiosyncratic distortions in the foreign exchange and labor markets. Overall, the findings suggest that Cuba possesses significant untapped productivity reserves and that advancing the structural reforms initiated in 2011 is among the most effective ways to reduce misallocation and revive economic growth.

In Chapter 2, the application of the seminal framework developed by Hsieh and Klenow (2009) enabled the quantification of the aggregate impact of specific heterogeneous distortions on firms. The results reveal significant dispersion in TFPR, indicating substantial potential gains from transitioning to a more efficient allocation of resources. Specifically, TFP gains are estimated to range between 45% and 75%. The pronounced productivity disparities among firms and the extent of misallocation highlight the critical role of structural reforms in boosting aggregate productivity—an urgent priority for Cuba given the multiple existing distortions and imbalances. The importance of such reforms is evidenced by the improvements in allocative efficiency documented in this thesis following the 2011 reforms (at least until 2017), which, despite their limited scope, contributed to accelerating growth.

Chapter 3 concludes that exchange rate policy is a cornerstone of Cuba's reform agenda, as addressing distortions in the foreign exchange market is closely linked to significant improvements in allocative efficiency. Firm-level data generated under Cuba's unique monetary arrangement provides valuable insights into how the system of implicit taxes and subsidies resulting from an overvalued currency leads to inefficient resource allocation. Specifically, when TFPR is adjusted to account for currency transfers, manufacturing value added doubled—an increase that is 50 percentage points higher than the efficiency gains obtained when TFPR is not corrected for this issue. These findings highlight the risk of underestimating the true costs of misallocation when exchange rate misalignment is overlooked.

In Cuba's case, unifying the official and parallel exchange rates could significantly enhance aggregate productivity by eliminating the burden (exchange tax) imposed on the export sector, which is forced to operate at the overvalued official exchange rate. Exchange rate overvaluation not only disincentivizes exports but also undermines the country's external financial balance and hinders the expansion of industries aligned with its comparative advantages. This, in turn, slows structural transformation and weakens external competitiveness.

Additionally, this study highlights the relevance of the parameters ϕ^d and ψ^d—which provide redundant information under equilibrium conditions—in fully characterizing resource misallocation in contexts of exchange rate misalignment. This finding suggests that in economies experiencing significant currency misalignment, misallocation studies should be revisited to ensure a more accurate assessment of inefficiencies and potential policy responses.

Finally, Chapter 4 concludes that the Cuban labor market holds untapped potential for productivity improvements through the reduction of labor misallocation. The fact that firms lack autonomy in hiring and dismissing workers constitutes a policy burden that results in redundant employment, with significant consequences for profitability, competitiveness, and short-term financial health.

The chapter applies a two-stage optimal choice model to data from 2009 to 2013. The results show that in 2009, 93% of manufacturing SOEs in Cuba were affected by this issue, with an estimated 66% of state employment in this sector classified as redundant. However, by 2013, a substantial reduction of 15 percentage points in the redundancy rate was documented. This decline is attributed to a combination of rising prices, increased availability of raw materials, productivity growth, and improved financial performance during the period. Restoring firms' autonomy to manage their workforce size could serve as a catalyst for productivity. However, the results suggest that reforms in other areas, such as pricing policies, could also help mitigate the negative effects of redundant employment.

Bibliography

Acosta, M. d. l. C. and Riverí, C. (2017). Propuesta de regulación de precios mayorista. *Revista Cubana de Finanzas y Precios*, 1(4):38–45.

Alonso, J. A. and Vidal, P. (2020). Las dificultades del cambio institucional en cuba. *Políticas sociales y reforma institucional en la Cuba pos-COVID*, page 149.

Alpysbayeva, D. and Vanormelingen, S. (2022). Labor market rigidities and misallocation: Evidence from a natural experiment. *Labour Economics*, 78:102229.

Alvarez, O. (2014). Impacto del tamaño de la empresa en las brechas de productividad: La industria manufacturera en cuba. Tesis, Facultad de Economía, Universidad de La Habana, La Habana.

Belser, P. and Rama, M. (2001). *State ownership and labor redundancy: estimates based on enterprise-level data from Vietnam*, volume 2599. World Bank Publications.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007). Firms in international trade. *Journal of Economic perspectives*, 21(3):105–130.

Bils, M., Klenow, P. J., and Ruane, C. (2021). Misallocation or mismeasurement? *Journal of Monetary Economics*, 124:S39–S56.

Blundell, R. and Bond, S. (2000). Gmm estimation with persistent panel data: an application to production functions. *Econometric reviews*, 19(3):321–340.

Braguinsky, S., Ohyama, A., Okazaki, T., and Syverson, C. (2015). Acquisitions, productivity, and profitability: evidence from the japanese cotton spinning industry. *American Economic Review*, 105(7):2086–2119.

Bresnahan, T. (1989). Empirical studies of industries with market power. *The Handbook of Industrial Organization/North Holland.*

Buera, F. J., Kaboski, J. P., and Shin, Y. (2015). Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Review of Economics*, 7(1):409–436.

Busso, M., Madrigal, L., and Pages, C. (2013). Productivity and resource misallocation in latin america. *The BE Journal of Macroeconomics*, 13(1):903–932.

Caselli, F. (2005). Chapter 9 accounting for cross-country income differences. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1, pages 679–741. Elsevier.

Castro, E. and Saenz, T. (2021). La inversión extranjera en cuba: retos y desafíos en la actual coyuntura. *Economía y Desarrollo*, 165(2, Esp.):e5.

Castro, Y. (2015). Sacudir el lastre de las pérdidas. Granma, Online edition. Consultado el 17/11/2024.

CEPAL (1999). *La economia cubana. Reformas estructurales y desempenyo en los noventa.* CEPAL, Sede Subregional del Mexico, 2dn edition.

CEPAL (2024). Balance preliminar de las economías de américa latina y el caribe-2024. Reporte, Comisión Económica para América Latina y el Caribe (CEPAL). Disponible en: https://repositorio.cepal.org/server/api/core/bitstreams/33ec9026-4546-48c3-a6ca-e151cdedb332/content.

Conner, K. R. (1991). A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm? *Journal of management*, 17(1):121–154.

Coudert, V. and Couharde, C. (2009). Currency misalignments and exchange rate regimes in emerging and developing countries. *Review of International Economics*, 17(1):121–136.

Cribeiro, Y. (2012). Contribución de la fuerza de trabajo calificada al crecimiento económico en cuba. principales determinantes. *Economía y Desarrollo*, 148(2).

de la Torre, A. and Ize, A. (2014). Exchange rate unification: The cuban case. *Brooking Institution.*

del Castillo, L. (2017). El marco institucional en la transición socialista y el sistema de dirección y gestión de las empresas de propiedad social. *International Journal of Cuban Studies*, 9(1):37–62.

Demir, F. and Razmi, A. (2022). The real exchange rate and development theory, evidence, issues and challenges. *Journal of Economic Surveys*, 36(2):386–428.

Di Bella, G., Grigoli, F., Romeu, R., et al. (2020). A note on the algebra of multiple exchange rates. *Economics Bulletin*, 40(1):382–387.

Di Bella, G. and Wolfe, A. (2008). A primer on currency unification and exchange rate policy in cuba: lessons from exchange rate unification in transition economies. *Cuba in Transition*, 18:50–62.

Dias, D. A., Marques, C. R., and Richmond, C. (2016). Misallocation and productivity in the lead up to the eurozone crisis. *Journal of Macroeconomics*, 49:46–70.

Diaz, I. (2018). La innovación en empresas estatales cubanas: análisis para un debate. *Economía y Desarrollo*, 159(1):166–184.

Diaz, I., Gonzalez, R., and Zipitria, L. (2024). ¿la empresa estatal cubana? ¿por qué no han funcionado los cambios asociados al proceso de actualización? In *Empresas Estatales Cubanas. Situacion y Propuesta de Transformacion*, pages 126–148. Ediorial Ciencias Economicas ANEC.

Doimeadios, Y. (2012). El crecimiento económico en cuba: un análisis desde la productividad total de los factores. Technical report, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).

Doimeadios, Y. and Hidalgo, V. (2011). *Políticas macroeconómicas en economías parcialmente dolarizadas. La experiencia internacional y Cuba*, chapter Dualidad monetaria de la economía cubana. Retos de la política económica, pages 67–116. Zona Libro.

Doimeadios, Y. and Sanchez, A. (2015). Productividad y eficiencia en la economía cubana: una aproximación empírica. *Economía y Desarrollo*, 153:90–107.

Dong, X.-Y. and Putterman, L. (2003). Soft budget constraints, social burdens, and labor redundancy in china's state industry. *Journal of Comparative Economics*, 31(1):110–133.

Dong, X.-Y. and Xu, L. C. (2009). Labor restructuring in china: Toward a functioning labor market. *Journal of Comparative Economics*, 37(2):287–305. Special Symposium in Honor of Padma Desai.

Eden, M. and Nguyen, H. (2012). Correcting real exchange rate misalignment: Conceptual and practical issues. *World Bank Policy Research Working Paper*, (6045).

Feinberg, R. (2013). Foreign investment in the new cuban economy. *NACLA Report on the Americas*, 46(1):13–18.

Fernandez, O. and Torres, R. (2020). Las empresas estatales en las economias de planificación central: el caso cubano e ideas para su transformación. *International Journal of Cuban Studies*, 12(1):13–28.

Frenkel, R. and Ros, J. (2006). Unemployment and the real exchange rate in latin america. *World development*, 34(4):631–646.

Gandhi, A., Navarro, S., Rivers, D., et al. (2017). How heterogeneous is productivity? a comparison of gross output and value added. *Journal of Political Economy, forthcoming.*

Garcia, F. and Sovilla, B. (2017). Características y vulnerabilidades del sector exportador cubano en el periodo revolucionario (1959-2013). *América Latina en la historia económica*, 24(1):140–170.

González, R. (2021). Transferencias del presupuesto al sector empresarial en cuba¿ subsidio cambiario o subsidio por ineficiencia? *Economía y Desarrollo*, 165.

Gonzalez, R. and Cribeiro, Y. (2018). Dynamic panel data estimates of a production function: An application to cuban manufacturing industry. *Investigación Operacional*, 39(2):139–206.

Gonzalez, R. and Torres, R. (2024). Removiendo restricciones presupuestarias blandas en cuba: Hacia una reforma productiva real. Accessed: 2024-12-16.

Granma (2010). Pronunciamiento de la central de trabajadores de cuba. Sitio Web del periódico Granma. News article published on September 13, 2010. [Accessed: October 11, 2024].

Gray, M. S. T. (2021). Recognizing reality: Unification of official and parallel market exchange rates. *IMF*, page 45.

Griliches, Z. and Mairesse, J. (1995). Production functions: the search for identification.

Gutiérrez, S. D. (2017). Propuesta de esquema óptimo de aplicación de los métodos de formación de precios en cuba. *Revista Cubana de Finanzas y Precios*, 1(4):46–54.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The quarterly journal of economics*, 114(1):83–116.

Hernández-Catá, E. et al. (2017). Reforming cuba s s exchange system: The perils of gradualism. In *Annual Proceedings*, volume 27. The Association for the Study of the Cuban Economy.

Hernández, L. M. (2018). Examina presidente cubano programas de empleo y de inversiones en el país. Granma, Online edition. Consulted: October 15, 2024.

Hidalgo, V. and Cribeiro, Y. (2015). Estrategia de crecimiento y equilibrio macroeconómico en cuba. *Economía y desarrollo*, 153:30–48.

Hinkle, L. E. and Monteil, P. J. (1999). *Exchange rate misalignment: Concepts and measurement for developing countries*. Oxford University Press.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Jones, C. I. (2016). The facts of economic growth. In *Handbook of macroeconomics*, volume 2, pages 3–69. Elsevier.

Kabiraj, S. (2020). Analyzing the sources of misallocation in indian manufacturing: A gross-output approach. *Asian Development Review*, 37(2):134–166.

Kirmani, A. and Rao, A. R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of marketing*, 64(2):66–79.

Kornai, J. (1980). *Economics of Shortage*. North-Holland, Amsterdam.

Kornai, J., Maskin, E., and Roland, G. (2003). Understanding the Soft Budget Constraint. *Journal of Economic Literature*, 41(4):1095–1136.

Kunzmann, M. (2021). "ni plan ni mercado". problemas y coherencia del enfoque de reforma gradualista. In Hoffman, B., editor, *Políticas sociales y reforma institucional en la Cuba pos-COVID*, pages 173–200. Verlag Barbara Budrich.

Lage, C. (2016). *Esquema General de Politica Monetaria en el Nuevo Escenario de la Economia Cubana*. PhD thesis, Facultad de Economia, Universidad de La Habana. Consultado el 1/11/2022.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341.

Liao, F., Sun, Y., and Xu, S. (2023). Do the policy burdens of state-owned enterprises affect the efficiency of resource allocation of tax incentives? *Environmental Science and Pollution Research*, 30(30):75957–75972.

Lin, J. Y., Cai, F., and Li, Z. (1998). Competition, policy burdens, and state-owned enterprise reform. *The American economic review*, 88(2):422–427.

Lin, J. Y. and Li, Z. (2008). Policy burden, privatization and soft budget constraint. *Journal of Comparative Economics*, 36(1):90–102.

Lin, J. Y. and Tan, G. (1999). Policy burdens, accountability, and the soft budget constraint. *American Economic Review*, 89(2):426–431.

Lixing, L. (2008). Employment burden, government ownership and soft budget constraints: Evidence from a chinese enterprise survey. *China Economic Review*, 19(2):215–229.

Manning, A. (1987). An integration of trade union models in a sequential bargaining framework. *The Economic Journal*, 97(385):121–139.

Marion, N. P. (1994). Dual exchange rates in europe and latin america. *The World Bank Economic Review*, 8(2):213–245.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6):1695–1725.

MFP (2013a). Resolucion 441/2013. Ministerio de Finanzas y Precios de Cuba. Documento en formato PDF no publicado.

MFP (2013b). Resolución Nro. 38/2013. Ministerio de Finanzas y Precios (MFP), Publicada en la Gaceta Oficial No. 010 Ordinaria, 27 de febrero de 2013.

Nguyen, P. T. and Nguyen, M. K. (2020). Misallocation and reallocation of resources in vietnamese manufacturing firms. *Journal of Economic Studies*, 47(7):1605–1627.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

ONE (2009). Anexo no.1 a la resolucion no.163/2009. Web Site.

ONEI (2016). Anuario estadistico de cuba 2015, ed. 2016. Website. Accessed 11/6/2022.

ONEI (2017). Gaceta oficial con resolucion no.498/2016. Website. Accessed 11/6/2022.

ONEI (2022a). Series Estadísticas. [Accessed: October 11, 2024]. Table 5.4.

ONEI (2022b). Series Estadísticas. [Accessed: October 11, 2024]. Table 4.2.

ONEI (2022c). Series Estadísticas. [Accessed: October 11, 2024]. Table 6.3.

ONEI (2022d). Series Estadísticas. [Accessed: October 11, 2024]. Table 7.4.

ONEI (2023). Series Estadísticas. [Accessed: October 11, 2024]. Table 7.2.

Palacios, J. C. (2013). Crecimiento del sector productivo cubano entre 1986 y 2009. análisis de sus determinantes y restricciones estructurales. *Revista de la CEPAL*, 110.

Petrin, A. and Sivadasan, J. (2013). Estimating lost output from allocative inefficiency, with an application to chile and firing costs. *Review of Economics and Statistics*, 95(1):286–301.

Posner, R. A. (1978). The chicago school of antitrust analysis. *U. Pa. L. Rev.*, 127:925.

Quiñones, N. and Torres, R. (2015). Crecimiento y déficit comercial sostenible en cuba: el papel de las exportaciones de bienes. *Economía y Desarrollo*, 153:141–158.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720.

Restuccia, D. and Rogerson, R. (2017). The causes and costs of misallocation. *Journal of Economic Perspectives*, 31(3):151–74.

Roland, G. (2000). *Transition and economics*. Cambridge, MA: MIT Press.

Romero, A. (2015). El comercio internacional actual y la inserción externa de países en desarrollo: desafíos para la economía cubana. *Economía y Desarrollo*, 153:190–207.

Sheehan, J., Morris, J., and Hassard, J. (2000). Redundancies in chinese state enterprises: a research report. *Industrial Relations: A Journal of Economy and Society*, 39(3):486–501.

Shy, O. (1996). *Industrial organization: theory and applications*. MIT press.

Svejnar, J. and Terrell, K. D. (1991). *Reducing labor redundancy in state-owned enterprises*, volume 792. World Bank Publications.

Talavera, O., Yin, S., and Zhang, M. (2024). Political motives of excess leverage in state firms. *International Review of Financial Analysis*, 94:103155.

Tamayo, R. (2022). Alejandro gil: En enero, más de 400 empresas originan pérdidas. Sitio web Cubadebate. Consultado el 17/11/2024.

Torres, R. (2013). Structural problems and changes in cuba's economic model. In *No More Free Lunch: Reflections on the Cuban Economic Reform Process and Challenges for Transformation*, pages 5–22. Springer.

Travieso-Diaz, M. F. and Trumbull, C. P. I. (2003). Foreign investment in cuba: Prospects and perils. *George Washington International Law Review*, 35(4):903–946.

Triana, J. (2021). Ied en tiempos de covid-19: ¿qué podemos esperar? In Rosales, H. B. and Cruz, B. A., editors, *Apuntes sobre economía cubana y COVID-19*. Centro de Estudios de la Economía Cubana.

Vidal, A. P. (2017). ¿ qué lugar ocupa la economía cubana en la región?: Una medición a la tasa ppa de las brechas de ingreso y productividad. Technical report, Inter-American Development Bank.

Vidal, P. (2008). La encrucijada de la dualidad monetaria. *Nueva Sociedad*, 216:90–106.

Vidal, P. (2012). *Monetary duality in Cuba: Initial stages and future prospects*, chapter Cuban Economic and Social Development: Policy Reforms and Challenges in the 21st Century, pages 39–54. Harvard University Press.

Wadhwani, S. B. and Wall, M. (1991). A direct test of the efficiency wage model using uk micro-data. *Oxford Economic Papers*, 43(4):529–548.

Weinberger, A. (2020). Markups and misallocation with evidence from exchange rate shocks. *Journal of Development Economics*, 146:102494.

World Bank (1996). China - Reform of State-Owned Enterprises. Economic or Sector Report 14924, World Bank, East Asia and Pacific Region. Disclosed on July 1, 2010.

Yacoub, J. P. F., Hamadeh, N., and Reinhart, C. (2022). The pitfalls of parallel currency markets: higher inflation and lower growth.

Resum

Aquesta tesi aborda el problema de la mala assignació de recursos a les economies en transició, focalitzant l'estudi en Cuba. Tot i que el país va dur a terme algunes reformes orientades a la liberaització del mercat a mitjans dels anys noranta, no va ser fins al 2011 que va començar un procés més significatiu d'obertura econòmica, marcat per l'expansió gradual del sector privat. No obstant això, malgrat els canvis innegables en comparació amb el passat, la reforma del 2011 va ser molt menys exhaustiva que les implementades a, per exemple, la Xina i al Vietnam; i és una de les raons que explica els resultats relativament decebedors de la darrera dècada.

El meu treball examina causes que poden haver impedit una reducció significativa de la mala assignació dels recursos i explicarien la limitada efectivitat de les reformes. L'estudi identifica les distorsions generals (Capítol 2), així com les causes més específiques relacionades amb el mercat de divises (Capítol 3) i el mercat laboral (Capítol 4). Cada capítol ofereix una discussió conceptual, teòrica i empírica d'aquestes qüestions en el context cubà.

El Capítol 2 es basa en el treball seminal de Hsieh and Klenow (2009) per mesurar els costos de benestar associats a la mala assignació de les recursos. L'estudi s'aplica a la indústria manufacturera cubana duranr el període 2004-2017. Segons les estimacions del model, si totes les distorsions s'eliminessin, la productivitat agrega (TFP) augmentaria entre un 45% i un 75%. Aquest resultat subratlla el marge que polítiques estructurals podrien tenir per augmentar la productivitat cubana. A més, el capítol identifica possibles mecanismes que podrien conduir a aquestes millores.

El Capítol 3 es centra en les distorsions en el mercat de divises. Durant anys, Cuba ha mantingut un sistema de tipus de canvi altament peculiar, caracteritzat per múltiples tipus de canvi, segmentació monetària i una moneda oficial sobrevalorada. Aquesta última va crear un sistema de transferències de divises (subvencions i impostos), que es va convertir en l'arrel de múltiples problemes d'eficiència, amb un impacte particularment negatiu en la competitivitat del sector exportador.

En aquest capítol, el model de Hsieh and Klenow (2009) s'amplia per incorporar el tipus de canvi. El model ampliat s'aplica a Cuba utilitzant dades del

2015, amb una bretxa canviària persistent del 2.400%. Els resultats mostren una major dispersió del TFPR quan es té en compte la bretxa canviària, i es documenten guanys de TFP 50 punts percentuals superiors en comparació amb el model base (sense la bretxa). Aquests resultats suggereixen que la política de tipus de canvi és un factor clau per la producctivitat del país.

Finalment, el Capítol 4 investiga les conseqüències d'eficiència de la redundància laboral, un problema persistent en el mercat laboral cubà. Les empreses estatals (SOEs) no tenen control sobre la contractació i l'acomiadament dels seus treballadors—una restricció institucional amb implicacions significatives per a la rendibilitat, competitivitat i salut financera a curt termini d'aquestes empreses.

El capítol utilitza un model de decisió òptima en dues etapes, aplicat a dades del període 2009-2013. Els resultats mostren que, el 2009, el 93% de les empreses estatals del sector manufacturador cubà estaven afectades per aquest problema, amb un 66% de l'ocupació estatal considerada redundant. No obstant això, el 2013 es va observar una reducció significativa de 15 punts percentuals en la taxa de redundància. Aquesta disminució s'atribueix a una combinació de l'augment dels preus, una major disponibilitat de matèries primeres, el creixement de la productivitat i una millora del rendiment financer durant el període. Els resultats suggereixen que restaurar l'autonomia de les empreses per determinar la seva plantilla podria servir com a estímul per a la productivitat. Tot i així, també indiquen que reformes en altres àrees, com la política de preus, podrien ajudar a mitigar les conseqüències de la redundància laboral.

Paraules clau: *Productivitat Total dels Factors, Desalineació del Tipus de Canvi Real, Redundància Laboral, Indústria Manufacturera*