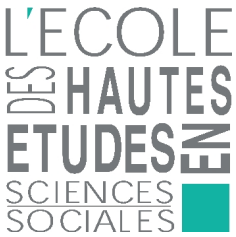


ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
PARIS SCHOOL OF ECONOMICS

Doctoral School nº 465
Économie Panthéon Sorbonne

PhD Thesis

For the Degree of Doctor of Philosophy in Economics

Prepared and defended on October 1st, 2021 by:

Yajna GOVIND

Essays on Inequality and Migration: A Post-Colonial & Global Perspective

Thesis Advisor: Thomas PIKETTY
Co-advisor: Marion LETURCQ

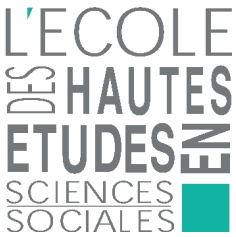
Jury

Rapporteurs

Johan FOURIE	University of Stellenbosch
Ioana MARINESCU	University of Pennsylvania School of Social Policy & Practice

Jury

Janet GORNICK	The Graduate Center, CUNY
Marion LETURCQ	INED
Thomas PIKETTY	EHESS-PSE
Hillel RAPOPORT	University Paris-Sorbonne-PSE



ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
ÉCOLE D'ÉCONOMIE DE PARIS

École Doctorale n° 465
Économie Panthéon Sorbonne

THÈSE

pour l'obtention du grade de docteur en sciences économiques

Présentée et soutenue publiquement le 1er octobre 2021 par :

Yajna GOVIND

Essais sur les inégalités et la migration: Une perspective post-coloniale et globale

Directeur de thèse: Thomas PIKETTY
Co-directrice de thèse: Marion LETURCQ

Membres du Jury

Rapporteurs

Johan FOURIE	Université de Stellenbosch
Ioana MARINESCU	University of Pennsylvania School of Social Policy & Practice

Jury

Janet GORNICK	The Graduate Center, CUNY
Marion LETURCQ	INED
Thomas PIKETTY	EHESS-PSE
Hillel RAPOPORT	Université Paris-Sorbonne-PSE

Acknowledgement

I owe a large part of who I am today to a great deal of people, and circumstances.

First, and foremost, I would probably not have so eagerly pursued a PhD if it was not for the trust and support that my advisor, Thomas Piketty, has provided me. I can appreciate the progress I have made since the day I approached him with my master thesis idea. He has constantly helped me to grow as a researcher, whether it was through feedback on my work, my presentations, or through simple, yet powerful advice[1]. His fervent passion for important questions and his methodological and intellectually stimulating approach to research have, and continue to inspire me. He has been ever-ready to help me, whatever the situation, and for that, I am indebted to him.

I am also grateful to Marion Leturcq, my co-supervisor, who has provided tremendous support from the very beginning. The main qualities (of many) that I had the chance to appreciate during my PhD is her openness to new ideas and discussions, her determination and willingness in helping and finding solutions, her pragmatism, her listening abilities, her kind and supportive words, her constant good mood and humor, and the list could go on. I am indebted to Marion for the technical, and day-to-day aspects of this journey. I hope that we will continue to collaborate in the future. In all, Thomas and Marion have been perfect complements in the supervision of this thesis, and I am grateful to have benefited from both of their supervisions.

I would also like to thank all the members of my jury: Johan Fourie, Janet Gornick, Ioana Marinescu, and Hillel Rapoport. I got the chance to meet Johan at an Economic History seminar at PSE, and I admire his work at Stellenbosch, where I got the chance to present my research. I have crossed Janet's path in several conferences and recently on Twitter, and I have always been in admiration of her achievements. I have long admired Ioana's top-notch work, and could not be happier when she accepted to review my work. I am also thankful to Hillel who has followed my thesis and helped me throughout. I would like to mention that I am thankful to Marco Tabellini who has been a nice support on the job market.

[1] One of which always reminded me to take a step back: "Don't drown in a glass of water!"

With a descending chronological order in mind, I will attempt to group people based on the places or institutions (PSE, INED, Cité U, Delhi, Mauritius, and elsewhere) where I have met them and shared memories with them. Since the line between colleague/fellow PhD/friend is quite blurry, it might not be perfectly organized as such.

Through my different affiliations, I had the chance to have several offices (official and unofficial ones) which meant that I have spent amazing times with many people. The gravitational office has been R4-54, which I like to believe that I was a full member of! Paolo, Sara, Antton, and Caro eventually kindly accepted my (very) frequent visits and chit-chats. I still think that a couch would have been game-changing especially during naptimes! I have spent most of my years with them, and I cannot think of a better or more supportive group. We took turns in being supportive and needing support, and we went through it all, like our emblematic office plant. When I think of that office, I relive our lively, and mainly loud talks, discussions, and laughter (sorry to the neighboring offices!).

Antton has enlivened my days with his great fashion sense and ability to color-match all his outfits. He's been a source of affection, and comfort, much like his cat, when I have needed it. Caro and I share a strong bond, and it is greatly tied to our mutual appreciation for the sun, warmth, and enjoying life! Sara has been an all-rounder: amazing friend, great flatmate, best job-market partner, super efficient seminar co-organizer and so much more (Yes, Julien, you are in here too)! Paolo's inspirational quotes always gave us the answer in difficult times. Grazie a voi per tutti (in the official office language, with only one italian!).

I've also shared great moments with the PSE community at large. Sofi, Martin & Lucio, who since our times at Cité U, I could undoubtedly count on. I'll learn Spanish soon-ish, then we'll enjoy Latino dances even more (and perhaps not Reggeaton anymore)! Alessandro, Cyprien for the sarcastic comments that I got used to, and which showed their constant appreciation. Grazie Francesco for the political discussions on Italy! Cyp and Paul for the nice (and very varied) singing repertoire while working in the CASD room. The World Inequality Lab team: Alvaro, Amory, Aurélie, Flore, Jeanne, Li (& Lisa), Lydia, Marc(k), Mathilde, Mauri, Morten, Nacho & Nitin (my ping-pong partners), Olivia, Oscar, Rowaida, Santi, Theresa, Tom(x2). This affiliation has also allowed me to enjoy one of the fun aspects of research, which is collaborating. I have thoroughly enjoyed working with Luis, Filip, and Daniel. A special thanks to Clara and Clementine for their tremendous support and mentorship during the job market and beyond. To Véronique Guillotin, without whom the

organization of the PhD lives and job market would be chaotic. Thanks to the job market covid-year crew! A sweet thought to the "Jourdan à Hendaye" team, plus Laura, Pepe, and Clément. To Giulio, Emma, Sarah x2, Caroline, Melanie, Juli, Victor, Helene, Morgan, Naomie, and the PSE PhD community for the nice times. To the M2 PPD "the-true-Nolay" crew. I have also enjoyed my days at INED, which has allowed to me make wonderful encounters, among whom, let me mention P-A, Marine, and Constance, with whom I have shared the office. I was moved by their support, and I largely owe them my successful naturalization.

I have also crossed the paths of many in Boston, Paris, at Cité U, the OECD, Delhi University, and in Mauritius, who are still important parts of my life: Elika, and *La Cooloc* for the homely feeling and fun together; Andrea, to the many spontaneous encounters; Aleksandra, to the care and attention she gave me; Miguel, to the amazing philosophical discussions; Julian, for the amazing friendship we share; Prash, Keshav, Krish, Nevin, Sujan, and Vivek, to the Delhi/Indian adventures and the cruise trip we should plan soon! To Mansa and Shaima, for our friendship that goes back to as far as I can remember.

I would also never have become the person I am today without the unconditional support of my family. We ended up being as physically far away as humanly possible, and yet, we have such an incredible closeness! To Daksha, for her generosity and kindness; Bhavna for having the right words to uplift me, to Aradhya (hurray!) & Sudeesh. To Mum who has been my role model, a strong pillar, and an unshakeable support. To Dad, for his humble smile, his devotion to our success, the champion of "dad jokes", & the person who would have taken the most pride in calling me "Dr. Govind". I hope you are smiling at this achievement of yours, wherever you are. To my spiritual master, whose guidance is always present.

Last, I'd like to thank Paolo who has been a part of most of the above - as a best friend, colleague, cheerleader, mentor, travel companion, and more. This section risks to be longer than the thesis if I had to put his contributions into words, so I'll simply leave it to grazie for these amazing years of support, from the "Dai, you can do this!" to the heavenly-cooked meals (mostly pasta with many ingredients). ???[2]

While writing this section, I ran out of synonyms for the word "support" and this gives a good sense of the overwhelming encouragement that I have received throughout this journey. I am so grateful to many more!

Copenhagen, 26/08/2021

[2]These question marks are intentionally left here.

Pu twa papa, pu to blessings, to lamour, ek to fierte...JGD

Summary

This PhD dissertation explores different aspects of inequality and migration.

Chapter 1 looks at the causal effect of naturalization on the labor market integration of foreigners. It is acknowledged that better integration is beneficial for both migrants and the host country. In this respect, granting citizenship could be an important policy to boost migrants' integration. In this chapter, I estimate the causal impact of obtaining citizenship on migrants' labor market integration. I exploit a change in the law of naturalization through marriage in France in 2006. This reform amended the eligibility criteria for applicants by increasing the required number of years of marital life from 2 to 4, generating an exogenous shock and thus a quasi-experimental setting. Using administrative panel data, and a difference-in-differences approach, I estimate the labor market returns to naturalization. I find that, among those working, citizenship leads to an increase in annual earnings. While the gain in earnings is similar for both men and women, the effect for men is mostly driven by an increase in hours worked compared to an increase in hourly wages for women. I provide suggestive evidence that naturalization helps reduce informality and discrimination. This chapter thus provides strong evidence that naturalization acts as a catalyst for labor market integration.

Chapter 2 studies the post-colonial trends of income inequality in four ex-French colonies. Most ex-colonies have gained their independence during the decolonization wave in the last century. Recent research on the colonial legacy in terms of inequality has thus mostly focused on these independent states, overlooking the few territories which were assimilated by their ex-colonizers. This chapter analyzes the post-colonial inequality in four such territories- La Reunion, Guadeloupe, Martinique, and Guyane. Drawing on a new income tax dataset put together in this chapter, I study the evolution of income inequality since their decolonization in 1946 until recent years. The results of the top 1% income shares reflect a rapid decline of inequality since

decolonization and stabilization in the recent decade. Despite the general catch-up of the overseas departments, the top 10% income share remained consistently higher than in the metropolis. Going further, I investigate the hidden underlying cleavage: the metropolitan-native divide. Using administrative data, I show that metropolitans are over-represented at the top of the distribution and that there exists a "metropolitan income premium" in the overseas departments, even after controlling for observable characteristics.

Chapter 3 is joint work with Luis Bauluz, Filip Novokmet, and Daniel Sanchez Ordoñez. It aims at measuring land inequality in a large variety of countries across different regions. It is known that agricultural land is vital for three out of four of the poorest billion individuals in the world yet little is known about its distribution. Existing cross-country estimates of land inequality, based on agriculture census data, measure the size distribution of agricultural holdings. These neither reflect land ownership inequality nor value inequality and often do not account for the landless population. In this chapter, we tackle these issues and provide novel and consistent estimates of land inequality across countries, based on household surveys. We show that i) land-value inequality can differ significantly from land-area inequality, ii) differences in the proportion of landless across countries vary substantially, affecting markedly inequality estimates and, iii) regional patterns in inequality according to our benchmark metric contradict existing estimates from agricultural censuses. Overall, South Asia and Latin America exhibit the highest inequality with the top 10% landowners capturing up to 75% of agricultural land, followed by Africa and 'Communist' Asia (China and Vietnam) at levels around 55-60%.

Résumé

Cette thèse s'intéresse aux liens entre les inégalités, la migration et la colonisation.

Le **premier chapitre** s'intéresse à l'effet causal de la naturalisation sur l'intégration des étrangers sur le marché du travail suite à l'obtention de la citoyenneté. J'exploite un changement dans la loi de naturalisation par mariage en France en 2006. Cette réforme a modifié les critères d'éligibilité des candidats en augmentant le nombre d'années de vie commune requis de 2 à 4 ans, générant un choc exogène et un cadre quasi-expérimental. En utilisant des données administratives de panel et une approche de différence-de-différences, j'estime le rendement de la naturalisation sur le marché du travail. Je constate que, parmi les personnes qui travaillent, la citoyenneté entraîne une augmentation des revenus annuels. Je démontre que la naturalisation contribue à réduire l'emploi informel et la discrimination. Ce chapitre permet de conclure que la naturalisation agit comme un catalyseur de l'intégration sur le marché du travail.

Le **deuxième chapitre** étudie les tendances post-coloniales de l'inégalité des revenus dans quatre anciennes colonies françaises. Les recherches récentes sur l'héritage colonial en termes d'inégalité se sont principalement concentrées sur des États indépendants, négligeant les quelques territoires qui ont été assimilés par leurs anciens colonisateurs. Ce chapitre analyse l'inégalité post-coloniale dans quatre de ces territoires : La Réunion, la Guadeloupe, la Martinique, et la Guyane. En m'appuyant sur un nouvel ensemble de données sur l'impôt sur le revenu constitué dans ce chapitre, j'étudie l'évolution de l'inégalité des revenus depuis leur décolonisation en 1946 jusqu'à ces dernières années. Les résultats de la part des revenus des 1% supérieurs reflètent un déclin rapide des inégalités depuis la décolonisation et une stabilisation au cours de la dernière décennie. Malgré le rattrapage général des départements d'outre-mer, la part des 10% de revenus les plus élevés est restée constamment supérieure à celle de la métropole. En allant plus loin, j'étudie le clivage métropole-natifs. En utilisant des données administratives, je montre que les métropolitains sont surreprésentés dans le

haut de la distribution et qu'il existe une "prime de revenu métropolitaine" dans les départements d'outre-mer, même en contrôlant les caractéristiques observables.

Le **troisième chapitre** est un travail en collaboration avec Luis Bauluz, Filip Novokmet, et Daniel Sanchez Ordoñez. Il vise à mesurer l'inégalité des terres dans plusieurs pays à travers différentes régions du monde. Les estimations existant de l'inégalité des terres, basées sur les données du recensement agricole, mesurent la distribution de l'utilisation des exploitations agricoles. Elles ne reflètent ni les inégalités de la possession des terres, ni les inégalités de la valeur des terres et, souvent, elles ne tiennent pas compte de la population sans terres. Dans ce chapitre, nous abordons ces questions et fournissons des estimations nouvelles et cohérentes de l'inégalité des terres entre les pays, basées sur des enquêtes auprès des ménages. Nous montrons que i) l'inégalité de la valeur des terres peut différer considérablement de l'inégalité de la superficie des terres, ii) les différences dans la proportion de sans-terre entre les pays varient considérablement, ce qui affecte fortement les estimations de l'inégalité et, iii) les niveaux d'inégalité régionaux selon notre métrique de référence contredisent les estimations existantes des recensements agricoles. Dans l'ensemble, l'Asie du Sud et l'Amérique latine présentent les inégalités les plus fortes, les 10 % de propriétaires fonciers les plus importants s'appropriant jusqu'à 75 % des terres agricoles, suivies de l'Afrique et de l'Asie " communiste " (Chine et Vietnam), avec des niveaux d'environ 55-60 %.

Table of Contents

List of Tables

List of Figures

General Introduction

"All animals are equal, but some animals are more equal than others"
George Orwell, *Animal Farm, 1945*

0.1 Overview

Studies on inequality are gaining momentum in recent years, with the rise in disparities in all regions of the world and an increasing awareness of its consequences. Inequality can be present in different forms: income, wealth, gender, race, discrimination and social mobility among others. This thesis focuses on the intricate link between inequality, migration, and colonization, by looking at three main research questions: the integration of foreigners in the host country, the post-colonial level of inequality in ex-colonies that remained part of their ex-colonizers, and finally, land inequality in developing countries.

The link between colonization and inequality is pretty straight-forward. Colonization, in fact, represents an extreme form of inequality, and its consequences can last in the post-colonial period (Frankema, 2010; Nunn, 2008; Engerman and Sokoloff, 2002; Engerman and Sokoloff, 2005). This can take the form of persisting unequal access to resources and opportunities, both among the native population and between them and their ex-colonizers. This is particularly relevant and striking in the cases of colonial territories which have remained part of their ex-colonizers.

Colonization and migration are also closely linked. Colonization represents a domination by a foreign class, adding new ethnic fractures to the society. This social structure may also be perpetuated in the post-colonial era, albeit to a lesser extent. Often, ex-colonizers are at the top of the socioeconomic ladder, while natives remain at the bottom. In ex-colonies that remain part of the metropolis, this is exacerbated by the continued internal migration flows between the metropolis and these territories. In addition to this direct link between colonization and migration, very often these colonial heritages or, more generally, unequal access to opportunities, are also at the

root of the South-North migration waves.

Finally, there exists an direct relationship between migration and inequality. Inequality can in itself foster migration (Chiquiar and Hanson, 2005). Developed countries have been impacted by various migration waves which pose many challenges, and the successful integration of foreigners is a crucial one. In the face of socioeconomic difficulties, migrants move to developed countries with the aim of securing a better future for themselves and their families. However, they are often confronted with barriers in the form of discrimination or hostilities. Migrants thus remain at a disadvantage in the host countries, with a persisting gap in their socio-economic situation as compared to natives (Chiswick, 1978).

In this thesis, I approach these questions from different angles to shed new light on the link between inequality, migration, and colonization. Among the different facets of these links, this thesis will focus on the unequal situations between different groups: inequality between migrants and natives, between ex-colonizers and ex-colonized, and finally, between those with different ownership of land. A crucial step in tackling inequality is to have a good measure of the problem at hand. While there is renewed interest in the measurement of inequality, some areas are still not fully covered. This thesis fills the gap in two such strands of the literature- the distribution of land, and economic inequality in post-colonial contexts. In particular, the novelty of the measurement exercises in these two chapters is to exploit different data sources, household surveys, or archival fiscal data — that have been under-utilized when looking at these questions. Once the stylized facts established, it is important to understand the underlying issues. Combining the descriptive analysis with administrative data, I attempt to uncover the peculiarities of inequality in the post-colonial settings. Finally, it is also important to establish causal relationships to provide the most relevant policy advice. This is the aim of the chapter on the integration of foreigners.

0.2 Thesis Structure and Contributions

This thesis is organized around three chapters that analyze these three themes from a post-colonial and global perspective using a variety of data sources and methodologies. The focus of chapter 1 is to understand whether the legal status of migrants can play a role in providing migrants with different labor market opportunities. In particular, I look at whether obtaining the nationality of the host country can help foreigners better integrate the labor market. In the second chapter, I estimate the post-colonial evolution of income inequality in the overseas departments of France, from the 1950s

until today. Finally, in the last chapter, together with L. Bauluz, F.Novokmet, and D. Sanchez, we revisit the existing estimates and provide new measurements of land inequality in developing countries. In all, this thesis sheds light on three major challenges in today's world- inequality, colonial legacies and its implications, and the movement of population.

In order to tackle these questions, this thesis analyzes inequality in various contexts spanning a large geographical area. In the first chapter, I look at the integration of foreigners in the host country, France. In the second chapter, I study the long-run post-colonial evolution of income inequality in ex-French colonies. Finally, in the last chapter, I estimate land inequality in a variety of developing countries.

Apart from the large geographical scope, another interesting feature of this thesis is the use of a variety of data sources. In the first chapter, to study the research question of the integration of foreigners, I use rich French administrative data. It is a panel dataset that allows the matching of information from the civil registries to the population census as well as the employment data. This data source, despite representing a rich source, has been underutilized in the study of the integration of foreigners. In the second chapter, I put together a novel fiscal database at the regional-level, from archival sources, to estimate income distribution since decolonization. I combine this with administrative fiscal data in the second part of the chapter to explore some of the underlying factors that could explain the level of inequality observed in these territories. Taking an historical perspective gives a better sense of the evolution of inequality. Finally, in the last chapter, we depart from the existing literature by using household surveys to measure land ownership inequality in a novel way. Each of these data sources have their own advantages and limitations, but they all provide valuable resources to tackle different dimensions of inequality.

The different chapters also exploit different analytical tools. In the first chapter, I adopt a difference-in-differences approach, exploiting the 2006 reform of the naturalization rules by marriage in France, to understand the causal effects of naturalization on the labor market integration of foreigners. While a positive association between naturalization and labor market situation has been established in the literature, it is crucial from a policy and political perspective to understand whether this link is causal or not. Beyond establishing causal inferences, the precise measurement of inequality is also crucial. Indeed, inequality might be perceived by all, but its measurement brings concrete evidence on the matter. For instance, it is well-known that the level of inequality in these ex-French colonies are very high, but so far the long-run evidence were sparse. This is studied in the second chapter. In the context

of land inequality, the standard definition of land inequality is revisited to provide a more relevant and encompassing definition, and accordingly, different estimates of land inequality are provided.

0.3 Naturalization

With each subsequent migration wave in developed countries, the question of foreigners' integration increasingly gains importance. It is known that the immigrant population largely suffers from high unemployment rates, poor economic situations, and face discrimination. This is problematic for migrants themselves and can, in addition, lead to hostility and foster anti-immigrant feelings among the native population. The successful integration of foreigners into the host societies is thus a crucial objective of policymakers. Lack of legal status tends to be a major barrier for the social, economic and political integration of foreigners (Monras, Vázquez-Grenno, and Elias Moreno, 2018; Bahar, Ibáñez, and Rozo, 2021; Pinotti, 2017; Dustmann, Fasani, and Speciale, 2017). Hence, granting of the host country's nationality can be a potential tool for integration. The fierce debate surrounding it is whether naturalization is only a reward for achieving integration, or whether it can in itself be a boost to integration (Hainmueller, Hangartner, and Pietrantuono, 2017).

In this chapter, I estimate the causal effect of naturalization on labor market outcomes, exploiting a national-level reform in the law of naturalization through marriage in France which occurred in 2006. The reform amended the eligibility criteria of applicants by increasing the required number of years of marital life from 2 to 4. De facto, this reform impacted foreigners that were married before and after 2004 differently, as the former could apply as early as within two years of marriage while the latter were constrained to wait four years of marriage. In other words, foreigners in two otherwise similar cohorts of marriage, expecting to naturalize within the same number of years, unexpectedly face a different length of waiting time to be eligible to naturalize.

Using a rich administrative French panel data, the Permanent Demographic Sample (EDP), I first document the differential propensity to be naturalized in the years following marriage between those facing a two-year waiting period (early-treated group) compared to those waiting four years (late-treated group). I exploit this unexpected change in the eligibility rule, and the resulting impact on these two groups, using a difference-in-differences approach, by comparing their labor market situations before and after they are likely to be naturalized. The results suggest that naturalization is followed by an increase in net annual earnings, explained by an

increase in the number of hours worked and hourly wages, with differences based on gender, age and professional categories. I explore the underlying channels that could explain the benefits of naturalization. While unrestricted access to the labor market, as proxied by public sector employment, does not seem to have played a role, the results suggest that naturalization helps reducing discrimination and informality. The nationality of the host country seems to be used as a signaling device in the labor market for integration and language proficiency. Additionally, naturalization leads to higher declared number of hours worked in sectors that are highly impacted by informal employment.

In all, this underlines the importance of naturalization for the economic integration of foreigners, with causal evidence. This debate on the benefits and disadvantages of restricting naturalization, being mostly a political one in nature, needs to better account for the fact that naturalization can in itself boost the economic integration of foreigners. Hence, if the aim of host countries is to better integrate foreigners, easier access to citizenship could potentially be a way to hasten the path to integration.

0.4 Post-colonial Income inequality

The colonial social structure was founded on very unequal distribution of resources, and with decolonization, territories have attempted to do away with this colonial legacy. The literature has looked at the level of inequality in ex-colonies, both during and after colonization (Atkinson, 2014; Atkinson, 2015a; Atkinson, 2015b; Atkinson, 2015c; Alvaredo and Atkinson, 2010; Alvaredo, Cogneau, and Piketty, 2021). However, not much is known about the situation in ex-colonies that have remained part of the metropolis. This is the case of the ex-colonies of France, now overseas departments, namely La Reunion, Guadeloupe, Martinique and Guyane, which have centuries of colonial history before being at par with other French metropolitan regions. While it is generally accepted that these territories are among the most unequal regions in France today, the measurement of the level of inequality since decolonization can provide a good overview of the picture to allow us to better grasp today's realities.

In this chapter, I build a novel fiscal dataset at the regional-level in France that allows me to estimate the income distribution in each of these four territories since their decolonization in the 1950s until 2014. By doing so, I provide a comprehensive picture of the long-run post-colonial evolution of economic inequality in these territories, and perform a comparative analysis of the situation in these territories compared to the metropolis. Combining archival and contemporary sources of fiscal data with

population census and income data, I first estimate the income shares accruing to the richest groups in these territories. I find that there was a small increase in inequality in the immediate post-decolonization period, followed by a large decline in inequality until the beginning of the 21st century. The top 1% has since stabilized at a level that is comparable to the metropolis, while the top 10% income shares are systematically above the ones of metropolitan France. In the second part of the paper, I use administrative fiscal data, the EDP, to uncover some specificities of these territories. I show that the public sector is overly-important, and discuss its implications for inequality. Additionally, I show that metropolitans occupy the highest-paying jobs in these territories. This intra-national migration pattern is even more relevant as these territories are already marked by ethnic disparities as a result of its colonial past.

In this chapter, by collecting data and measuring the post-colonial level of inequality in these overseas departments of France, I shed light on the inequality pattern after decolonization in territories that remained part of their ex-colonizers. I show that there has been a convergence in the top 1% but that this is not the case for the top 10% income shares between metropolitan France and the overseas departments. While these territories have benefitted from their link to the metropolis, they also have some peculiarities, given their situations, that have not helped reducing, or at worse exacerbated the inequality levels.

0.5 Land Inequality

A vast majority of the developing world depends on land for their livelihood. Lack of access to land hampers development, growth and may perpetuate poverty (Deininger and Squire, 1998). Unequal distribution of land is thus a crucial issue in these regions. In order to understand its role, having a precise and consistent measure of land ownership inequality is a starting point. In this paper, along with F.Novokmet, L.Bauluz and D.Sanchez, we revisit the existing literature, and provide novel estimates of land inequality. These estimates largely relied on census data to measure land inequality. However, censuses, by construction, measure landholdings rather than land owned. In addition, censuses do not measure differentials in land value, or quality, and do not account for the landless population.

We depart from the existing literature by using household surveys[3] rather than census data. This allows us to overcome some limitations faced by estimates relying on

[3]The World Bank's LSMS or other surveys overseen by the countries

census data when estimating land ownership inequality. We first re-estimate land inequality as defined by the existing literature, using the most up-to-date census rounds. We then estimate land area inequality from the survey and compare those to the estimates from the census. We find that the two are very similar. Going a step further, we estimate land value inequality and finally take into account the landless populations. Our result points to the importance of taking into account the different definitions. Our most preferred definition, the land value inequality including the landless, uncover some interesting regional patterns: South Asia and Latin America are among the most unequal regions, followed by Africa and "Communist"-Asia.

In a nutshell, in this article, we aim at providing a consistent measure of land inequality as a first step in analyzing its role in the development process of these territories. We put forward the need to define the different concepts used and the usefulness of household surveys in estimating land ownership inequality in terms of area, value, and accounting for a marginalized, yet very relevant section of the population, the landless households.

Introduction Générale

"Tous les animaux sont égaux, mais certains sont plus égaux que d'autres" George Orwell, *Animal Farm, 1945*

0.6 Synthèse

Les études sur les inégalités ont pris de l'ampleur ces dernières années, avec la montée des disparités dans toutes les régions du monde et une prise de conscience croissante de leurs conséquences. Les inégalités peuvent se présenter sous différentes formes : en termes de revenu, de richesse, du genre, de la race, de la discrimination ou de la mobilité sociale, entre autres. Cette thèse s'intéresse aux liens complexes entre les inégalités, la migration et la colonisation, en examinant trois questions de recherche principales : l'intégration des étrangers dans le pays d'accueil, le niveau d'inégalité post-colonial dans les anciennes colonies qui sont restées rattachée à leurs anciens colonisateurs, et enfin, les inégalités de terres dans les pays en développement.

Le lien entre colonisation et inégalité est assez simple. En effet, la colonisation représente une forme extrême d'inégalité et ses conséquences peuvent perdurer dans la période post-coloniale (Frankema, 2010; Nunn, 2008; Engerman and Sokoloff, 2002; Engerman and Sokoloff, 2005). Cela peut se traduire par la persistance des inégalités d'accès aux ressources et aux opportunités, tant au sein de la population native, qu'entre celle-ci et ses ancien colonisateurs. Ceci est particulièrement pertinent et frappant dans le cas de territoires coloniaux qui sont restés rattachés à leurs anciens colonisateurs.

La colonisation et la migration sont également étroitement liées. La colonisation représente une domination par une classe étrangère, ajoutant de nouvelles fractures ethniques à la société locale. Cette structure sociale se perpétue souvent dans l'ère post-coloniale bien que sous une forme amoindrie. Souvent, les anciens colonisateurs se trouvent en haut de l'échelle socio-économique, tandis que les indigènes restent en bas de l'échelle. Dans les anciennes colonies qui font toujours partie de la métropole,

cette situation est exacerbée par les flux migratoires internes continus entre la métropole et ces territoires. Outre ce lien direct entre la colonisation et la migration, bien souvent ces héritages coloniaux ou, plus généralement, l'inégalité d'accès aux opportunités, sont également à l'origine de vagues migratoires Sud-Nord.

Enfin, il existe un lien direct entre la migration et les inégalités. Les inégalités peuvent en soi favoriser la migration (Chiquiar and Hanson, 2005). Les pays développés ont été touchés par diverses vagues de migration qui posent de nombreux défis, et l'intégration réussie des étrangers en est un crucial. Face aux difficultés socio-économiques, les migrants se déplacent vers les pays développés dans le but d'assurer un meilleur avenir pour leurs familles et eux-mêmes. Cependant, ils sont souvent confrontés à des obstacles sous forme de discrimination ou d'hostilité. Les migrants restent donc désavantagés dans leur pays d'accueil, avec un écart persistant dans leur situation socio-économique par rapport aux natifs (Chiswick, 1978).

Dans cette thèse, j'aborde ces questions sous différents angles afin d'apporter un nouvel éclairage sur le lien entre les inégalités, la migration et la colonisation. Parmi les différentes facettes de ces liens, cette thèse se concentrera sur les inégalités entre différents groupes : les inégalités entre les migrants et les natifs, entre les anciens colonisateurs et les anciens colonisés, et enfin, entre ceux qui possèdent des terres et ceux qui n'y ont pas accès. Une étape importante dans la lutte contre les inégalités est de disposer d'une bonne mesure du problème. Bien qu'il y ait un regain d'intérêt pour la mesure des inégalités, certains domaines ne sont toujours pas entièrement examinés. Cette thèse vise à contribuer à deux thématiques : la distribution des terres et les inégalités économiques dans les contextes post-coloniaux. En particulier, la nouveauté des exercices de mesure dans ces deux chapitres est d'exploiter différentes sources de données - les enquêtes auprès des ménages et des données fiscales d'archives - qui ont été sous-utilisées pour répondre à ces questions. Une fois les faits établis, il est important de comprendre les canaux qui peuvent expliquer ces tendances. En combinant une analyse descriptive avec des données administratives, je démontre les particularités des inégalités dans les contextes post-coloniaux. Enfin, il est également important d'établir des relations causales afin de guider les politiques publiques. C'est l'objectif du chapitre sur l'intégration des étrangers.

0.7 Structure de la thèse et contributions

Cette thèse est organisée autour de trois chapitres qui analysent ces trois thématiques d'un point de vue postcolonial et global en utilisant plusieurs de sources de données, et de méthodologies. L'objectif du premier chapitre est de comprendre si le statut

légal des migrants peut jouer un rôle en donnant plus d'opportunités aux migrants sur le marché du travail. En particulier, je cherche à savoir si l'obtention de la nationalité du pays d'accueil peut aider les étrangers à mieux s'intégrer sur le marché du travail. Dans le deuxième chapitre, j'estime l'évolution post-coloniale des inégalités de revenus dans les départements d'outre-mer de France, des années 1950 à aujourd'hui. Enfin, dans le dernier chapitre, avec L. Bauluz, F.Novokmet, et D. Sanchez, nous revisitons les estimations existantes et fournissons de nouvelles mesures d'inégalité des terres dans les pays en développement. Au total, cette thèse apporte un éclairage sur trois défis majeurs du monde actuel : les inégalités, les héritages coloniaux et leurs implications, et les mouvements de population. Afin d'aborder ces questions, cette thèse analyse les inégalités dans différents contextes couvrant une large zone géographique. Dans le premier chapitre, je m'intéresse à l'intégration des étrangers dans le pays d'accueil, la France. Dans le deuxième chapitre, j'étudie l'évolution post-coloniale à long terme des inégalités de revenus dans les anciennes colonies françaises. Enfin, dans le dernier chapitre, j'estime l'inégalité des terres dans plusieurs pays en développement.

Outre la grande portée géographique, une autre caractéristique intéressante de cette thèse est l'utilisation de plusieurs sources de données. Dans le premier chapitre, pour étudier la question de recherche de l'intégration des étrangers, j'utilise de riches données administratives françaises. Il s'agit d'un ensemble de données de panel qui permet d'apparier les informations d'état civil au recensement de la population ainsi qu'aux données de l'emploi. Cette source de données, bien que représentant une source riche, a été peu utilisée dans l'étude de l'intégration des étrangers. Dans le deuxième chapitre, j'ai constitué une nouvelle base de données fiscales au niveau régional, à partir de sources d'archives, afin d'estimer la distribution des revenus depuis la décolonisation. Je combine cette base avec des données fiscales administratives dans la deuxième partie du chapitre pour explorer certains des facteurs qui pourraient expliquer le niveau d'inégalité observé dans ces territoires. L'adoption d'une perspective historique donne une meilleure idée de l'évolution des inégalités. Enfin, dans le dernier chapitre, nous contribuons à la littérature existante en utilisant les enquêtes auprès des ménages pour mesurer les inégalités de la propriété des terres. Chacune de ces sources de données a ses propres avantages et limites mais elles fournissent toutes des ressources précieuses pour aborder différentes dimensions des inégalités.

Les différents chapitres exploitent également différents outils analytiques. Dans le premier chapitre, j'adopte une approche de différence des différences, en exploitant la réforme de 2006 qui change les règles de naturalisation par mariage en France, pour

comprendre les effets causaux de la naturalisation sur l'intégration des étrangers sur le marché du travail. Alors qu'une relation positive entre la naturalisation et la situation sur le marché du travail a été établie dans la littérature, il est crucial d'un point de vue de politique publique de comprendre si ce lien est causal ou non. Au-delà de l'établissement d'une analyse causale, la mesure précise des inégalités est également cruciale. En effet, les inégalités peuvent être perçues par tous, mais sa mesure apporte des preuves concrètes sur le sujet. Par exemple, il est bien connu que le niveau d'inégalité dans ces anciennes colonies françaises est très élevé, mais jusqu'à présent, les preuves à long terme sont rares. Ce point est étudié dans le deuxième chapitre. Dans le contexte des inégalités de répartition des terres, la définition standard des inégalités de répartition des terres est revisitée pour fournir une définition plus pertinente et plus englobante, et en conséquence, différentes estimations des inégalités de répartition des terres sont fournies.

0.8 Naturalisation

Avec chaque nouvelle vague de migration dans les pays développés, la question de l'intégration des étrangers prend de plus en plus d'importance. On sait que la population immigrée souffre en grande partie de taux de chômage élevés, de mauvaises situations économiques et qu'elle est confrontée à la discrimination. Cette situation est problématique pour les migrants eux-mêmes et peut, en outre, engendrer de l'hostilité et nourrir des sentiments anti-immigrants au sein de la population native. L'intégration réussie des étrangers dans les sociétés d'accueil est donc un objectif crucial des décideurs politiques. L'absence de statut juridique tend à être un obstacle majeur à l'intégration sociale, économique et politique des étrangers (Monras, Vázquez-Grenno, and Elias Moreno, 2018; Bahar, Ibáñez, and Rozo, 2021; Pinotti, 2017; Dustmann, Fasani, and Speciale, 2017). L'octroi de la nationalité du pays d'accueil peut donc être un outil potentiel d'intégration. Le débat qui l'entoure est de savoir si la naturalisation n'est qu'une récompense pour la réalisation de l'intégration, ou si elle peut en soi être un catalyseur d'intégration (Hainmueller, Hangartner, and Pietrantuono, 2017).

Dans ce chapitre, j'estime l'effet causal de la naturalisation sur la situation sur le marché du travail, en exploitant une réforme au niveau national de la loi de naturalisation par le mariage en France qui a eu lieu en 2006. Cette réforme a modifié les critères d'éligibilité des candidats en augmentant le nombre d'années de vie maritale requis de 2 à 4. De facto, cette réforme a eu un impact différent sur les étrangers qui se sont mariés avant et après 2004, puisque les premiers pouvaient faire

une demande dès deux ans de mariage alors que les seconds devaient attendre quatre ans de mariage. En d'autres termes, les étrangers de deux cohortes de mariage par ailleurs similaires, qui s'attendaient à être naturalisés après le même nombre d'années, ont été confrontés de manière inattendue à une durée d'attente différente pour pouvoir être naturalisés. En utilisant un riche panel de données administratives françaises, l'Échantillon Démographique Permanent (EDP), je documente d'abord la propension différentielle à être naturalisé dans les années suivant le mariage entre ceux qui font face à une période d'attente de deux ans (groupe traité tôt) et ceux qui attendent quatre ans (groupe traité tard). J'exploite ce changement inattendu dans la règle d'éligibilité, et l'impact qui en résulte sur ces deux groupes, en utilisant une approche de différence-de-différences, en comparant leurs situations sur le marché du travail avant et après qu'ils soient susceptibles d'être naturalisés. Les résultats suggèrent que la naturalisation est suivie d'une augmentation du revenu annuel net, expliquée par une augmentation du nombre d'heures travaillées et des salaires horaires, avec des différences en fonction du sexe, de l'âge et des catégories professionnelles. J'explore les canaux qui pourraient expliquer les avantages de la naturalisation. Alors que l'accès illimité au marché du travail, tel que représenté par l'emploi dans le secteur public, ne semble pas avoir joué un rôle, les résultats suggèrent que la naturalisation contribue à réduire la discrimination et l'informalité. La nationalité du pays d'accueil semble être utilisée comme un signal sur le marché du travail pour l'intégration et la maîtrise de la langue. En outre, la naturalisation entraîne une augmentation du nombre déclaré d'heures travaillées dans les secteurs fortement touchés par l'emploi informel.

Cela souligne l'importance de la naturalisation pour l'intégration économique des étrangers, cela démontrée de manière causale. Le débat sur les avantages et les inconvénients de la restriction de la naturalisation, qui est essentiellement de nature politique, doit mieux prendre en compte le fait que la naturalisation peut en soi stimuler l'intégration économique des étrangers. Par conséquent, si l'objectif des pays d'accueil est de mieux intégrer les étrangers, un accès plus facile à la citoyenneté pourrait potentiellement être un moyen d'accélérer l'intégration.

0.9 Inégalité des revenus post-coloniaux

La structure sociale coloniale était fondée sur une distribution très inégale des ressources, et avec la décolonisation, les territoires ont tenté de se débarrasser de cet héritage colonial. La littérature s'est penchée sur le niveau d'inégalité dans les anciennes colonies, tant pendant qu'après la colonisation (Atkinson, 2014; Atkinson,

2015a; Atkinson, 2015b; Atkinson, 2015c; Alvaredo and Atkinson, 2010; Alvaredo, Cogneau, and Piketty, 2021). Cependant, on ne sait pas grand-chose de la situation dans les anciennes colonies qui sont restées rattachées à la métropole. C'est le cas des anciennes colonies de France, aujourd'hui départements d'outre-mer, à savoir la Réunion, la Guadeloupe, la Martinique et la Guyane, qui ont des siècles d'histoire coloniale avant d'obtenir le même statut que d'autres régions métropolitaines françaises. S'il est généralement admis que ces territoires font actuellement partie des régions les plus inégalitaires de France, la mesure du niveau des inégalités depuis la décolonisation peut fournir un bon aperçu du tableau pour nous permettre de mieux appréhender les réalités actuelles.

Dans ce chapitre, je construis un nouvel ensemble de données fiscales au niveau régional en France qui me permet d'estimer la distribution des revenus dans chacun de ces quatre territoires depuis leur décolonisation dans les années 1950 jusqu'en 2014. Ce faisant, je fournis une image complète de l'évolution post-coloniale à long terme des inégalités économiques dans ces territoires, et je réalise une analyse comparative de la situation de ces territoires par rapport à la métropole. En combinant des sources d'archives et contemporaines de données fiscales avec des données de recensement de la population et des données sur les revenus, j'estime d'abord les parts de revenus revenant aux groupes les plus riches dans ces territoires. Les résultats montrent qu'il y a eu une légère augmentation des inégalités dans la période suivant immédiatement la décolonisation, suivie d'une baisse significative des inégalités jusqu'au début du 21ème siècle. Les 1% supérieurs se sont depuis stabilisés à un niveau comparable à celui de la métropole, tandis que les parts de revenu des 10% supérieurs sont systématiquement supérieures à celles de la France métropolitaine. Dans la seconde partie de l'article, j'utilise des données administratives fiscales, l'EDP[4], pour mettre en évidence certaines spécificités de ces territoires. Je montre que le secteur public est très important, et je discute des implications de ce fait sur les inégalités. De plus, je montre que les métropolitains occupent les emplois les mieux rémunérés dans ces territoires. Ce modèle de migration intra-nationale est d'autant plus pertinent que ces territoires sont déjà marqués par des disparités ethniques résultant de leur passé colonial.

Dans ce chapitre, en collectant des données et en mesurant le niveau d'inégalité post-colonial dans ces départements d'outre-mer de la France, je mets en lumière les inégalités après la décolonisation dans les territoires qui sont restés rattachés à leurs anciens colonisateurs. Les résultats montrent qu'il y a eu une convergence dans les 1% supérieurs mais que ce n'est pas le cas pour les parts de revenu des

[4] l'Echantillon Démographique Permanent

10% supérieurs entre la France métropolitaine et les départements d'outre-mer. Si ces territoires ont bénéficié de leur lien avec la métropole, ils présentent aussi des particularités, compte tenu de leur situation, qui n'ont pas permis de réduire, ou, pire, ont contribué à exacerber les niveaux des inégalités.

0.10 Inégalité des terres

Une grande majorité du monde en développement dépend de la terre pour sa subsistance. Le manque d'accès à la terre entrave le développement, et la croissance et peut perpétuer la pauvreté (Deininger and Squire, 1998). La répartition inégale des terres est donc un problème crucial dans ces régions. Afin de comprendre son rôle, disposer d'une mesure précise et cohérente des inégalités de répartition des terres est un point de départ. Dans cet article, avec F.Novokmet, L.Bauluz et D.Sanchez, nous revisitons la littérature existante et fournissons de nouvelles estimations des inégalités de répartition des terres. Les estimations existantes s'appuient largement sur les données de recensement pour mesurer l'inégalité des terres. Cependant, le recensement, par définition, mesure les terres gérées par un ménage plutôt que les terres possédées. En outre, les recensements ne mesurent pas les différences de valeur ou de qualité des terres et ne tiennent pas compte de la population sans terre.

Nous nous écartons de la littérature existante en utilisant les enquêtes au sein des ménages (LSMS de la Banque mondiale ou autres enquêtes supervisées par les pays) plutôt que les données de recensement. Cela nous permet de surmonter certaines limites rencontrées par les estimations qui s'appuient sur les données de recensement pour estimer les inégalités de propriété des terres. Nous ré-estimons d'abord l'inégalité des terres telle que définie par la littérature existante, en utilisant les vagues de recensement les plus récents. Nous estimons ensuite les inégalités de la propriété des terres à partir des enquêtes et nous les comparons aux estimations du recensement. Nous constatons que les deux sont très similaires. Pour aller plus loin, nous estimons les inégalités de la valeur des terres et prenons enfin en compte les populations sans terre. Notre résultat montre l'importance de prendre en compte les différentes définitions. Notre définition choisie, la mesure des inégalités de la valeur des terres incluant les ménages sans-terre, met en évidence certains modèles régionaux intéressants : L'Asie du Sud et l'Amérique latine sont parmi les régions les plus inégalitaires, suivies par l'Afrique et l'Asie "communiste".

En bref, dans cet article, nous visons à fournir une mesure cohérente des inégalités de répartition des terres comme première étape dans l'analyse de son rôle dans le processus de développement de ces territoires. Nous mettons en avant la nécessité de

définir les différents concepts utilisés et l'utilité des enquêtes auprès des ménages pour estimer les inégalités de terre en termes de taille, de valeur et en prenant en compte une section marginalisée, mais très pertinente de la population, les ménages sans terre.

Bibliography

Alvaredo, Facundo and Anthony B Atkinson (2010). "Colonial Rule, Apartheid and Natural Resources: Top Incomes in South Africa, 1903-2007". In:

Alvaredo, Facundo, Denis Cogneau, and Thomas Piketty (2021). "Income inequality under colonial rule evidence from French Algeria, Cameroon, Tunisia, and Vietnam and comparisons with British colonies 1920–1960". In: *Journal of Development Economics*, p. 102680.

Atkinson, Anthony B (2014). *The colonial legacy: Income inequality in former British African colonies*. Tech. rep. WIDER Working Paper.

– (2015a). "The distribution of top incomes in former British West Africa 1". In:

– (2015b). "Top incomes in Central Africa: Historical evidence". In:

– (2015c). "Top incomes in East Africa before and after independence". In:

Bahar, Dany, Ana Marıa Ibáñez, and Sandra V Rozo (2021). "Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees". In: *Journal of Development Economics* 151, p. 102652.

Chiquiar, Daniel and Gordon H Hanson (2005). "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States". In: *Journal of political Economy* 113.2, pp. 239–281.

Chiswick, Barry R (1978). "The effect of Americanization on the earnings of foreign-born men". In: *Journal of political Economy* 86.5, pp. 897–921.

Deininger, Klaus and Lyn Squire (1998). "New ways of looking at old issues: inequality and growth". In: *Journal of development economics* 57.2, pp. 259–287.

Dustmann, Christian, Francesco Fasani, and Biagio Speciale (2017). "Illegal migration and consumption behavior of immigrant households". In: *Journal of the European Economic Association* 15.3, pp. 654–691.

Engerman, Stanley L and Kenneth L Sokoloff (2002). *Factor endowments, inequality, and paths of development among new world economics*.

– (2005). *Colonialism, inequality, and long-run paths of development*.

Frankema, Ewout (2010). "The colonial roots of land inequality: geography, factor endowments, or institutions?" In: *The Economic History Review* 63.2, pp. 418–451.

Hainmueller, Jens, Dominik Hangartner, and Giuseppe Pietrantuono (2017). "Catalyst or crown: Does naturalization promote the long-term social integration of immigrants?" In: *American Political Science Review* 111.2, pp. 256–276.

Monras, Joan, Javier Vázquez-Grenno, and Ferran Elias Moreno (2018). "Understanding the effects of legalizing undocumented immigrants". In:

Nunn, Nathan (2008). "Slavery, inequality, and economic development in the americas". In: *Institutions and economic performance* 15, pp. 148–180.

Pinotti, Paolo (2017). "Clicking on heaven's door: The effect of immigrant legalization on crime". In: *American Economic Review* 107.1, pp. 138–68.

Chapter 1

Is Naturalization a Passport for Better Labor Market Integration?

Abstract

Better integration is beneficial for both migrants and the host country. In this respect, granting citizenship could be an important policy to boost migrants' integration. In this paper, I estimate the causal impact of obtaining citizenship on migrants' labor market integration. I exploit a change in the law of naturalization through marriage in France in 2006. This reform amended the eligibility criteria for applicants by increasing the required number of years of marital life from 2 to 4, generating an exogenous shock and thus a quasi-experimental setting. Using administrative panel data, I first show evidence of the impact of the reform on naturalization rates. I then use a difference-in-differences model to estimate the labor market returns to naturalization. I find that, among those working, citizenship leads to an increase in annual earnings by 29%. This effect is driven by a significant increase in the number of hours worked, as well as a positive effect on hourly wages. While the gain in earnings is similar for both men and women, the effect for men is mostly driven by an increase in hours worked compared to an increase in hourly wages for women. I provide suggestive evidence that naturalization helps reduce informality, and discrimination. This chapter thus provides strong evidence that naturalization acts as a catalyst for labor market integration.

1.1 Introduction

Obtaining the nationality of the host country is deemed as a sign of integration for foreigners. There is however a growing political debate on whether naturalization is merely a reward for integration or can rather boost integration[1]. This debate has led to the support of opposing policies. On the one hand, supporters of the former position prescribe that the path to citizenship should be hardened to screen the best-integrated migrants. On the other hand, others support the relaxation of the rules since naturalization could help accelerate the integration process of migrants, in which case, it would be a potential tool for governments to provide better labor market prospects to foreigners. Despite substantial interest around this question, there is so far scarce evidence of the causal impact of naturalization on labor market integration.

This paper estimates the causal effect of naturalization on labor market outcomes and provides evidence of the mechanisms at play. Having well-integrated migrants, be it culturally, socially, or economically is a desirable condition for migrants themselves, as well as for the host country. Economic integration through labor market participation leads to less dependence on welfare benefits and even positive net fiscal contributions (Dustmann and Frattini, 2014, d'Albis, Boubtane, and Coulibaly, 2016). Lack of integration, on the other hand, could lead to hostility and anti-immigrant feelings from the native population. In fact, there are evidence that economically integrated immigrants tend to commit less crime (Freedman, Owens, and Bohn, 2018, Mastrobuoni and Pinotti, 2015). Their integration is hence crucial to ensure social cohesion in the host country.

The literature puts forward different factors that can boost migrants' economic integration: better language skills (Dustmann and Fabbri, 2003; Lochmann, Rapoport, and Speciale, 2019), networks, or marriage to a national (Safi and Rogers, 2008; Meng and Gregory, 2005; Meng and Meurs, 2009). However, it is well-documented that migrants tend to remain at a disadvantage on the labor market when compared to natives (Chiswick, 1978; Baker and Benjamin, 1994; Dell'Aringa, Lucifora, and Pagani, 2015). Hence, if naturalization could in itself boost foreigners' labor market

[1] As an example, an extract of an article from France 24 (28/08/2013) with a statement by the UMP, a center-right party in France on the Socialist Party's move to ease citizenship (emphasis added): "(The Socialist Party) wants to increase the number of naturalisations to *facilitate the integration* of immigrants...On the contrary, we think that becoming French must be the *result of a successful journey of assimilation* into the French community."

integration, then it could help in narrowing this gap.

There are different potential channels through which citizenship could boost labor market integration. First, in most countries, part of the labor market, often the public sector, tends to be conditional on nationality. This is the case in France, where an estimated 30% of the labor market falls within such a category (GED, 2000). Obtaining nationality thus provides foreigners with unrestricted access to the labor market. Additionally, citizenship can provide foreigners with a long-term stable legal stay, which can also be used as a signal of a commitment to stay in the host country to the employer. Finally, when applying for jobs, it can help reduce statistical or taste-based discrimination faced by foreigners. Altogether, citizenship can help to overcome barriers to employment, promotions, and economic opportunities.

In this paper, I estimate the causal effect of naturalization on labor market outcomes, by exploiting a national-level reform in the law of naturalization through marriage in France in 2006. As laid down in section 2.2, the reform amended the eligibility criteria of applicants by increasing the required number of years of marital life from 2 to 4. This provides a quasi-experimental setting whereby foreigners in two otherwise similar cohorts of marriage expecting to naturalize within the same number of years, unexpectedly face different length of waiting to be naturalized. To do so, I adopt a simple difference-in-differences. In order to compare the two cohorts of marriage while controlling for general economic trends, as a robustness check, I also adopt a triple difference approach using foreigners married to foreigners as a never-treated group, as this group of foreigners are not eligible for naturalization through the channel of marriage and hence, not affected by the reform.

I use the Permanent Demographic Sample[2], a rich administrative French panel data, described in section 2.3. This data source allows me to match the marriage registry, population censuses, and employee panel data. First, based on the reported nationality at the time of marriage, I identify the two main marriage groups: i) marriages of interest: foreigners married to French (in the main analysis), and ii) never-treated group: foreigners married to foreigners (for placebo checks and robustness analyses). By matching the marriage information to the population census data, I build a proxy for naturalization, defined as foreigners who declare being french in subsequent rounds of the census. Finally, I also follow the labor market trajectory of individuals who are salaried workers. The empirical strategy is to compare the foreigners among the group of interest who face a shorter compared to a longer waiting period, in a difference-in-differences analysis.

[2]Échantillon Démographique Permanent (EDP)

I establish three main results. First, I show that, as expected, the reform led to a gap in the propensity to be naturalized between foreigners married to French facing a shorter waiting time compared to those waiting longer, in the years following marriage. I then compare their labor market outcomes from their marriage up to 10 years after. This corresponds to the period before and after they become eligible for naturalization. I show that naturalization increases net annual earnings, through an increase in the number of hours worked and log hourly wages. I further show that these effects differ by gender. Both men and women experience an increase in earnings, but it is entirely driven by an increase in the number of hours worked for men.

Finally, in section 2.4, I explore different potential mechanisms. The mixed marriage context allows me to rule out some expected channels since the channels through work authorization and networks are less at stake for this group of foreigners as they already obtain those benefits prior to naturalization through their marriage. I show that the channel of unrestricted access to the labor market, as proxied by the probability of being employed in the public sector, does not seem to have played a role in explaining the naturalization premium. Instead, I provide suggestive evidence that naturalization is used as a signaling device for integration and language proficiency. I also show that naturalization helps reduce informality.

This paper looks at a context of relatively moderate access to citizenship. In terms of immigration policies and naturalization laws, France can be placed closer to the traditional immigrant countries such as Canada, Australia, and the US, where the average number of years of residence required is of 5 years or less, in contrast to more than 10 years on average in Germany, Switzerland, Italy, and Spain.

My paper has three main contributions. First, it estimates the causal effect of naturalization, overcoming challenges of self-selection, endogeneity, and reverse causality that have so far limited this literature. Chiswick (1978), at the onset of this literature, found a positive correlation between naturalization and labor market outcomes, by comparing the situation of naturalized to non-naturalized individuals in the U.S using cross-sectional data. A more recent strand of the literature has exploited panel data to take into account time-invariant individual characteristics, and also finds a positive association (Bratsberg, Ragan, and Nasir, 2002; Fougère and Safi, 2009; Steinhardt, 2012). In France, comparing naturalized and non-naturalized migrants using panel data, Fougère and Safi (2009) found that obtaining French nationality is associated with a significantly higher probability of being employed.

While being the first to show a link between the two, these studies suffer from issues

of self-selection, endogeneity, and reverse causality. Naturalization involves a double positive selection: firstly, individuals who choose to apply for the nationality are normally positively selected among the pool of all immigrants and secondly, those who end up obtaining the nationality are also positively selected among the group of applicants. Hence, comparing naturalized immigrants to non-naturalized ones leads to biased estimates. Establishing the causal impact of naturalization is furthermore complicated given that while citizenship might lead to better labor market outcomes, the reverse is also likely to be true as well-integrated immigrants have higher chances of being naturalized. This might induce those who wish to apply for the nationality to invest most in their human capital.

Second, this paper exploits a novel and clean setting to estimate the direct causal effect of naturalization. In doing so, it comes closer to the few noticeable exceptions in this literature that provides evidence using quasi-experimental designs. This is the case of Hainmueller, Hangartner, and Ward (2019), which compares those who are naturalized or not by a close margin in local referendums in Swiss municipalities, showing that naturalization has a long-run positive effect on earnings in a rather conservative context[3]. In effect, they compare the positive effect of obtaining citizenship to the negative effect of having a rejection[4]. I depart from this paper by looking at a change in the access to naturalization that allows me to compare a group of foreigners that obtain the nationality to another similar group that eventually obtains it. My analysis is thus closely related to the paper by Gathmann and Keller (2018) which focuses on the labor market effects of differential access to citizenship. Exploiting two reforms that altered residency requirements in Germany, they find that naturalizing earlier has a significant positive and persistent effect on the long-run economic integration of migrants. In my paper, I look at individuals who naturalize when they are in the working-age group. I follow their labor market outcomes before and after naturalization, and I hence provide the direct and immediate effect of naturalization.

Third, this paper contributes to the literature on employment discrimination (see Bertrand and Duflo, 2017 for a comprehensive review of the literature on field experiments on discrimination). In France, Duguet et al. (2010) and Petit, Duguet, and L'Horty (2015) show through testing that the CV of a Moroccan with a French

[3]Applicants need to have spent 12 years in Switzerland and passed a culture and language test.

[4]Critics of the direct referendum to grant citizenship in Switzerland put forward the potentially discriminatory practice. It has been declared unconstitutional and deemed to violate international laws in 2003 by the Swiss Federal Court. The resulting unwelcome perception felt by those who are refused the nationality by members of their own locality, might have induced an adverse behavioral change on the labor market of rejected candidates, negatively affecting the control group.

nationality receives more callbacks from employers than that with a Moroccan nationality. This might be explained by statistical discrimination against foreigners based on perceived language skills. In this paper, I bring suggestive evidence that nationality can help reduce hiring discrimination by signaling better language proficiency and more integration. I conclude in section 2.5 that naturalization is a catalyst for the economic integration for foreigners and can thus be a crucial policy tool.

1.2 Context & Design

Like most developed countries, France has had a long history of political debate about the softening or hardening of its migration policies (Weil, 2002). Foreigners can obtain a legal stay in France through different types of visas, depending on their status and purpose of stay. These may be short-term ones (e.g. student visa, short-term work permits..) or longer-terms (e.g. 10 years residence permit). Upon satisfactory integration in France, foreigners become eligible for naturalization.

Applicants to naturalization are generally assessed based on their degree of integration in the country, by the French authorities. The two main channels through which a foreigner can apply for naturalization are through decree and through declaration[5]. The first channel, being the general process, requires significant proof of socio-economic integration. The second channel applies to individuals born in France to foreign parents, as well as to foreigners married to French nationals, which is the focus of this paper. Since both situations, in themselves, constitute some level of integration, naturalization through declaration is deemed part of the natural order. While foreigners applying through decree have to show proof of substantial integration in the social and professional life in France, foreign spouses of French citizens are only required to fulfill three criteria: a certain number of years of marital life, a valid marriage, and a sufficient knowledge of French, their marriage to a French national being an adequate proof of integration.

The success rate among admissible files is estimated to be at around 70% for applications by decree and 90% for those through declaration. This gives an insight into the relative preference for the latter channel whenever possible. Rejections of applications of naturalization through marriage are rare and only occur in cases of ineligibility with respect to the main criteria or for invalid marriages determined through an in-depth inquiry by the local authorities. Despite the screening measures

[5]The bulk of applications (around 60% are through decree and 40% through declaration, of which half is through marriage.

in place, this somewhat privileged access to naturalization has led many to fear that mixed marriages could be wrongly instrumentalized to obtain the French nationality. As a result, throughout time, the French government has attempted to harden the rules to applying through the channel of marriage, mainly by increasing the number of years of marriage to a French national required to be eligible. Apart from the 1998 reform, when this condition had been relaxed, all the other reforms increased this duration, the underlying justification being that longer marriage duration requirements are more costly and will deter individuals from contracting marriages for the sole reason of obtaining the French nationality.

Similarly, the reform announced in March 2006 and acted in July 2006 changed the eligibility criteria of naturalization through marriage by increasing the number of years of marriage required to be eligible from 2 years to 4 years. Given the retroactive nature of the law, the relevant eligibility criteria for any given foreigner married to a French depended on their application date. It differently affected applicants before and after the reform in 2006 which translated into the unintended consequence of penalizing the cohort of marriage after 2004 compared to the couples married before 2004. In other words, in terms of application dates, applicants before July 2006 were required to have at least two years of marriage to be eligible, and conversely, any application after July 2006 had to fulfill the new requirement of at least four years of marriage to be eligible. This translated into the fact that only marriages that were contracted before July 2004 could have the possibility to apply for naturalization after 2 years of marriage, while those married after July 2004 faced the hardened eligibility criteria and had to wait 4 years[6].

The identification strategy relies on the fact that the reform was unanticipated at the time of marriage: any couple married before the announcement of the reform in March 2006 expected to wait only two years after marriage to be eligible to apply for naturalization. Hence, there is no reason for mixed married couples before and after 2004 to be any different except for their differential probability of obtaining the nationality, due to this exogenous shock. The treatment is defined as the higher propensity to being naturalized and marriages within a window before July 2004 are thus defined as the "early treated" group (by naturalization) and those after July 2004 as the "late treated" group (with respect to naturalization) as in Figure 1.2.

Conceptually, under a full compliance setting and no administrative delay, we would

[6]Only mixed marriages between January 2002–February 2006 are kept in the sample. As an example, a foreigner married to a French national in January 2004 would be eligible as soon as January 2006 while a similar foreigner married in December 2004 would only have 2 years of marriage in December 2006, not enough to be eligible under the new law.

expect a 0% naturalization rate among the early treated and late treated group up to two years after marriage. If every individual applied as soon as they were eligible, that there were no administrative delays, and no rejection in obtaining the French nationality through marriage, then there would be a 100% naturalization rate among the early treated group as from the third year after marriage. Under similar conditions, the late treated group would have a 0% naturalization rate up to 4 years after marriage and a 100% rate as from the fifth year after marriage (See Appendix Figure A.1.1).

However, in practice, this is not likely to be the case. First, the announcement of the reform in March 2006 might lead to changes in behavior and hence to changes in the composition of marriages after the announcement. This is taken care of by restricting the end date of the sample to marriages up to February 2006. Second, among the early-treated group, while everyone is eligible to apply within two years of marriage, not everyone might have enough time to apply before the reform passes. This is mostly a concern for marriages closer to the July 2004 cutoff. To address this, the end date of the early treated cohort is limited to February 2004. Hence, the early-treated cohort are marriages which occurred between January 2002 and February 2004, and the late-treated cohort consists of marriages between July 2004 and February 2006.

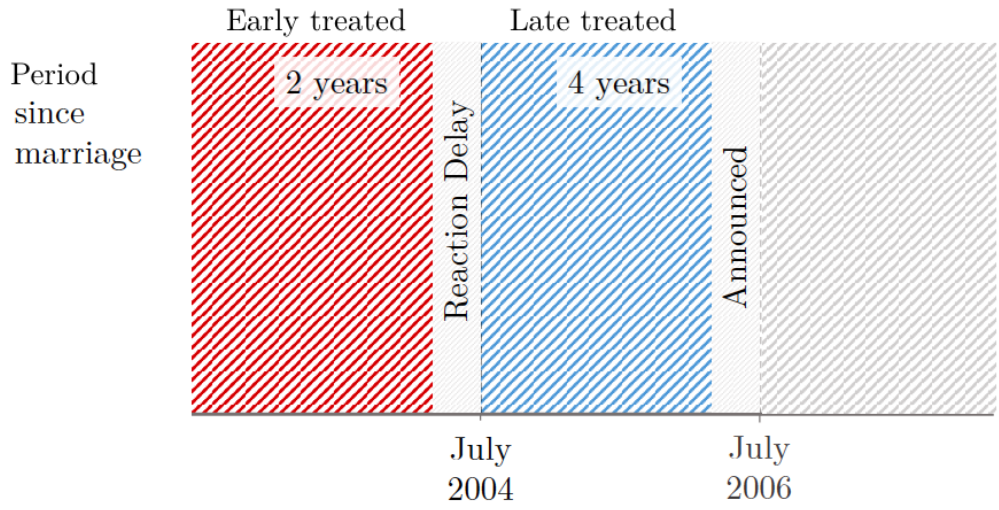


Figure 1.1: Empirical Strategy Design

In addition, there might be non-compliance, making this a fuzzy design. Some marriages in the early-treated cohort might not be treated within two years of marriage due to two main reasons: if they do not apply before July 2006 (despite limiting this risk as explained above); and if they would not apply for the nationality

irrespective of the eligibility criteria, known as the never-takers. Additionally, there might be foreigners treated prior to four years of marriage in the late-treated group since foreigners married to French nationals can also choose to apply for the French nationality through the general channel if they are eligible[7], despite not having the incentives to do so. Since there is no direct information on naturalization in the data, only a proxy of naturalization is used, as explained in Section 2.4. This setting is thus similar to an intention to treat (ITT) design. There are also administrative delays between the date of application and obtaining the nationality that is estimated to be almost a year on average[8]. Due to all of these reasons, the differential naturalization rate between the early and late-treated group.

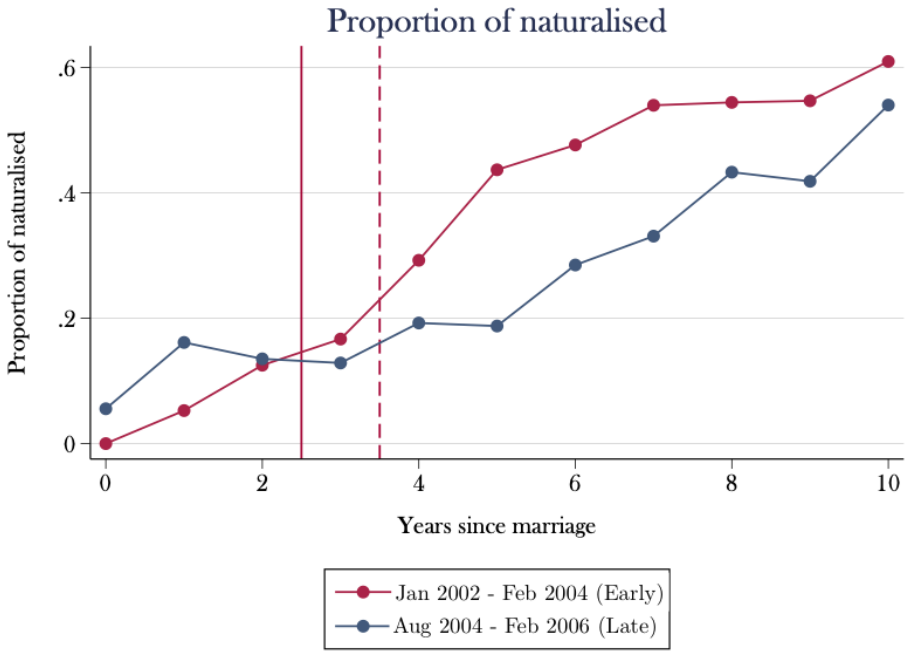


Figure 1.2: Proportion of naturalized in treated and control groups

Figure 1.2 shows that there is a sizeable gap in the proportion of naturalized between the early-treated and the late-treated group as from the fourth year of marriage. As expected, assuming a one-year administrative delay, there is a change in the trend of the share of naturalized in the early-treated group beyond the third year of marriage and the same change in trend occurs beyond the fifth year for the late-treated group. This is empirically tested in the first-stage analysis in section 2.4. For placebo and robustness checks, the same analysis is performed on a similar yet never-treated

[7] If for instance those who have been married for less than 4 years of marriage, have resided for at least 5 years on the French territory and can prove sufficient integration in the socio-economic life in France, then they could apply through the decree channel

[8] Acquisitions et pertes de la nationalité française- Rapport annuel de la sous-direction des naturalisations, 2005

group. These are foreigners married to foreigners as they are not eligible to apply for naturalization through the marriage channel, and are hence not affected by the reform. The reduced-form analysis exploits the gap in naturalization rates to estimate its effect on labor market outcomes[9]. The reduced-form coefficients[10] corresponds to the effect of the ITT and the local average treatment effect (LATE) can be recovered under some assumptions, by dividing the ITT by the differential propensity of being naturalized, obtained in the first-stage.

Given that the reform only impacted the naturalization channel through marriage, in this paper, I focus on foreigners in mixed marriages. Marriages between French and foreign nationals account for 13% of all marriages in France on average. These foreigners are generally relatively more integrated among the pool of foreigners. First, upon marriage to a French national, foreigners are eligible for a special[11] visa, renewable every two to four years, which allows them to have a legal stay and authorization to work in France. They might also be more familiar to the French institutions, labor laws, taxation, and social security systems.

When compared to the foreign-born population in the same age window, they tend to be more educated on average, have a higher probability of being employed as manual workers and relatively less in executive positions, and finally they are more likely to be from a francophone country on average (see Table A.1). These specificities imply that some of the mechanisms put forward by the literature on the effect of naturalization on labor market outcomes are not relevant in this case, allowing me to disentangle and pin down other mechanisms. For instance, through the spouse visa, foreigners already acquire a stable stay and employers do not face additional costs in hiring them. Hence, the main channels that are still relevant are: unrestricted access to jobs which are conditional on the nationality; reducing discrimination; and reduced informality, which will be tested in Section 2.5.

1.3 Data & Empirical Setting

I exploit the French administrative panel data known as the *Echantillon Demographique Permanent* (EDP). It is a panel that matches different administrative data sources for individuals born on certain dates of the year, providing the

[9] Extensive robustness checks show that there are no differential rates of migration out of the country due to this reform and the sample composition based on observable characteristics remains similar between early-treated and late-treated group over time.

[10] Given the structure of the data and the sample under study, merging the three sources of data to perform a second-stage analysis is not feasible.

[11] "Vie Privée et Familiale (VPF)"- Private and Family Life

sociodemographic characteristics of individuals. Before 2004, the EDP constituted a sample of approximately 1% of the total population and around 4% thereafter[12]. In this paper, I focus on three main data sources of the EDP which are the civil registries of marriage, the population census, and the employees' panel data (part of the Déclaration Annuelle de Données Sociales - DADS).

First, the civil registry of marriage allows me to identify the date of marriage of couples with an EDP individual, as well as other characteristics for both spouses at the time of marriage. This includes their nationality, gender, and age among others. Through this data source, I can categorize individuals into different types of marriage, namely endogamous marriages between two French or two foreigners, as compared to mixed marriages[13]. In line with the identification strategy, mixed marriages are defined as any marriage contracted between a foreign individual and a French person[14], as recorded at the date of marriage. Since Europeans are less likely to be affected by this reform, they are excluded from the analysis[15]. Only marriages contracted between January 2002 and February 2006[16] are kept in the analysis. The sample is limited to February 2006 to ensure that marriages affected by the announcement of the reform are excluded. Marriages between February 2004 and July 2004 are also taken out to minimize the number of foreigners that were eligible for the short waiting time, but did not have enough time to apply before the change of the law in July 2006.

Second, I match the marriage registries to the different rounds of population censuses. As of 2004, the population census is based on a five-year rotating sample of around 14% of the population yearly. This annual structure of the population census gives information on the nationality of the individual every time they are surveyed in the census. While there is no direct information on naturalization (date of application, date of naturalization, naturalization channel), individuals report their nationality each time they are surveyed. This provides a proxy for naturalization. An individual is considered naturalized if he/she is recorded as non-French at the time of marriage and reports being French in subsequent years in the census[17]. Population censuses

[12]Before 2004, the EDP included individuals born on 4 dates of the year. The sample has increased to 16 dates of the year as of 2004. This was effectively applied to the civil registries in 2004 but to the population census only as of 2008. Independently, the employer-employee data had a sample of 4% of the population until 2001, and it has increased to 8% in 2002.

[13]Interchangeably used with the term "intermarriage".

[14]Irrespective of whether the French spouse is herself/himself a naturalized citizen or is a second-generation immigrant. Further distinction and heterogeneous analyses are carried out in section 2.5.

[15]Reference to non-French in this paper is interchangeable with non-Europeans.

[16]Excluding marriages between February 2004 and July 2004.

[17]Despite some measurement errors, this remains the best tool to measure naturalization. There

also contain extensive sociodemographic information such as country of birth, level of education, and marital status, providing an indication of divorces. Since the interest of this analysis is to look at the labor market outcomes, only the working population is kept in the sample (aged between 20 and 65 years old).

Finally, to look at the effect of naturalization on labor market outcomes of individuals, I match the marriage registry data to the employees panel data. This data is originally derived from a panelized version of the employer-employee linked data (DADS)[18]. It provides extensive annual information on employed individuals, namely their salary, type of contract, type of occupation, number of hours worked among others. Only foreign individuals who have worked at least once before 2002, hence entered the employee panel before their date of marriage, are kept in this panel to ensure that the results are not driven by new entrants. To be able to meaningfully interpret the result, I also restrict the sample to include only foreigners who have worked in the baseline period.

The empirical strategy takes the form of a difference-in-differences in the static form and an event-study analysis in the dynamic form, centered around the date of marriage. In other words, each time period is expressed in terms of its distance from the date of marriage or simply the duration since marriage (Dur). A reasonable event-window of up to 10 years after marriage[19] is included in the analysis. In the static double difference analysis, a pre- and post-treatment period is defined. Given the one-year administrative delay on average, the pre-period are defined as the first three years since marriage and the post-period is set at more than 3 years since the year of marriage.

In the first-stage, I show evidence of the effect of the reform on the naturalization rates among the early-treated and late-treated groups. To do so, I match the marriage registry to the population census. I build an indicator of naturalization (Nat_{it}) for whether the foreign individual i at the time of their marriage, reports being French or foreigner at time t in the census. I estimate equation (1) where i is the individual, t is the calendar year, $Treat_i$ is a dummy of whether individual i is in early-treated or late-treated cohort, $Post_{it}$ is a dummy for more than two or three years of marriage, depending on the specification. The coefficient of interest, λ gives the differential rate of naturalization between early-treated and late-treated group. The specification

is otherwise no official dataset that tracks naturalized foreign individuals, hence no information on the exact date and type of naturalization of foreign individuals.

[18] Déclaration Annuelle des Données Sociales

[19] It corresponds to 11 time periods, whereby d ranges from 0 (the year of marriage) to 10 (ten years after marriage)

for dynamic form is similar, whereby $Post_{it}$ is replaced by a duration dummy for each time period since marriage and these are interacted with $Treat_i$.

$$Nat_{it} = \alpha + \delta Treat_i + \beta Post_{it} + \lambda(Post_{it} * Treat_i) + \epsilon_{it} \quad (1)$$

In the second step, I estimate the reduced form effect of naturalization on labor market outcomes[20]. To do so, I match the marriage registry data to the employee panel data. The static specification for the difference-in-differences strategy is as follows:

$$Y_{it} = \eta + \delta Post_{it} + \gamma(Post_{it} * Treat_i) + \mu_i + \epsilon_{it} \quad (2)$$

where γ is the coefficient of interest. Y are labor market outcomes such as annual earnings, no of hours worked, and hourly wage. As in the first-stage analysis, $Post$ is a dummy for being up to or more than three years since marriage and $Treat$ is a dummy for being in the early-treated or late-treated group. Individual fixed effects are also included in this analysis (μ_i). The standard errors are clustered at the individual level. In the equivalent dynamic model, duration fixed-effect are included to account for any potential effects that are specific to a particular number of years of marriage[21]. Coefficients of interest in the dynamic form are the interaction between each duration dummy and $Treat$. In the static form, the interaction term between duration and treatment group, γ in equation (2) estimates the differences between the early-treated and late-treated group, hence the reduced-form effect of naturalization. The underlying common trend assumption holds if the early-treated and late-treated group evolve in a similar way in the pre-treatment period, especially in their labor market situations.

A potential threat to this identification strategy is the fact that foreigners married to French who fulfill the requirements are also eligible to apply through the general channel. A toughening of the criteria to apply through the marriage channel can lead some of those in the late-treated group to apply for and obtain the nationality in this way to overcome the slightly longer waiting time. This would lead to a positive share of naturalized individuals in the late-treated group. As long as the early-treated group has a sizable higher share of naturalized individuals, this is not a concern. However, if the late-treated individuals exercise more effort on the labor market to maximize their chances of obtaining the nationality, the reduced form estimates may

[20] Given the structure of the data and the sample under study, merging the three sources of data to perform a second-stage analysis is not feasible.

[21] For instance, couples might have kids in the first few years following marriage.

suffer from an attenuation bias due to the better labor market outcomes among the late-treated group induced by their behavioral response to the longer waiting time. In practice, since foreigners married to french have a legal stay and an authorization to work through their marriage, these behavioral responses are likely to be marginal.

Descriptive Statistics

Table 1.1 shows the descriptive statistics on demographic and labor market characteristics for the period under study. The main group of interest are foreigners married to french, and the never-treated group as a point of comparison are foreigners married to foreigners. The average age and age difference between spouses at marriage is lower on average among mixed married couples compared to the average foreign couples. There are on average more foreign men married to french women than marriages between foreign women and french men, as seen by the proportion of women in the sample of mixed marriages, which is at 34%. Around 60% of the foreigners are from francophone countries and the majority comes from Algeria, Morocco, and Tunisia, which accounts for 54% of the sample of mixed marriages.

	Foreigner - French		Foreigner - Foreigner	
	Mean	SD	Mean	SD
Demographic characteristics				
Age	36.14	7.41	38.82	8.57
Age diff	5.35	5.17	6.14	5.41
Female	0.34	0.47	0.45	0.50
Francophone	0.63	0.48	0.57	0.49
Nationality of origin				
Algerian	0.18		0.22	
Morrocan	0.27		0.15	
Tunisian	0.09		0.03	
Others	0.46		0.60	
Labor Market charateristics				
Prob. Panel	0.72	0.45	0.68	0.47
Net annual earnings	17216.6	13111.2	16763.3	12975.7
Number of hrs worked	1334.0	675.1	1349.0	676.5
Hourly wages	12.6	6.6	12.0	6.0
Full-time	0.72	0.45	0.70	0.46
Public Empl.	0.07	0.26	0.06	0.24
Obs	4919		3403	

Table 1.1: Descriptive Statistics

In terms of labor market characteristics, the probability of observing the foreign individual in the married couple as being employed is around 70% for both groups. Mixed couples tend to earn slightly higher annual earnings on average due to higher hourly wages, despite a lower number of hours worked on average. Around 70% of employed are employed with a full-time contract.

The balancing test of the main covariates at baseline for the two groups as well as the difference of the differences are reported in Table A.2. The average age at marriage has generally been increasing, and spouses have an average age of 33 years at the time of marriage. Given this trend, the average age at marriage among the early-treated group (married before 2004) is automatically lower than that of the late-treated group (married after 2004), especially in the group of interest (Column 1-3 of Table A.2). None of the labor market characteristics are significantly different between early-treated and late-treated groups among the group of interest.

1.4 Results

4.1. First-Stage

This section tests whether the reform has had an effect on the naturalization rates in the early-treated and late-treated group, by estimating equation (1). Table 1.2 summarizes the results of the first stage analysis for the main group of interest (foreigners married to french) and the never-treated or placebo group (foreigners married to foreigners) with a difference-in-differences approach. In the conservative approach and under the scenario of no administrative delay, the post-period is defined as after the second year of marriage, T2 (columns 1 and 3). As reported by the official statistics on naturalization, the average delay between the time of application and an administrative answer is a year on average. Hence, in a more likely scenario, the post-period can also be set after T3 (columns 2 and 4). These results show that the probability of being naturalized is between 13 to 15% higher in the early-treated group compared to the late-treated group for mixed marriages, the difference being highly significant. The non-significant result for the never-treated group confirms that the gap only exist for the group of interest, and it is most likely driven by the reform.

Figure 1.3 shows the underlying dynamic effects whereby each point estimate is the differential rate of naturalization in the early-treated foreigners compared to the late-treated foreigners married to a French, at each year since marriage[22]. T0

[22]The estimate of the rate of naturalization is conditional upon being observed in the population census. For instance, the coefficient of T4 is interpreted as a 20 percentage point higher naturalization

	(1)	(2)	(3)	(4)
	Foreigner-French		Foreigner-Foreigner	
Cutoff	After T2	After T3	After T2	After T3
Treat x Post	0.13***	0.15***	-0.04	-0.04
	(0.02)	(0.03)	(0.04)	(0.04)
Observations	1,804	1,804	687	687
Adj R-squared	0.04	0.06	0.02	0.02
Individual FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.2: First Stage regressions

corresponds to the year of marriage and T10 refers to 10 years after marriage. Since the early-treated group are married before the 2004 threshold, they become eligible to apply for naturalization through marriage as soon as 2 years after marriage. On the contrary, having contracted a marriage after July 2004, the foreigners in the late-treated group will only become eligible through this channel after 4 years of marriage. In addition, it takes a year on average for the French administration to process the application.

The rates of naturalization between the two groups do not seem to significantly differ in the "pre-treatment" period- from the year of marriage to two years after marriage, since none of the groups are eligible for naturalization through the channel of marriage. The difference gradually sets in as from the fourth year of marriage, likely due to the one-year administrative delay, at about 20-25 percentage points. The gap seems to close off as from 6 years of marriage, consistent with the timing at which the late-treated group is likely to witness an increasing probability of being naturalized[23].

Different placebo analyses are undertaken to confirm the validity of the first stage. First, since foreigners married to foreigners are not eligible to apply to the nationality through the marriage channel, they are not impacted by the reform. Column 3 and 4 of table 1.2, as well as Figure 1.4 shows the result of a similar analysis with non-mixed foreign marriages. As expected, there are no significant difference in

rate among the early-treated group compared to the late-treated group conditional of being in the population census 4 years after marriage. A series of robustness checks are carried out to show that there is no differential rate of attrition and stable population composition.

[23]Robustness checks show that there are no differential probability of observing individuals in the early-treated and late-treated group over time and the sample composition based on observable characteristics remains similar between early-treated and late-treated group over time.

the naturalization rates between the equivalent "early-treated" and the equivalent "late-treated" groups in this never-treated group. The coefficients of the dynamic analysis are not different from zero when taken together. This supports the claim that the patterns seen in Figure 1.3 are driven by the reform for naturalization through marriage, and it validates the use of the foreigners married to foreigners as a never-treated group in the triple-difference analysis in the robustness test section. In addition, a second set of placebo tests are presented in Appendix A.4, whereby the reform dates are altered and the dynamic first-stage exercise for the group of interest, foreigners married to French citizens, are presented.

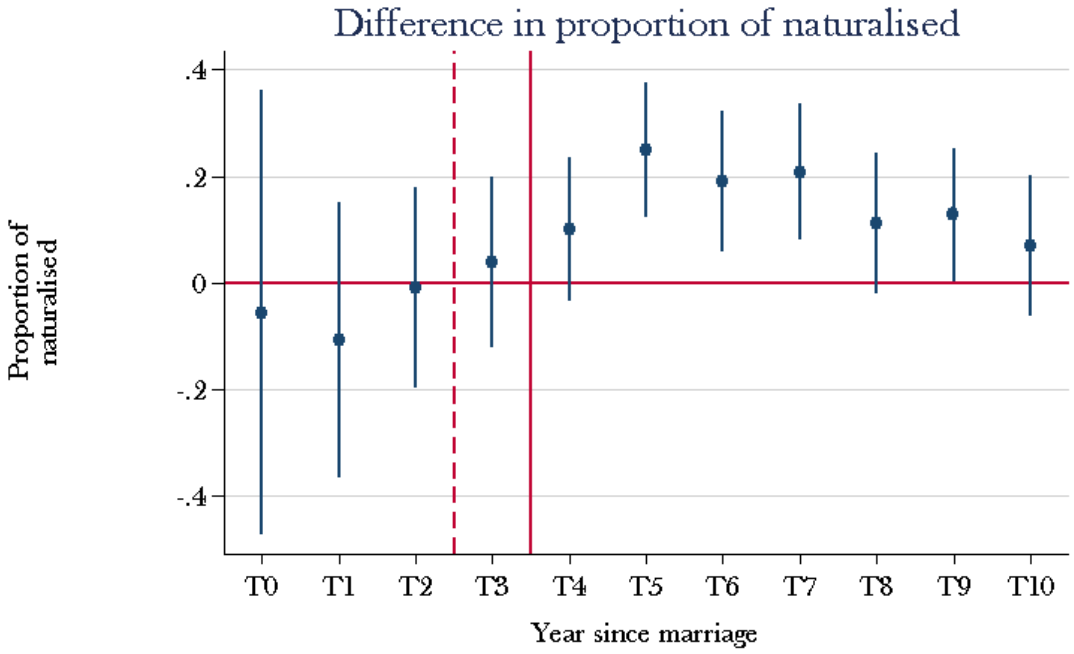


Figure 1.3: Difference in the proportion of naturalized between treated and control groups

4.2. Reduced Form

Exploiting the 2006 reform shock on the naturalization propensity of two otherwise comparable groups, I estimate the causal effect of naturalization on the labor market outcomes of foreigners. In this section, the reduced-form equation (2) is estimated and results based on the difference-in-differences approach are reported in static and dynamic forms[24].

The main result of the difference-in-differences analysis is presented in Table 1.3[25]. Naturalization led to approximately 2300 € or a 29% increase in annual earnings.

[24] All confidence intervals are at the 95% as standard in the literature.
[25] The results are conditional of working in the first 3 years since marriage

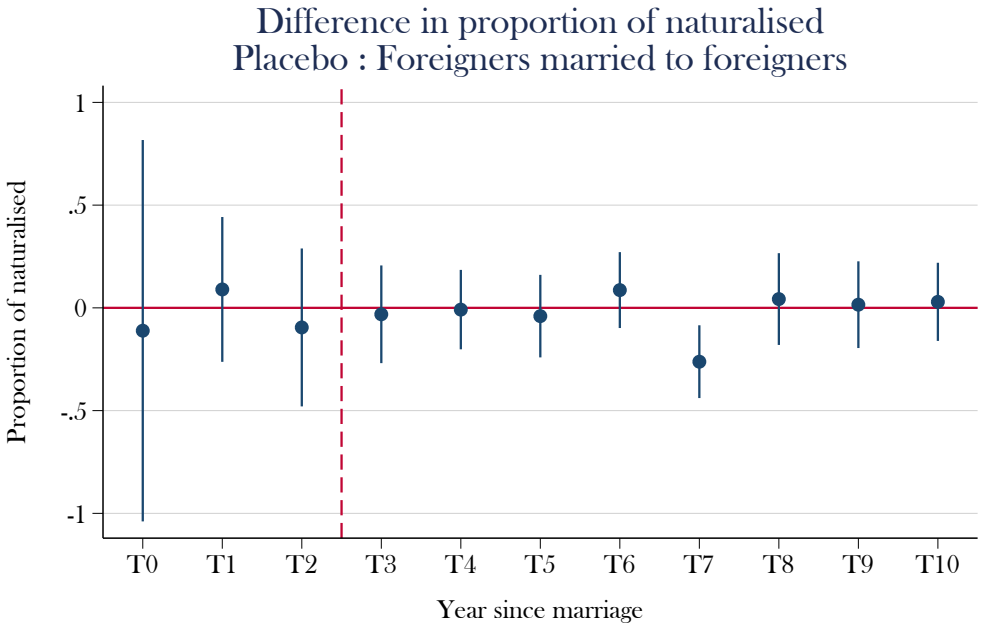


Figure 1.4: Difference in the proportion of naturalized between treated and control groups among foreign non-mixed marriages

This can be decomposed into a positive effect on the number of hours worked and hourly wages. The model explains up to 65% of the variations in annual earnings. These results are similar to the triple differences analysis in magnitude, as reported in the robustness test section. This suggests that accounting or not for the year effects does not significantly change the results.

	(1) Net annual earnings	(2) Log earnings	(3) No of hours worked	(4) Hourly wages	(5) Log Wages
Post x Treat	2,293.28**	0.29***	111.92*	0.92**	0.07***
	(1,094.35)	(0.11)	(66.07)	(0.42)	(0.03)
Observations	2,040	2,040	2,040	2,040	2,040
Adj R-squared	0.65	0.41	0.43	0.61	0.65
Ind. FE	Yes	Yes	Yes	Yes	Yes
Mean	17103	10.11	1407	11.83	3.108

The table present the difference-in-differences coefficient for foreigners married to french citizens before and after 2004. The pre-period consist of the first three years of marriage (T0 - T3) and the post-period is defined as time periods beyond the third year of marriage (T4 - T10). Results are conditional on working in the pre-treatment period. Standard errors clustered at the individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.3: Difference-in-differences: Effects of naturalization on labor market outcomes of foreigners

Figure 1.5 shows the dynamic effect of naturalization on earnings. There is no

significant difference between the early-treated and late-treated group up to 3 years since marriage, hence no pre-trend. The effect of naturalization kicks in as from the fourth year since marriage, as expected. The effect of naturalization on annual earnings can be decomposed into its effect on the number of hours worked and hourly wages. Figure 1.6 shows an increase in the number of hours worked as well as in the hourly wages.

The gradual decrease in the labor market effect over time is the mechanical result of the catch-up of the late-treated group with the early-treated group, as late-treated foreigners also obtain the French nationality as from the fifth year since marriage. This implies that there are no discernible persistent effect on annual earnings of being naturalized earlier in this context. The gap between the two group narrows down completely in terms of the number of hours worked. This does not seem to be the case in terms of the average hourly wages for which the gap remains sizeable, up to ten years after marriage, even if not significantly different as seen in Figure 1.6.

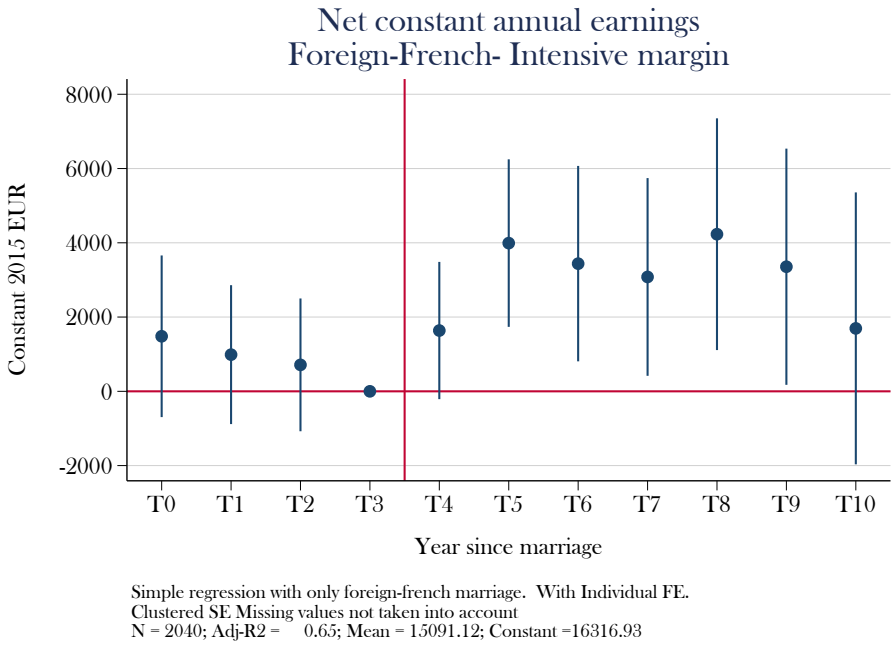


Figure 1.5: Dynamic effect of naturalization on annual earnings

Gender decomposition

These results mask underlying gender differences. Figure 1.7 reveals that the effects on annual earnings are much larger for women in absolute terms compared to men. When decomposed in terms of its effect on the number of hours worked (Figure

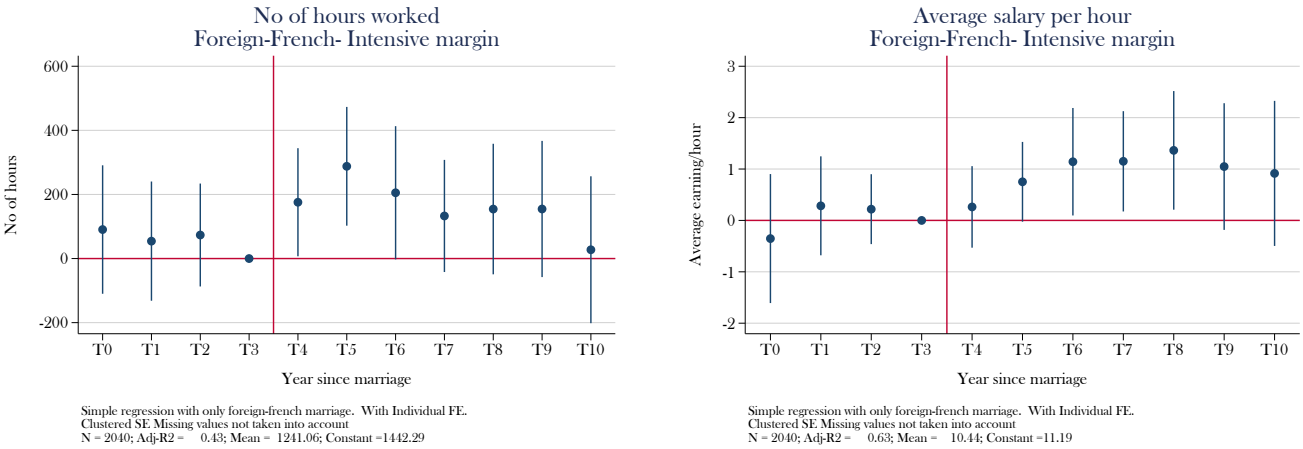
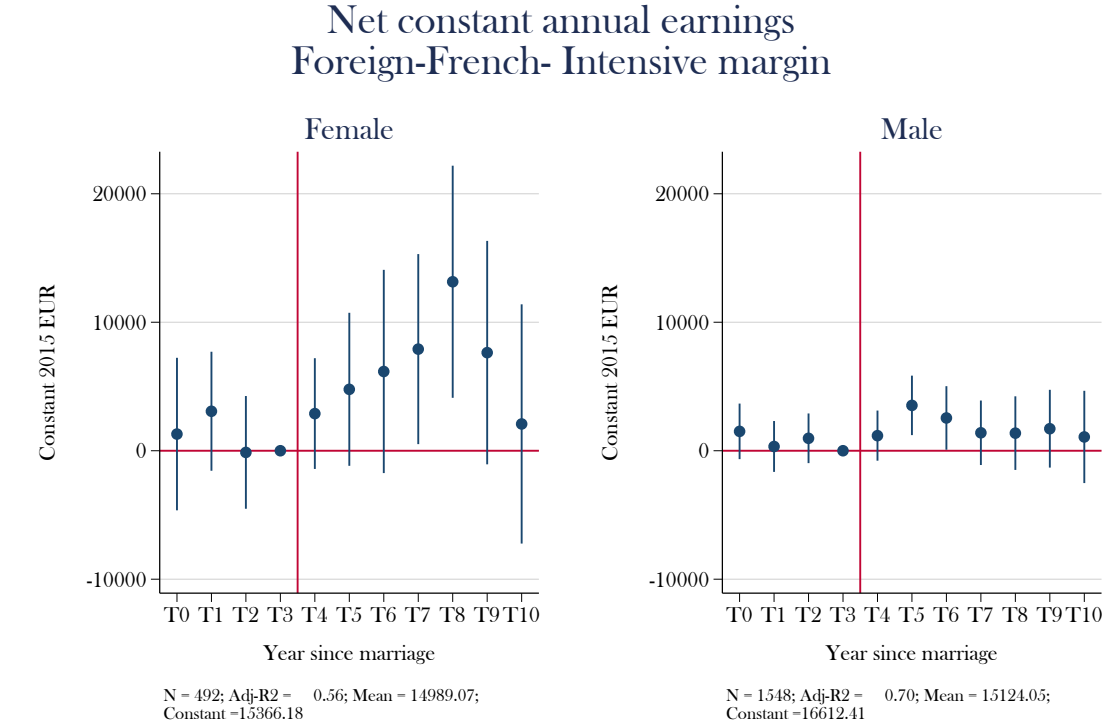


Figure 1.6: Dynamic effect of naturalization on (a) Number of hours worked and (b) Hourly wages

A.5.1), and hourly wages (Figure A.5.2), it seems that men gain more in terms of an increase in the number of hours worked while the effect for women seem to go relatively more through an increase in hourly wages.



Simple regression with only foreign-french marriage. With Individual FE.
Clustered SE Missing values not taken into account

Figure 1.7: Heterogeneous analysis by gender: Effect on annual earnings

4.3. Robustness Tests

Since the main analysis is a cohort comparison design, the calendar year effects

cannot be directly accounted for by including year fixed effects. To reduce any bias related to this, a similar group that is not affected by the reform is included to capture any year-specific effects through a triple-difference approach. Foreigners married to foreigners are not eligible to apply for naturalization through the marriage channel. Marriages between two foreigners are thus considered as never-treated groups since they are unaffected by the reform. To make sure that foreigners married to foreigners are similar in characteristics to those married to french, I implement a Coarsened Exact Matching (CEM) (Iacus, King, and Porro, 2012) on baseline characteristics such as the age group, year, gender, sector of employment, working full-time or not and earnings[26].

Given the setting, there should be no differential rate of naturalized between a similarly-defined "early-treated" and "late-treated" group among the never-treated foreigner group. In terms of the first-stage analysis, this is the case as shown in Column 3-4 of Table 1.2. The reduced form estimates of the effect of naturalization on labor market outcomes with a triple difference approach is obtained through the following specification:

$$Y_{it} = \eta_2 + \delta_2 Post_{it} + \theta(Post_{it} * Mixed_i) + \gamma_2(Post_{it} * Treat_i) + \rho(Post_{it} * Treat_i * Mixed_i) + \mu_i + \epsilon_{it} \quad (3)$$

where all variables are as described for equation (2) and; $Mixed_i$ is a dummy for whether the foreign individual is married to a french (group of interest) or to a foreigner (never-treated group). As in the specification (2), $Post_{it}$, as well as the interaction between $Post_{it}$ and treatment are included. In addition, in this specification, the interaction between the three are included. The term of interest, ρ, provides the effect of naturalization on labor market outcomes at each duration since marriage for the early-treated group compared to the late-treated group of the group of interest compared to the never-treated group. In other words, in the triple difference approach, the estimate is net of any effect that might arise due to the calendar year. This relies on the plausible assumption that both $Mixed_i$ groups are affected in similar ways by calendar effects. As before, this model is also estimated in its dynamic form by including duration fixed effects and the corresponding interactions.

The balancing test for both the groups of interest (foreigners married to French) and the never-treated group (foreigners married to foreigners), as well as the difference of the differences at baseline is reported in A.2. Column 4-6 shows the basic characteristics in the "early-treated" and "late-treated" group and the difference

[26]Baseline here refers to pre-treatment period $Dur = 0$ to 3.

between the two in the never-treated group. It is noted that the difference of the differences in column 7 is significant for the age and the age difference. However, none of the labor market outcomes have significant differences in the baseline period.. The results for the triple difference estimates are shown in table 1.4. While this model introduces noise and hence results in a loss in the precision of the estimates, the signs and the magnitude are comparable to the ones in table 1.3. This suggests that accounting for year effects does not significantly alter the results.

	(1) Net annual earnings	(2) Log earnings	(3) No of hours worked	(4) Hourly wages	(5) Log Wages
Post x Treat x Mixed	2,758.52 (1,907.57)	0.36** (0.17)	115.55 (114.46)	1.31* (0.72)	0.11*** (0.04)
Observations	3,238	3,238	3,238	3,238	3,238
Adj R-squared	0.67	0.40	0.41	0.66	0.69
Ind. FE	Yes	Yes	Yes	Yes	Yes
Mean	17321	10.13	1421	11.88	3.106

The table present the triple difference coefficient. The pre-period consist of the first three years of marriage (T0 - T3) and the post-period is defined as time periods beyond the third year of marriage (T4 - T10). Results are conditional on working in the pre-treatment period. Standard errors clustered at the individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4: Triple Differences: Effects of naturalization on labor market outcomes of foreigners

1.5 Mechanisms

The literature puts forward different potential mechanisms through which naturalization could lead to better economic integration. Obtaining the nationality of the host country can provide foreigners with a stable legal stay. This can in turn lead to behavioral changes such as country-specific human capital accumulation. Naturalization can lead to better job matches and reduce skill-downgrading for foreigners. Since part of the labor market, generally the public sector, is conditional on nationality, obtaining citizenship helps by giving foreigners access to those jobs. In addition, naturalization can reduce the hiring costs of foreigners in cases where there are administrative and financial costs on employers hiring foreigners. Finally, naturalization can play a role in reducing discrimination.

The unique setting exploited in this paper allows me to rule out many of the above-mentioned channels, while enabling me to better isolate the remaining ones. In fact, foreigners married to French are eligible for a spouse visa which provides them with

a long-term stay and authorization to work, as well as complete access to the welfare benefits in the country. In addition to this, employers do not face any additional burden in hiring them, implying that these foreigners have relatively easier access to the labor market. However, they still face three main constraints due to their nationality: first, restricted access to jobs conditional on the nationality; second, reduced access to the formal sector employment; and third, they can still be subject to taste-based and statistical discrimination. In this section, I test these channels and provide suggestive evidence for whether they matter.

Unrestricted access to the labor market

In France, Fougère and Safi (2009) based on GED (2000), document that around 20% of the labor market, of which a large part of the public sector, is not accessible to non-citizens. These restrictions also exist in other countries, such as the US and Canada. Since not all restricted positions can be identified in the data, I look at a broad proxy, the public sector employment. I test this channel by looking at the effect of naturalization on the probability of employment in the public sector.

The result reported in Column 2 in Table A.3 shows that there did not seem to have been a sizeable effect on public sector employment. Figure A.5.3 shows the dynamic effects over the ten years after marriage, and the null effect seems to hold over the whole period. This might be explained by the fact that only part of the public sector employment is conditional on nationality, and this proxy might hence be too noisy to detect an effect. It can also be driven by the fact that entry in the public sector[27] is costly, and is thus less likely at later stages in a person's career. There is a need to further analyze the employment in other jobs that are restricted to French nationals.

Informality

Foreigners tend to lack negotiating power when looking for a job, and employees take advantage of their situation to hire them informally. In France, the construction sector is one of the main sectors that is massively impacted by informal employment. I test the effect of naturalization for foreigners in the construction sector compared to those employed in other sectors. I use a triple difference specification, similar to equation (3), replacing $Mixed_i$ by a dummy ($Construction_i$) for whether an individual is employed in the construction sector or not. The results in table 1.5 show that on average, foreigners in the construction earned less than in other sectors.

[27] In France, public sector jobs are obtained through national competitions.

As a result of naturalization, they seem to obtain a much higher increase in their earnings, almost entirely explained by an increase in the number of hours worked.

VARIABLES	(1) Net annual earnings	(2) No of hours worked	(3) Hourly wages
Post x Construction	-4,336.32**	-253.37*	-0.65
	(1,872.08)	(128.54)	(0.65)
Post x Treat x Construction	5,179.08**	422.50**	0.02
	(2,406.37)	(182.99)	(0.80)
Observations	2038	2038	2038
Adj R-squared	0.65	0.43	0.61
Ind. FE	Yes	Yes	Yes

Clustered standard errors at the individual-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The pre-period consist of T0 - T3 and the post-period is defined as T4 - T10.

Table 1.5: Effect of naturalization in the construction v/s non-construction sector

The heterogeneous analysis by sector, construction and others, are presented in Table A.4. It shows that among workers in the construction sector, naturalization seems to have a massive effect on the number of hours worked, with no effect on hourly wages. On the contrary, in the non-construction sector, the effect is solely in terms of an increase in hourly wages and no significant effect on the number of hours worked. Given that the construction sector is heavily affected by informal work, these results might be driven by an increase in declared work following naturalization. Apart from the gain for foreigners in terms of less precarious situations, this would also represent a fiscal gain for governments.

Reduced Discrimination

The literature on discrimination in hiring has extensively shown that employers tend to discriminate against foreigners or foreign-sounding names. In France, a recent study has shown that French individuals with foreign-sounding names had a 20 – 30 % lower chance of being called back when compared with a fellow citizen with a french-sounding name. Duguet et al. (2010), and Petit, Duguet, and L'Horty (2015) show that obtaining the nationality for a Moroccan-origin with a foreign-sounding name increases the call-back success rate of an application by 1.45%.

Non-citizens might be subject to statistical discrimination and taste-based discrimination if employers have a preference for their own group (French citizens). In that case, naturalization can help foreigners in reducing nationality-related taste-based discrimination. It can also potentially send a signal of higher integration and language proficiency. A proxy to test this channel is to look at the benefits of naturalization for foreigners from non-francophone compared to francophone countries. Having the French nationality, irrespective of the foreign-sounding name, could help send a signal of better language skills, and better integration in general. In this case, the benefit of obtaining the nationality would be lower for foreigners coming from francophone countries compared to those with a nationality from a non-francophone country.

To test this, I adopt a triple difference approach and estimate an equation similar to equation (3). Instead of $Mixed_i$, in this setting, I include a dummy for having the nationality of a non-francophone country. Column 1 of Table 1.6 shows the coefficient of the interaction terms. A foreigner from a non-francophone country tends to have lower annual earnings on average, even if not significant. Obtaining the nationality significantly increases their earnings compared to foreigners from francophone countries. This seems to be driven by a higher increase in hourly wages for those from non-francophone countries. This suggests that naturalization can help to reduce discrimination through a signaling effect.

VARIABLES	(1) Net annual earnings	(2) No of hours worked	(3) Hourly wages
Post x	-2314.73	-69.52	-1.57***
Non-Francophone	(1,755.87)	(136.89)	(0.54)
Post x Treat x	6436.08*	124.69	3.08**
Non-Francophone	(1432.9)	(164.67)	(1.39)
Observations	2040	2040	2040
Adj R-squared	0.65	0.43	0.62
Ind. FE	Yes	Yes	Yes

Clustered standard errors at the individual-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The pre-period consist of T0 - T3 and the post-period is defined as T4 - T10.

Table 1.6: Effect of naturalization for non-francophones v/s francophones

1.6 Conclusion

Given the known benefits of economically well-integrated migrants, efforts should be put in further integrating them into the labor market of the host country. One of the policies at the disposal of every government is the naturalization process of migrants. Due to the rising fear towards migrants, countries tend to become stricter in terms of their naturalization rules. In the same line, France has increasingly hardened the rules and thus restricting the path to naturalization. The channel of naturalization through marriage, traditionally thought to be a natural process for well-integrated citizens, has not been spared by the tightening of rules.

In this paper, I exploit such a reform in the law of naturalization through marriage in France in 2006 as an exogenous shock on mixed married couples in France. To the best of my knowledge, it is the first paper to exploit a national-level reform that provides a quasi-experimental setting, allowing me to overcome the main issues of the existing literature: endogeneity, selection and reverse causality. Using a difference-in-differences strategy, I show that naturalization has a positive effect on annual earnings. This is explained by a positive effect on the number of hours worked, as well as the hourly wages. A gender decomposition reveals that the effects on earnings are stronger for women as compared to men. The effect is driven by an increase in the number of hours worked for men, and an increase in hourly wages for women.

Of the potential mechanisms put forward by the literature for the positive association between naturalization and labor market outcomes, unrestricted access to the local labor market, as proxied by public sector employment, does not seem to have played a role. Instead, there are suggestive evidence that naturalization helps in reducing informal employment, hence representing a gain for the foreigners as well as the host country. Naturalization also helps in reducing discrimination by signaling better language proficiency, and integration. In all, these results confirm the relevance of naturalization as a powerful tool to foster integration.

Bibliography

Baker, Michael and Dwayne Benjamin (1994). "The performance of immigrants in the Canadian labor market". In: *Journal of labor economics* 12.3, pp. 369–405.

Bertrand, Marianne and Esther Duflo (2017). "Field experiments on discrimination". In: *Handbook of economic field experiments*. Vol. 1. Elsevier, pp. 309–393.

Bratsberg, Bernt, James F Ragan Jr, and Zafar M Nasir (2002). "The effect of naturalization on wage growth: A panel study of young male immigrants". In: *Journal of Labor Economics* 20.3, pp. 568–597.

Chiswick, Barry R (1978). "The effect of Americanization on the earnings of foreign-born men". In: *Journal of political Economy* 86.5, pp. 897–921.

d'Albis, Hippolyte, Ekrame Boubtane, and Dramane Coulibaly (2016). "Immigration policy and macroeconomic performance in France". In: *Annals of Economics and Statistics/Annales d'Économie et de Statistique* 121/122, pp. 279–308.

Dell'Aringa, Carlo, Claudio Lucifora, and Laura Pagani (2015). "Earnings differentials between immigrants and natives: the role of occupational attainment". In: *IZA Journal of Migration* 4.1, p. 8.

Duguet, Emmanuel et al. (2010). "Are young French jobseekers of ethnic immigrant origin discriminated against? A controlled experiment in the Paris area". In: *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, pp. 187–215.

Dustmann, Christian and Francesca Fabbri (2003). "Language proficiency and labour market performance of immigrants in the UK". In: *The Economic Journal* 113.489, pp. 695–717.

Dustmann, Christian and Tommaso Frattini (2014). "The fiscal effects of immigration to the UK". In: *The economic journal* 124.580, F593–F643.

Fougère, Denis and Mirna Safi (2009). "Naturalization and employment of immigrants in France (1968-1999)". In: *international Journal of manpower* 30.1-2, pp. 83–96.

Freedman, Matthew, Emily Owens, and Sarah Bohn (2018). "Immigration, employment opportunities, and criminal behavior". In: *American Economic Journal: Economic Policy* 10.2, pp. 117–51.

Gathmann, Christina and Nicolas Keller (2018). "Access to citizenship and the economic assimilation of immigrants". In: *The Economic Journal* 128.616, pp. 3141–3181.

GED (2000). "Une forme méconnue de discrimination: les emplois fermés aux étrangers". In: *Groupe d'Etudes sur les Discriminations*.

Hainmueller, Jens, Dominik Hangartner, and Dalston Ward (2019). "The effect of citizenship on the long-term earnings of marginalized immigrants: Quasi-experimental evidence from Switzerland". In: *Science advances* 5.12, eaay1610.

Iacus, Stefano M, Gary King, and Giuseppe Porro (2012). "Causal inference without balance checking: Coarsened exact matching". In: *Political analysis* 20.1, pp. 1–24.

Lochmann, Alexia, Hillel Rapoport, and Biagio Speciale (2019). "The effect of language training on immigrants' economic integration: Empirical evidence from France". In: *European Economic Review* 113, pp. 265–296.

Mastrobuoni, Giovanni and Paolo Pinotti (2015). "Legal status and the criminal activity of immigrants". In: *American Economic Journal: Applied Economics* 7.2, pp. 175–206.

Meng, Xin and Robert G Gregory (2005). "Intermarriage and the economic assimilation of immigrants". In: *Journal of Labor economics* 23.1, pp. 135–174.

Meng, Xin and Dominique Meurs (2009). "Intermarriage, language, and economic assimilation process: A case study of France". In: *International Journal of Manpower* 30.1-2, pp. 127–144.

Petit, Pascale, Emmanuel Duguet, and Yannick L'Horty (2015). "Discrimination résidentielle et origine ethnique: une étude expérimentale sur les serveurs en Île-de-France". In: *Economie prevision* 1, pp. 55–69.

Safi, Mirna and Godfrey Rogers (2008). "Intermarriage and assimilation: Disparities in levels of exogamy among immigrants in France". In: *Population* 63.2, pp. 239–267.

Steinhardt, Max Friedrich (2012). "Does citizenship matter? The economic impact of naturalizations in Germany". In: *Labour Economics* 19.6, pp. 813–823.

Weil, Patrick (2002). *Qu'est-ce qu'un français?* Grasset.

Chapter 2

Post-colonial Trends of Income Inequality: Evidence from the Overseas Departments of France

Abstract

Most ex-colonies have gained their independence during the decolonization waves in the last century. Recent research on the colonial legacy in terms of inequality has thus mostly focused on these independent states, overlooking the territories which have been assimilated by their ex-colonizers. This chapter analyzes the post-colonial inequality in four such territories- La Réunion, Guadeloupe, Martinique and Guyane. Drawing on a new income tax dataset put together in this chapter, I study the evolution of income inequality in the four oldest French colonies, now overseas departments of France, since their decolonization in 1946 until recent years. I find that the top 1% income share rapidly declined since decolonization and stabilized at the level of metropolitan France in the recent decades. Despite this general catch-up, the top 10% income share remained consistently higher than in the metropolis. Matching contemporary fiscal data to the population census, I uncover underlying cleavages and show that public-sector employment and metropolitans are over-represented at the top of the distribution in the overseas departments. This is partly due to the existence of a metropolitan income premium, likely to further exacerbate the perception of inequality in a post-colonial setting.

2.1 Introduction

The colonial origins of inequality is well-established in the literature. This is unsurprising as colonies, especially slaves ones, have been built on extremely unequal foundations. The existing literature has substantially broadened our understanding of the potential underlying factors explaining persisting inequality in now independent states. However, the situation of territories which have gone through a less common path of decolonization, by assimilation to the metropolis, remains largely unaddressed by the literature. This paper attempts to fill this gap by investigating the post-colonial levels of inequality in four such territories- La Réunion, Guadeloupe, Martinique and Guyane, attached to France in 1946 after three centuries of colonial domination by the latter. In addition to the measurement of inequality, I shed light on the underlying racial cleavage by comparing the native population to the ex-colonizers, now metropolitans.

The existing literature consists of cross-country analyses that look at the economic, political and institutional settings inherited from the colonial period to explain the persisting inequality observed after independence (Nunn, 2008; Engerman and Sokoloff, 2002; Engerman and Sokoloff, 2005). Angeles (2007) finds that independence in "settler colonies"[1] did not mark the end of the income concentration in the hands of the white-descendants minority. In the same line, Engerman and Sokoloff (2005) identify the colonial population composition as a major factor in the setting up and persistence of more or less unequal institutions, even after independence.

A second strand of the literature has focused more on country-specific colonial and post-colonial evolutions of income inequality. However, no single pattern can be discerned[2]: most of the literature on Africa[3] and Latin America[4] shows a declining income concentration leading up to independence followed by a rapid increase immediately after[5]. On the other hand, countries like Mauritius, Singapore and India witnessed the opposing trends before and after independence[6]. The four territories studied in this paper provides a different case altogether, as detailed in Section 2.2.

[1] Countries are identified as settler colonies if European settlers accounted for 10 to 30% of the colonial population

[2] World Inequality Lab

[3] Anthony B Atkinson, Piketty, and Saez (2011), Anthony B Atkinson (2014), Alvaredo, Cogneau, and Piketty (2021)

[4] Williamson (2010); Williamson (2015))

[5] Post-colonial data is not available in some of these countries.

[6] A. Atkinson (2010); Anthony B Atkinson (2011); Banerjee and Piketty (2005)

Amid a general wave of decolonization through independence, the political status of these territories transformed from the "old" four colonies of France to its "overseas departments" in 1946, with an explicit agenda of bringing more equality within the territories as well as with the metropolis. These territories thus represent peculiar post-colonial contexts. On the one hand, one could expect high post-colonial levels of inequality, as they were slave colonies, and the colonial population composition corresponds to the "settler colonies" category (Angeles, 2007). On the other hand, the fact that they remained part of the metropolis could imply a convergence to the national-level, and hence lower inequality levels. All these factors combined make predictions about the expected post-colonial inequality situation, based on the existing literature, quite unclear. The present paper thus contributes by shedding light on the evolution of income inequality in peculiar post-colonial settings that remained attached to their ex-colonizers.

Almost 75 years after their formal decolonization, the overseas departments of France remain outliers on all socio-economic aspects. Extremely high levels of unemployment and poverty rates have led many to argue that departmentalization has failed to reach one of, if not its main goal. However, the current state of the literature around this question is too limited to have an informed debate. Maurin and Bernier (2013) ranks La Réunion (with a Gini of 0,53), Paris (0,5) and Martinique (0,47) as the most unequal departments of France, compared to an average Gini index of 0,31 in metropolitan France. Official publications also point towards the prevalence of high levels of inequality in the recent decades (Michel, Theulière, and Missègue, 2010). These few studies provide a snapshot of the current situation in these territories, without providing an in-depth analysis of its transformation from the colonial period and the past remnants today.

The lack of studies on this topic simply results from an acute lack of data on these territories. So far, analyses on inequality in the overseas departments have primarily relied on survey data that are only systematically available as from the mid-1990s[7]. These have largely restricted the coverage period of previous studies, confining them to cross-sectional or contemporary analyses. In addition, the lack of coordination among local statistical bodies, and thus the lack of comparable data, meant that these departments have hardly been analyzed together. I thus contribute by building a novel dataset based on income tax tabulations at the overseas departmental-level since the 1950s until 2014[8], allowing me to overcome the existing challenges.

[7] While they are available in metropolitan France since the 1960s, they are only recently fully extended to the overseas departments

[8] See Appendix B.2 for details of period coverage

This paper is divided into two parts. In the first part, I estimate the historical evolution of income inequality in the four overseas departments since their departmentalization. I combine fiscal data with population censuses and income data, applying the Generalized Pareto Interpolation technique (Blanchet, Fournier, and Piketty, 2017) and following the corrections made in Piketty (2001), as detailed in Section 2.3. The analysis in this part is at the tax-unit level using a fiscal income definition. Departing from the national-level focus in Garbinti, Goupille-Lebret, and Piketty (2018), I conduct the first regional-level analysis. I thus provide inequality series that are comparable across the four territories and with the national-level estimates.

I find that, despite an increasing trend of inequality in the immediate post-colonial period, the overseas departments witness a spectacular decline in their levels of inequality since the 1960s, as shown in section 2.4. The top 1% income share was as high as 25% of total income in the 1960s and declined to a level comparable to that of the metropolis at 10% in recent years. Similarly, the top 10% share decreased from around 60% in the mid-1980s to 40% today. This decline in the level of inequality can be attributed to general economic factors, such as the decay of the sugar industry, as well as institutional ones, such as the setting up of the public sector, regulated migration and the minimum wage.

However, despite a complete convergence in the top 1% income share to the national-level, the top 10% income share in the overseas departments have stabilized at a higher level than the french-level in the last decade. This resonates with the continued perception of high levels of inequality as signalled by the recurrent protests, strikes and riots[9] in these territories. The disparity between the overseas departments and France lies largely in labor market differences. In fact, evidence suggests a polarization of the labor market between a low-paid private sector and highly paid public sector.

The second part of the paper analyses the underlying cleavages in the contemporary period. I exploit individual-level fiscal data matched with the population census in section 2.6, focussing on labor income inequality. The results show two peculiar elements in the overseas departments. First, there is an overly important public sector compared to metropolitan France. Second, I provide suggestive evidences of

[9]Some of the major riots/protests in the overseas departments: Violent riot in 1959 Martinique leading to anti-colonial protests against oppression; Riot in Guadeloupe in 1967 which erupted due to racism, leading to workers protests demanding improved economic conditions; Riot in 1996 in Guyane which started with demands for the local education system; Riots in La Reunion in 2005 and 2012 and general strike in the Antilles in 2009 against the high cost of living and the unacceptable low standard of living

the existence of a "metropolitan premium" in the overseas departments. I argue that these factors emanate from the peculiar post-colonial setting of being attached to a metropolis. It is even more salient in the overseas departments due to the added post-colonial complexity of the presence of metropolitans at the top of the distribution, to the already existing ethnic frictions inherited from the colonial period. This may have important consequence, as seen in the post-colonial South African case of ethnic segregation and the ensuing rapidly increasing inequality trends (Alvaredo and Anthony B Atkinson, 2010).

In section 2.7, I conclude that this paper contributes to the post-colonial literature by first, taking a regional perspective to uncover specificities masked by aggregate estimates, in this case stemming from the colonial past and secondly, by analyzing a post-colonial setting so far largely under-studied. It also contributes to a more informed debate on the issue of inequality in France and its overseas departments. It has substantial policy-relevance given the renewed political will to tackle inequality in the overseas departments as seen by the recent enactment of the bill on "Real equality for overseas department"[10]. It can also contribute to the debate on quotas and positive discrimination in favor of natives in public employment in the overseas departments.

2.2 Background

The overseas departments, once known as the "four old colonies" of France, were among the first colonial possessions of the French empire in the 17$^{\text{th}}$ century,[11] and share a long common history with France. These ex-colonies present interesting characteristics owing to their peculiar colonial and post-colonial history. They have mostly been populated by colonial settlement, slaves and indentured laborers. There have been two major turning points in the history of these territories: first, the abolition of slavery in 1848 through which the population were granted the French citizenship; and second, a century later, with the transformation of these colonies into French departments. This process of decolonization by assimilation to the French Republic in 1946[12] occurred after three centuries of colonial domination and at a time of great uncertainty of the subsequent path of the French empire in Africa.

Politically, these colonies had parliamentary representation in Paris as early as 1789[13], albeit with periods of interruption under the different subsequent regimes. In

[10] Loi n$^{\text{o}}$ 2017-256 passed on the 28th February 2017 on "Egalité réelle des outre-mers"
[11] Even before some metropolitan territories such as Nice, Savoy and Corsica
[12] Loi no 46-451 du 19 mars 1946 also known as the "loi de departementalisation"
[13] The Constitution of 1795[14] further integrated these four colonies and few others-Saint-Domingue,

addition, quite remarkably different from other French colonies, the population of the four old colonies[15] were granted full-fledged French citizens and universal suffrage as from 1848. All these factors combined gave these territories a unique status within the French empire.

Despite being grouped under the umbrella term of "four old colonies" and the overseas departments of France in the post-colonial period, they hide different realities. These differences are rooted in the colonial era- while the Antilles and La Réunion, to a lesser extent, were used for slave trade and plantations, Guyane was initially a prison, later famous for gold exploitation and the French space center in the post-colonial period. These territories also differ in their population compositions. To begin with, they had different shares of white settlers- In 1848, at the time of the abolition of slavery, Guyane counted 6% of its population as white while the Antilles had a share of 10% and an even greater share of 20% in La Réunion (See Appendix B.1). Demographically, the Antilles and La Réunion shared the common feature of a binary white-black population until 1848, thereafter witnessing an influx of Indian and Chinese populations [16]. Guyane has always been a marked by a high influx of population from poorer neighboring countries- in search of better economic conditions.

When compared to the metropolis, there have been undeniable differences on the socio-economic front. At the turn of the 20th century, these four territories were marked with deep social divides on various lines. As part of the colonial heritage, the segregation between the white economic elite[17] and the African and Asian descendants persisted in la Réunion and the Antilles. Guyane, on the other hand, faced the challenge of its border disputes, mass migration inflows and dismal inequality. In view of this situation, the law of departmentalization and the underlying assimilation process, brought forward by local intellectuals, was sought to bring increased legal, social and economic equality within these territories as well as with their metropolitan counterparts.

The immediate post-colonial period witnessed a generally alarming socio-economic situation in the overseas departments. These territories were marked by high illiteracy

Sainte-Lucie et Tabago, Ile-de-France and French Indian colonies into the French Republic, subjecting them to the French constitutional law and dividing them into administrative departments. See https://www.conseil-constitutionnel.fr/les-constitutions-dans-l-histoire/constitution-du-5-fructidor-an-iii

[15] Irrespective of their origins. There is no difference between the "colonial citizens" and the metropolitan citizens unlike the system of indigenats in the other colonies

[16] few in the Antilles and more important in La Réunion

[17] commonly known as the Békés in the Antilles

rates, low sanitation level, low life expectancy and the decline of the sugar industry. In the face of this alarming situation, as from the 1960s, the French government undertook a step-by-step action plan to gradually tackle the problems at hand. These consisted of the setting up of post-colonial institutions to tackle the most pressing issues- initially health and sanitary issues, followed by the social and eventually the economic aspect.

Being attached to the metropolis led to two main peculiarities in the post-colonial context. First, in the absence of the development of a local productive sector, the French government largely invested in the labor market to absorb the large share of unemployed working age population. These territories thus witnessed massive job creation in the public sector, remunerated with a wage premium[18]. Second, being fully integrated in France entailed a free movement of population between the metropolis and the overseas departments. This has led to waves of out-migration from the overseas departments towards the metropolis and more importantly, significant inflow of metropolitans, mostly in the top occupations in these territories.

2.3 Data & Methodology

2.3.1 Data

In order to establish the long-term evolution of income inequality in the overseas departments, this paper gathers and exploits annual income tax data published by the French tax administration. Despite some caveats of tax data[19], it constitutes a valuable source of data for the analysis of income inequality. It is especially so in the DOM, since the only other potential source of data is the household budget surveys which face major limitations in many respects. In general, survey data is known to suffer from issues at the upper end of the distribution, which takes the form of top coding or under-reporting. In addition, the period coverage of surveys in the DOM is significantly shorter and at a lower frequency[20] as compared to the annual tax data which are available since its onset in the mid-20$^{\text{th}}$ century.

I construct a novel historical dataset of income tax data in the overseas department, thus contributing to the general pool of data available for these territories. It takes the form of tabulated tax data at the departmental level from the 1950s to 2014

[18] A higher pay relative to the metropolis pay

[19] For instance, issues of tax avoidance and evasion or the focus on pre-tax and transfer income inequality which does not take into account the redistributive efforts of public policies, especially in the DOM.

[20] The Enquête Budget de Famille only starts in 1995 and are available every five years

intermittently. These data contain information about the number of tax filers and the total income in the different brackets of income[21]. These income tax data have been collected from different sources and can be categorized accordingly in three periods:

i) 1950s - 1985: *Annuaire Statistique* of the Overseas Departments and *INSEE*
ii) 1986 - 1998: *Etats 1921*- Centre des Archives Economiques et Financières (CAEF)
iii) 2000 - 2014: *Direction Générale des Finances Publiques (DGFiP)*

Starting from the most recent period (2000 - 2014), income tax data is obtained from the online resources of the Direction Générale des Finances Publiques (DGFiP) for the four departments[22]. Data for the period 1986 - 1998 for all four territories are obtained in the form of paper-based tables, annually published in a pamphlet format. Known as the "Etats 1921", it was originally published for internal use by the Ministry of Finance[23]. For the preceding period, the income tax data is gathered from the various Annuaire Statistiques of La Réunion, Guadeloupe, Martinique and Guyane, published by INSEE over the period 1950 - 1974.[24]. Between 1972 - 1985 and 1988, partial tax data for La Réunion is retrieved from a retrospective compilation of statistical data from an INSEE publications. Unfortunately, the data reported are not as detailed as the previously-mentioned sources, as they were only published for expository purposes. The publication only reported the number of taxable taxpayers per income brackets, with no information about the corresponding incomes in the brackets. The additional corrections made to these partial data in order to estimate the income distribution is laid down in Appendix B.2.

The availability of data for the different departments are more or less sparse and do not cover the entirety of the period for all departments. A summary of the availability of the data over the whole period is presented in Figure B.2.1. As far as possible, the latest available tabulations are used in this paper to account for most updated corrections made to the tax data[25]. The comparability of the publications across time is generally consistent, except for changes in income definition, which is described in Appendix B.4.

[21] There has been noticeable changes in terms of the number of thresholds reported over time. The aim for this frequent update of the number of threshold is normally to provide more detail at the upper end of the distribution as taxpayers report increasingly higher taxable income.

[22] As of this date, data for 2004-2014 can be retrieved online from `www.impots.gouv.fr`

[23] These data do not violate any statistical confidentiality rule as it includes a large number of taxpayers. These data concern groups of more than 11 persons.

[24] The latest year corresponding to income perceived in 1972

[25] The tax administration normally publishes income tax data on income perceived in year n in both the following year at 31/12/(n+1) and the year after- 31/12/n+2. The latter is in principle the most up-to-date data as it takes into account tax audits, tax reliefs and changes in family status which occurs in the year after the imposition.

Moreover, as explained in more details in Appendix B.3- Control Total for Population, the unit of analysis in the tabulation tax data is the tax unit. While it is conceptually close to a household unit, which is the preferred unit of analysis in economic surveys, tax units refer to a person or group of persons that fills a unique tax form. Hence, the definition of household does not align perfectly to tax units[26]. As is done in this literature and for the sake of consistency over time, a tax unit is estimated as an adult above 20 years of age or a married couple (see Appendix B.3).

Apart from income tax data, this analysis also relies on population and income data. Demographic data are primarily obtained from population censuses over the whole period[27]. Departmental-level income estimates are primarily obtained from national accounts compiled by French Statistical Institute, INSEE. This covers the entire period for La Réunion and unfortunately exists only as from the 1970s for Guadeloupe, Martinique and Guyane. For the previous period, the national income series are estimated based on the known series of La Réunion (the assumptions are laid in Appendix B.4).

2.3.2 Methodology

Following the work of Piketty (2001) and Garbinti, Goupille-Lebret, and Piketty (2018)[28], this paper establishes a thorough study of top incomes at the overseas departmental-level. Given the truncated nature of the tabulated tax data, a generalized non-parametric Pareto interpolation technique Blanchet, Fournier, and Piketty (2017) is applied to the data.

In France, prior to 1985, only tax units subject to taxation were subject to income tax declarations. While it becomes mandatory as from the mid-1980s to fill in a tax form, it is only gradually applied in the overseas departments. Over time, a greater proportion of tax units is captured in the tax data, as seen in Figure 2.1. Hence, in order to estimate the whole income distribution, there is a need to estimate the total number of tax units and total income over the whole period, had every tax unit been required to fill in a tax form. These components, commonly known as control total for population and income, are detailed in this section.

[26]For instance, in the case whereby a cohabiting unmarried couple would constitute a single household but two tax units

[27]The population census are available for the following years: 1954, 1961, 1967, 1974, 1982, 1990, 1999, 2009 and 2014.

[28]Refer to the Appendices of GGP2018 for a detailed explanation of the estimations and corrections made.

Population Estimates

In the French fiscal system, individuals can choose to declare their income separately from their parents' declarations as from the age of 18 and a separate declaration became mandatory as from 21 years of age[29]. While single individuals fill independent declarations, married or PACSed[30] couples are required to jointly fill a tax form. Hence, a close estimate of the total number of tax unit would be the sum of single individuals and the number of married (or PACSed) couples[31]. Given the flexibility on the initial age of fiscal declarations and to be consistent with the literature, adult population is defined in this paper as individuals above 20 years of age. The long-run trends of adult population and total estimated tax units are presented in Appendix B.3.

Income Estimates

Similarly, there is a need to estimate the total fiscal income which would have been reported if all the tax units were required to fill a tax form. In order to obtain a coherent series over the long-run, I adopt the external control approach, which consist in correcting the national income accounts for non-household income and other non-relevant incomes to obtain the total taxable income. The relationship between the national income and the taxable income is shown in table 2.1. Fiscal income may hence diverge from national income due to production taxes and the part of income not subject to taxation and thus not declared in the tax data[32].

I first build a long-run series of national income at the overseas departmental level since the mid-20th century until recent years. Calibrating on the relationship between the national income and taxable income at the national French level, I then estimate taxable income at the DOM-level over the whole period (See details in Appendix B.4). While this process allows me to obtain an estimate of taxable income, the

[29] 25 years of age for students

[30] A civil solidarity pact- a contractual form of civil union

[31] Note that this only gives an approximate estimation of the total number of tax units since there may be cases of young students above 20 years of age attached to their parental tax unit or in cases of a marriage (or divorce) during the year would entail three declarations in total- two separate declarations for the income received before the marriage (or after divorce) and one declaration for the couple thereafter (before the divorce). However, despite not being a perfect estimate, it provides a precise enough estimate. A discussion on the choice of the age of the adult population and a detailed explanation of the steps in the estimation of is made in (Appendix B.3).

[32] It may include imputed rent (rental income from owner-occupied housing), employers' and employees' social security contribution, tax-exempt life insurance income and other tax-exempt income, for instance interest paid to deposits and savings accounts and non-taxable transfer payments. On the capital front, fiscal income also excludes corporate retained earnings and corporate taxes.

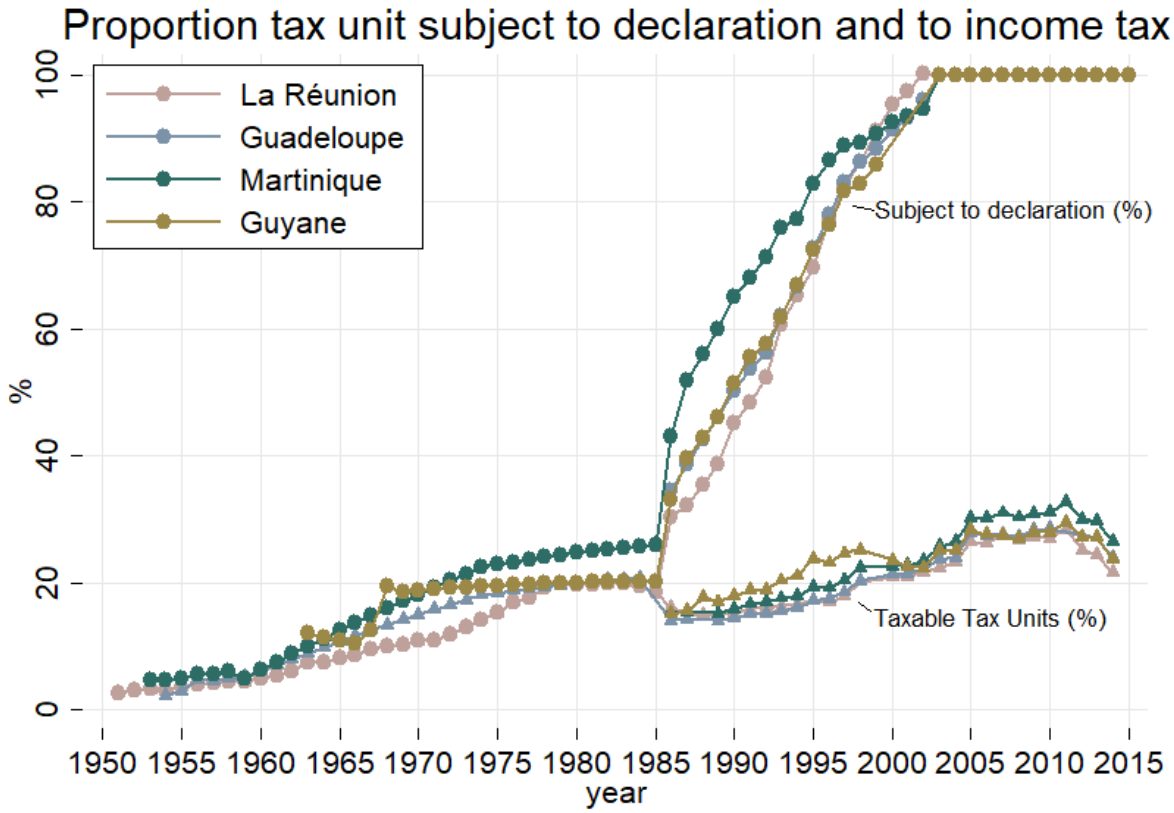


Figure 2.1: Proportion of tax declarations (1951 - 2014)

Balance of Primary Income
(-) Non-household incomes
Household sector total income
(-) Items not included in the tax base
Household Gross income
(-) Non-declared income
(-) Non-filers
Declared taxable income of filers

Table 2.1: Relationship between National income and Taxable income

definition of income of interest is the fiscal income. The latter refers to the income reported in the tax declarations before any adjustments. Fiscal income is preferred as taxable income is sensitive to changes in the tax administration and changes in deductions schemes over time, potentially leading to biased estimates of trends. Thus, corrections for the following deductions allowed for in the French tax laws are made to the taxable income series in order to estimate fiscal income:
i) A 10% lump-sum deduction for professional expenses of wage earners, currently capped at 12 183 € per member of the tax unit.
ii) An additional 20% deduction for wage income (up to a ceiling) which has been repealed in 2006.
Apart from the corrections made for these deductions, the series also take into account the capital gains based on Garbinti, Goupille-Lebret, and Piketty (2018). The resulting estimated average fiscal income in France and the overseas departments are presented in Figure 2.2[33].

Given the nature of tabulated tax data, estimates for the very top of the distribution (top 1, 0.1 and 0.01%) are available since the 1950s while the top 10% income share can be only be precisely estimated as from the mid-1980s. The tabulations in the recent decade allows for an estimation of the bottom 50% share, except for Guyane. In terms of data availability, the beginning of the period until 1986 is intermittently covered in the different departments, La Réunion having the most complete data[34]. An almost uninterrupted series is established for all four overseas departments from the mid-1980s up to 2014. The results for the overseas territories are put in perspective by comparing them to the French series by Garbinti, Goupille-Lebret, and Piketty (2018).

2.4 Results

2.4.1 Average Fiscal Income

The overseas departments were approximately 40% poorer than France at the beginning of the period. While France experienced rapid growth during the "Trente Glorieuses"[35], the overseas departments grew at a much lower pace, Guyane faring

[33] See Appendix B.4 for the trend of the taxable income based on the estimation described in the previous section and fiscal income based on the above-mentioned corrections. There is a clear jump in taxable income in 2006 due to the repeal of the 20% deductions for wage income.
[34] See Appendix B.2 for details about data coverage
[35] The 30-year period of post-war boom.

the worst[36]. This has widened the gap between the overseas departments and France to around 50-70% in the 1980s. Partly due to a stabilization of average fiscal income in France and partly to the increased pace of growth in the overseas territories as from the 1990s, there has been a slight convergence, with the gap stabilizing at around 30% (around 10 000 € per year difference per tax units in actual terms) in the recent years.

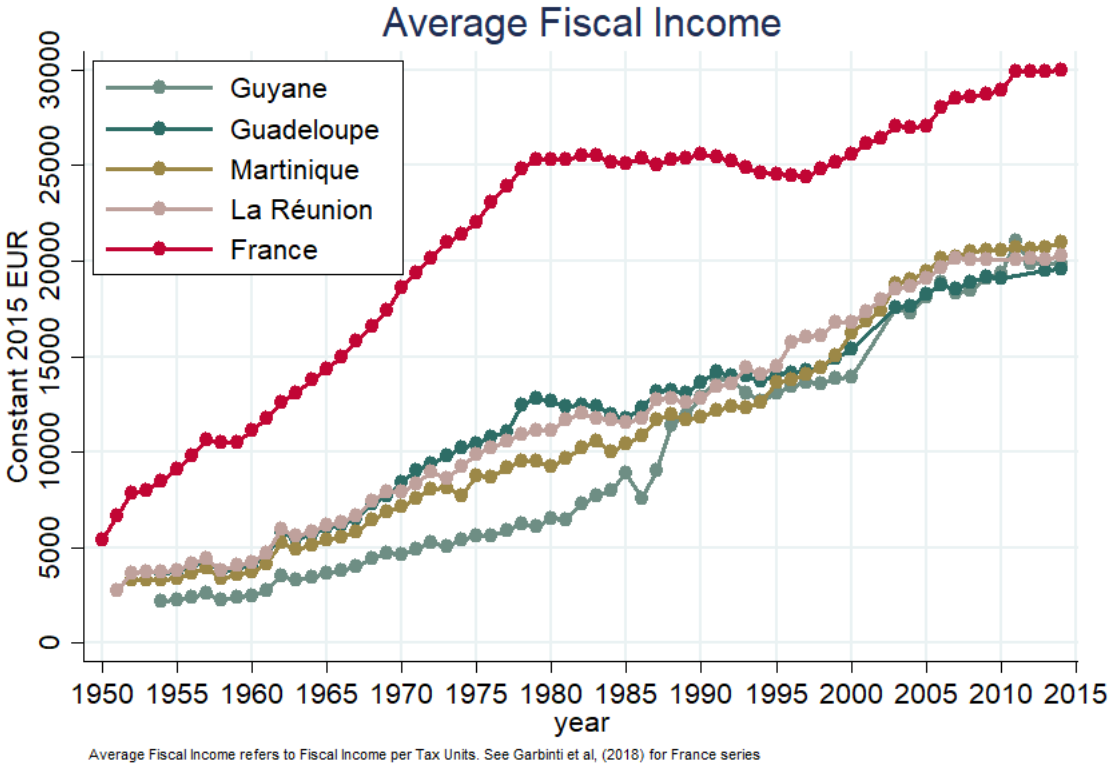


Figure 2.2: Average Fiscal Income (1950 - 2014)

2.4.2 Top Income Thresholds

Figure 2.3 and 2.4 depicts the minimum income required to be part of the top 10%, top 1% and top 0.1% of the distribution respectively. It is clear that the gap in the top income thresholds between the overseas departments and the metropolis are smaller than the one observed for the average fiscal income. The top 10% income threshold in the overseas departments have consistently been lower than the level in France. Despite the fact that the mid-1980s correspond to the period with the largest difference in average income between France and the overseas departments, the gap for the same years at the top of the distribution are 15 - 35%. This has

[36]Guyane's economy was very much fragile and dependent on the development of the Spatial Centre and the mass migration flows in the neighboring countries (Besson, 2017).

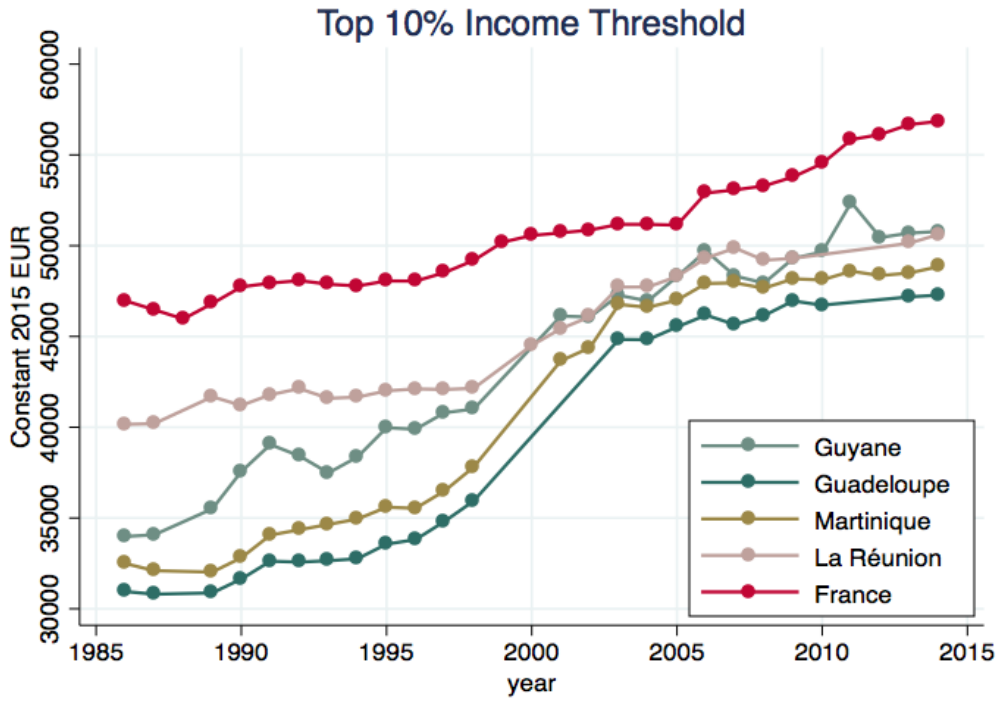


Figure 2.3: Threshold income of top 10% (1986 - 2014)

further reduced over the years, especially so in Guyane and the Antilles. In fact, in absolute terms, the 7 000 € - 16 000 € gap in the mid-1980s has narrowed down to 6 000 € - 10 000 € today. This translates to a relative gap in top 10% income of 15-30% in the mid-1980s to 10.5-17.5% today. This gap is 2 to 3 times smaller than the gap in the average fiscal income.

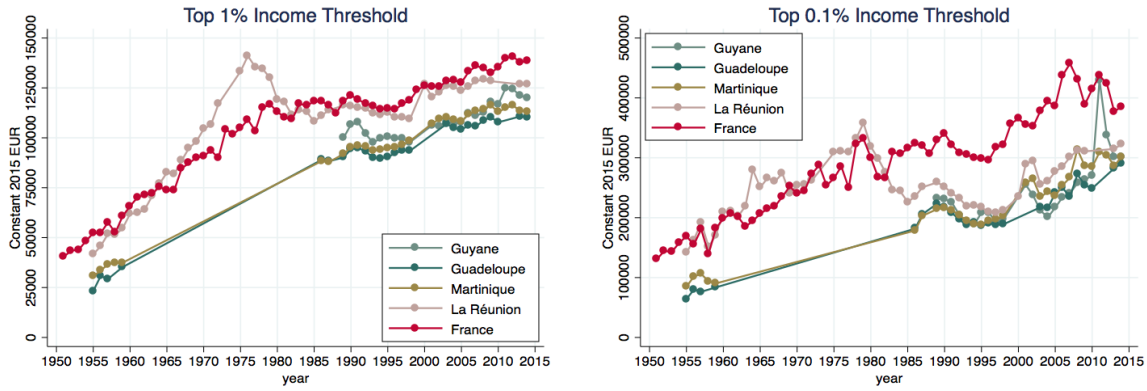


Figure 2.4: Threshold income of top 1% and top 0.1% (1950 - 2014)

The gap is even narrower at the very top of the distribution, especially for La Réunion (Figure 2.4). The top 0.1% threshold shows a slightly different trend, with almost no gap until the mid-1980s and a widening of the difference thereafter, but

the thresholds remain fairly close to the level of France. Altogether, figure 2.2 to 2.4 suggest that while the income at the top of the distribution in the overseas departments has remained close to the level in France throughout the period, the middle of the distribution has only moderately gained since the 1950s.

2.4.3 Top Income Shares

This section presents the results of the estimation of the top income shares[37] using the generalized Pareto method. Overall, the top income shares series spans over a 60-years period, from the 1950s to 2014.

Top 1% Income Shares

Figure 2.5 shows the top 1% income shares in the overseas departments[38] in comparison to France. Three main elements can be observed from figure 2.5. First,

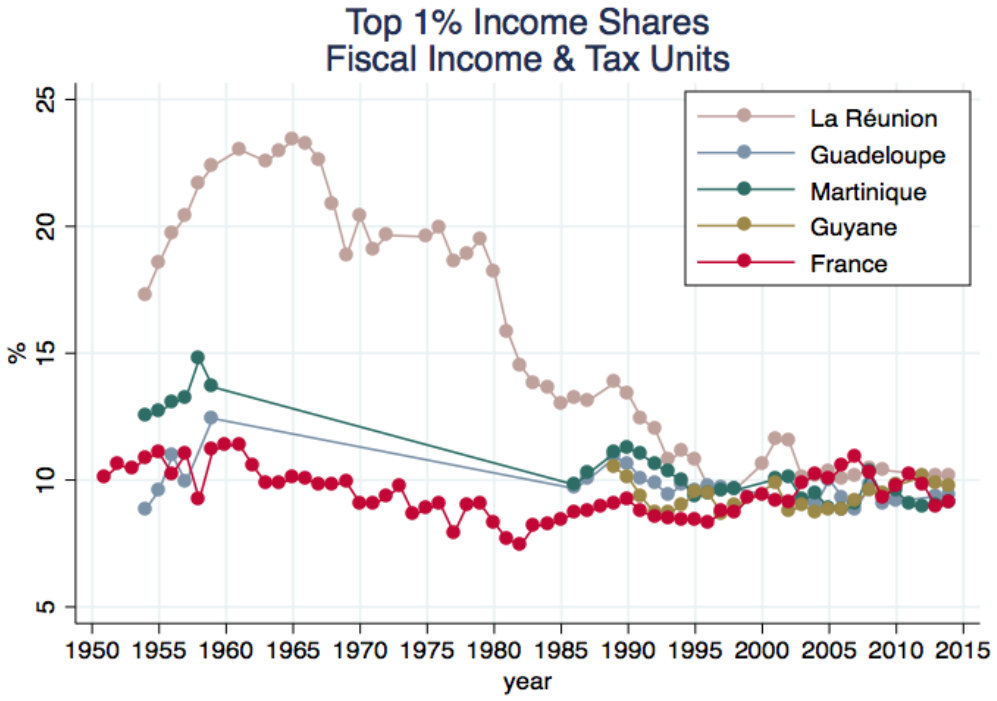


Figure 2.5: Top 1% Income shares (1951 - 2014)

[37] In order to understand the following series, one needs to grasp the concept of top income shares. As an illustration, in a perfectly egalitarian economy, the top 10% of the distribution would own 10% of total income. Similarly, the top 1% would own 1% of total income. If the share of the top 10% is estimated to be 20%, then the top 10% own twice the income they should have owned under a perfectly egalitarian economy.

[38] Since the 1950s for La Réunion and the Antilles with a gap in the data from 1960-85 in the latter territories and as from the late 1980s in Guyane.

the top 1% income shares in the overseas departments increased up to the 1960s, followed by a drastic decline and stabilization as from the 2000s[39]. There is an initial upward trend until the 1960s, peaking at 25%[40]. This extreme level of inequality is comparable to highly segregated societies such as Algeria under the French colonial rule (Alvaredo, Cogneau, and Piketty, 2021) or South Africa in the post-apartheid period (Alvaredo and Anthony B Atkinson, 2010). Thereafter, there has been an initial moderate decline until 1980 and a more rapid decline thereafter. Second, there are differences in the initial level of inequality between La Réunion and the Antilles, until the 1990s. Third, despite initial differences in the top 1% shares in the overseas departments and France, there has been a converging trend. In fact, the top 1% has stabilized at around 10% in the overseas departments, reaching the national level in the recent years.

Top 0.1% and Top 0.01% Income Shares

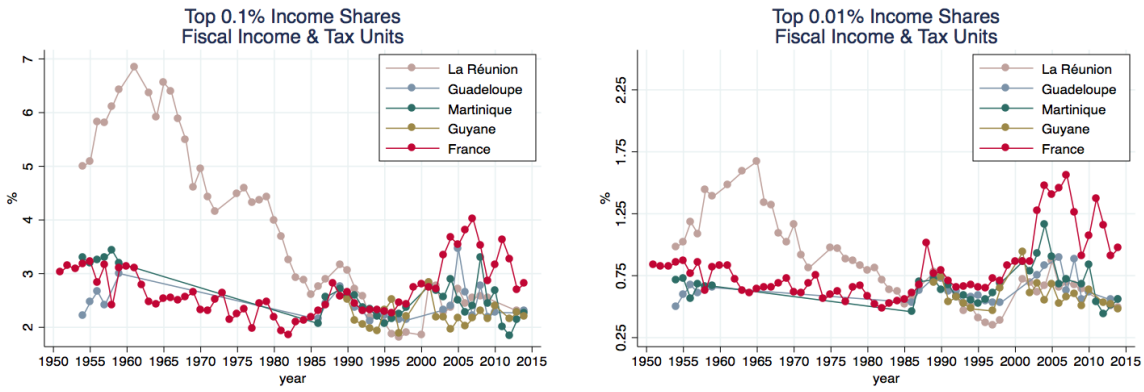


Figure 2.6: Top 0.1% and top 0.01% income shares (1951 - 2014)

Figure 2.6 shows the evolution in the income concentration at the very top of the distribution- the top 0.1% and the top 0.01% in the overseas departments and France. The shares were strikingly higher in La Réunion compared to the other overseas territories and France in the 1960s. Top 0.1% (0.01%) was at around 8% (1.7%), and reduced significantly to approximately 3% (0.8%) in the mid-1980s with a continued declining trend thereafter until the 2000s. Post 2000, the top 0.1 and 0.01% income shares of all four overseas departments hovered around the level of France. However, despite the complete convergence in the very top income shares to the metropolitan

[39]Based on the partial data for the Antilles and on the series of La Réunion which provides the most complete picture.

[40]France's had a similar level of top 1% income share in the inter-war period at 23%

level as seen in figure 2.5 and 2.6, top 10% shares remained higher than in the metropolis, as seen in the next section.

Top 10% Income Shares

The top 10% income shares followed a similar evolution as the top of distribution since the mid-1980s. The top 10% income shares were significantly higher in the overseas departments compared to France in the 1980s. This is especially the case in La Réunion, where top 10% income shares are above 60% and between 48 - 55% in the Antilles and Guyane. These levels of inequality are among the most extreme levels witnessed in the world. They are comparable to the Middle-Eastern regions in recent years and South Africa during and after apartheid (Alvaredo, Assouad, and Piketty, 2019; Alvaredo and Anthony B Atkinson, 2010). Similar to the top 1%, the levels of inequality are different between the four territories, La Réunion being the most unequal, followed by Guyane and the Antilles respectively. This period of high inequality in the overseas departments is followed by a general declining trend. As from the mid-1990s, there is a milder decrease and an eventual stabilization at the turn of the century.

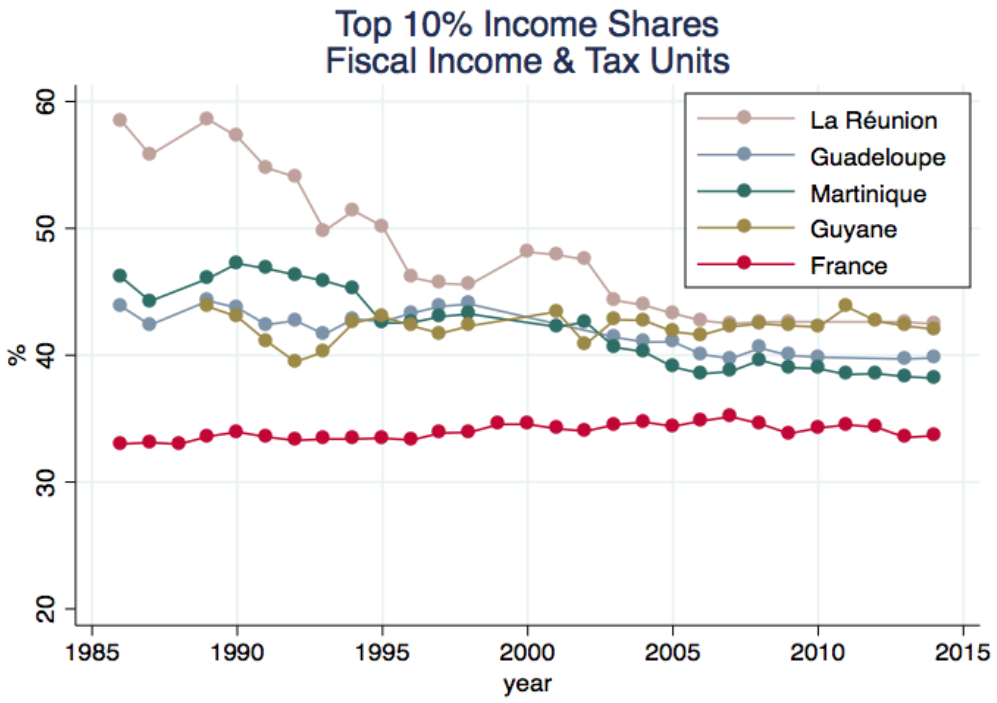


Figure 2.7: Top 10% income shares (1986 - 2014)

In contrast to the top 1% income shares, the top 10% income shares remain consistently higher in the overseas departments compared to France, despite the significant

declining trend. In the late 2000s, the top 10% share is around 33% in France compared to around 39-44% in the overseas departments, thus up to a 10% point difference. This goes in line with the official french statistical publications which concludes, based on survey data, that the overseas departments are one of the most unequal departments of France. Taken together, figures 2.5 and 2.7 imply that the higher level of inequality in the overseas departments compared to France is driven by the bottom 9% of the top 10% income group[41].

Bottom 50% Income Shares

This section lays down an estimate of the bottom 50% income share in La Réunion and the Antilles as from 2000[42]. In general, the share accruing to the bottom 50% is around 8-11% compared to 18% at the national level. Again, it is comparable to very unequal countries, namely the U.S. and the Middle East in recent years. This is in line with the extreme level of poverty in these territories. Since this paper looks at the pre-tax and transfer fiscal income, by definition, it does not include the informal sector and transfers from the government.

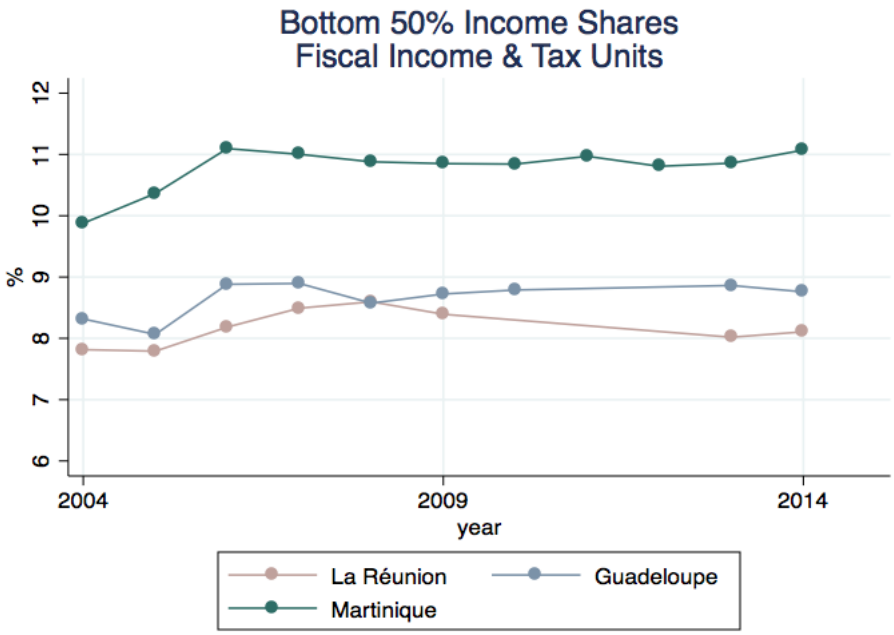


Figure 2.8: Bottom 50% Income shares

The overseas departments are, today, highly dependent on transfers from the metropo-

[41] Often denoted as the P90-P99

[42] Estimates prior to 2000 and for Guyane are less precise since the tabulation tax data does not allow to directly observe the bottom of the distribution and such estimation would require further assumptions.

lis and have among the highest number of public transfer dependencies per capita. Households in these territories also tend to live in communities and have a high financial dependence on close and extended family (Breton et al., 2009). While all these factors might correct the extremely low share of income accruing to the bottom 50% group, these reflect, at most, very precarious situations and should not be sought as the long-run sustainable solution in face of an unequal society.

2.5 Discussion

The main elements observed in Figure 2.5 to 2.8 are three-fold:
i) A rapid decline, followed by a stabilization of income inequality in the overseas departments[43] since departmentalization;
ii) Top 10% at a higher level in the overseas departments compared to metropolitan France; and
iii) Differences between the different overseas departments

In terms of the evolution of the inequality trends, first, the inequality trends in three main periods: An increasing trend in inequality since departmentalization until the 1960s, followed by a declining trend from the 1960s to the 2000s (with a sharper decline from the 1960s - 1990s and milder from the 1990s-2000s) and a stabilization of inequality thereafter. These evolutions are partly mechanical results of economic changes and can partly be attributed to institutional changes reflecting the French political will to address the situation in these territories.

1946 to 1960s: The local colonial economies were largely affected by the second world war due to the sudden detachment from the metropolis[44] leading to a period of severe blockage and a food crisis. Thus, in 1946, these territories were not only burdened by their colonial heritage but also by the impact of the war on the local economy. While the law of departmentalization was voted in 1946, there has been no sharp break between the colonial and post-colonial period in reality. Scholars consider the immediate post-departmentalization period until the mid-1960s as a period of status quo in these new french departments (Drozin, 2001). Thus, the starting points of the top income series of the overseas departments provide a fair insight into the degree of inequality at the end of the colonial period[45].

[43] Based on the top 1% income shares in La Réunion as the most complete data series exists for La Réunion. Few data points can be observed in the mid-50s for the Antilles and a full series as from 1986 in the Antilles and Guyane.

[44] France being under the German occupation

[45] The analysis for this period relies on the series estimated for the top 1% and top 0.1% income shares

The first decade after departmentalization marked the post-war recovery of the sugar production in the overseas departments (See Appendix B.5.1). At the same time, the public sector in the overseas departments was set up. Given the high level of illiteracy rate among the native population, the French government implemented incentives in the form of public sector premium wage to attract skilled metropolitans in these new departments [46]. These premiums (also known as "high cost of living premium"), which still exist today, stands at 40% of the metropolitan salary in the Antilles and Guyane and at 53% in La Reunion. As a result, these territories have received a massive influx of metropolitans over that period, taking up public service positions. Given the poor local economic situation, the highly paid public sector has likely played a role in the level of inequality observed during that period.

1960s to 2000s: Sugar production in the Antilles begun to decay as from the mid-1960s and around the 80s in La Reunion. At the same time, in an attempt to remedy for the highly unequal land ownership inherited from the colonial rule, the government undertook various land reforms in these territories (except for Guyane), aiming to redistribute large landholdings among a greater number of planters. In La Reunion for instance, SAFER[47], put in place in 1966[48], redistributed 24000 hectares of land since its creation, representing 40% of the agricultural land in that period.

This period was also marked by an institutional effort to encourage migration towards the metropolis, in a bid to tackle the exploding demographic situation in the overseas departments and the need for labor in the metropolis. Put in place in 1963 until 1981, the BUMIDOM[49] played both a direct role[50], and an indirect role in the native population outflow to the mainland. Migrants from these territories were highly positively selected along the education line (Haddad, 2018).

The phase, starting in the early 1980s, marked an intensified effort of the government to tackle the persistent levels of inequality. To begin with, there has been the decentralization of power from the central government to the regional-level in 1982. This led to a gradual catch-up of the social benefits to the metropolitan-level. It took the form of the extension of the (until then restrictive) family allocations and minimum old-age pensions to a larger share of the population. This period has

[46] Initially granted only to metropolitans, and it was extended to natives in 1953

[47] Société d'Aménagement Foncier et d'Etablissement Rural

[48] Following the loi du 2 août 1961, SAFER's main function included buying land to resell in smaller sizes to planters

[49] Bureau pour le développement des migrations dans les départements d'outre-mer replaced by the Agence nationale pour l'insertion et la protection des travailleurs d'outre-mer (ANT) in 1981

[50] Around 85000 individuals in total migrated through this institution from the Antilles and La Reunion representing around 5% of their total population in that period

also seen an alignment of benefits to the metropolitan level- the Revenu Minimum d'Insertion (RMI) in 1989, unemployment insurance in 1991, family allocations in 1993, the alignment of the minimum wage in 1996 and the facility for youth employment in 1997 among others. This social benefit alignment process to the metropolitan level was more or less completed by the beginning of the 21st century. The existence of a minimum wage in these territories is due to their attachment to France.

A major part of the effort to reduce social and economic inequality in this period were achieved through redistributive policies[51]. Despite the importance of transfers, focusing on the pre-tax and transfer income allows us to grasp the precarious situation of the overseas population and the need to tackle the issue at its roots.

2000 to 2014: With the completion of the catch-up of public policies with the metropolis in the 2000s, there was less space for comparably compelling policies in the following decades. This is reflected in the relative stabilization in the evolution of inequality in all four overseas departments as from the 2000s. The top 1% income shares in the overseas departments have converged to the level of the metropolitan, while the top 10% shares remained consistently higher than that of the metropolis.

It is also worth noting that despite the common inequality trends observed in the overseas departments, La Réunion experiences a much higher level of inequality at the beginning of the period compared to the Antilles. This can perhaps be traced back to their different colonial past and persisting differences between them. For instance, the level of education, proxied by the illiteracy rate, in these territories from 1954 to 1967 gives an insight into the differences inherited from the colonial period (See Table B.5.1).

A large part of the explanation for the higher level of inequality in the overseas departments compared to metropolitan France can be attributed to the higher level of labor income inequality. In the post-departmentalization period, the economies of the overseas departments have undergone sharp transitions from agrarian-based economies to a service-sector dominated economy, as can be seen in Figure B.5.6. As a result, there has been a massive loss of unskilled jobs in the agricultural sector, accompanied by a growing demand for skilled labor in the tertiary sector. This led to a marked polarization of the local labor market. On the one hand, there was highly qualified and better-paid public servants than in the metropolis, and on the

[51]Since this paper focuses on fiscal income[52], I only observe the effect of levelling up of specific policies such as the minimum wage and family allowances which were paid as part of the labor income until 1986.

other hand, a large segment of precarious unemployed or low-income earners paid the minimum wage that is lower than in the metropolis.

As a case in point, figure B.5.3 to B.5.5 depict the wage density distribution in La Reunion in 1988. A large share of workers in the private sector was paid around the minimum wage, while the wage distribution in the public sector was far above the minimum wage. In fact, according to the French Statistical Institute, the ratio between the minimum wage and the minimum public servant wage was around 0.40 in the 1980s and has increased to only 0.50 in the 1990s, compared to 0.94 in the metropolis. Hence, while the alignment of the minimum wage to the metropolitan level has undeniably played a role in pushing upwards a segment of the population, there still exists a gap between the private and public sector wages in the overseas departments compared to France[53].

These post-colonial trends and evidences points towards some underlying divides that are very particular to assimilated ex-colonies. In the next section, I shed some light on two such aspects: first, the public-private sector polarization and second, a metropolitan-native population divide. I show how these explain part of the inequality patterns that is observed in these territories and discuss their implications in a post-colonial context.

2.6 Underlying Divides

Being attached to the metropolis has led to various specificities in the overseas departments. First, in the face of the declining employment due to the rapid decay of the agricultural sector, the French government devoted financial resources to expand the public sector in these territories. One such policy was the public wage premium to attract qualified labor from the metropolis. This privileged pay for public civil servants in the overseas department is still in place today[54]. The inability of the private sector to take off, combined with the absence of any local productive sector, led the public sector to account for a significant share of total employment and total income paid in these territories. Given the existence of the wage premium specific to the public sector, it became a highly-paid sector, in comparison to the private sector and self-employment.

[53]The ratio between the average annual wage of the private sector to the public sector in 2010 was 0,71 in La Reunion compared to 0,98 in metropolitan France

[54]Discussions challenging the need for the maintenance of the public wage premiums have been met with fierce opposition. Today, neither cost of living differences and even less the qualification differences between the metropolitan and native population can fully give ground to a wage premium to a small group of civil servants in the overseas departments.

Given the qualification requirements, public sector employment has for long penalized the native population who were mostly under-qualified or unqualified in favor of metropolitans. In fact, the share of metropolitan population in the overseas departments went from around 1% in 1954 to 10% in recent years. While the contemporary racial aspect of inequality in the overseas might largely be the logical result of the assimilation of these territories into the French Republic and the political will for territorial continuity[55], it has serious implications in a post-colonial setting given the widespread unfair sentiment of inequality and discrimination perceived by the native population.

This analysis relates to a large literature on the economics of discrimination, beginning with the work of Becker (2010). The economic literature on discrimination has most commonly studied wage differentials among different groups of individuals to study the gap in economic outcomes based on gender and race, for instance. It is interesting as a first analysis to observe the raw gap between two groups, before controlling for individual characteristics. Unequal opportunities themselves contribute to the resulting wage differential observed in the labor market, which is likely to be the case in the overseas departments.

Using administrative fiscal data for the year 2014 for a sample of the population matched with the population census[56], I analyze the distribution of income in the overseas departments along different lines, namely by origin and sector of employment among other factors. First, I present a descriptive analysis, and an estimation of the actual level of labor income inequality in the overseas department. I then investigate the existence of, and estimate, the "metropolitan premium" and finally, I estimate counterfactual scenarios, in absence of these premiums and the corresponding levels of inequality.

Native-metropolitan divide

Two main interesting stylized facts that come up in this graphical analysis are the origins dimension and the public-private sector aspect. Figure 2.9 depicts the share of metropolitans in total population in each decile of labor income. While metropolitans represent around 10% of the total population, their share by deciles increases significantly higher up the distribution. In fact, they represent between 25 to 35% in the top decile in the Antilles and up to around 50% in La Réunion and

[55]For instance, most public sector jobs are contested in a national competition and would, even if distributed randomly, lead to the assignment of more metropolitans to the positions in the overseas departments than the native population, due to their relative sizes in the total French population.

[56]Using the Echantillon Demographic Permanent (EDP)

Guyane. The massive concentration of metropolitans at the top of the distribution in La Réunion is striking. This tends to confirm the widespread sentiment of an over-representation of metropolitans at the top of the income distribution in the overseas departments.

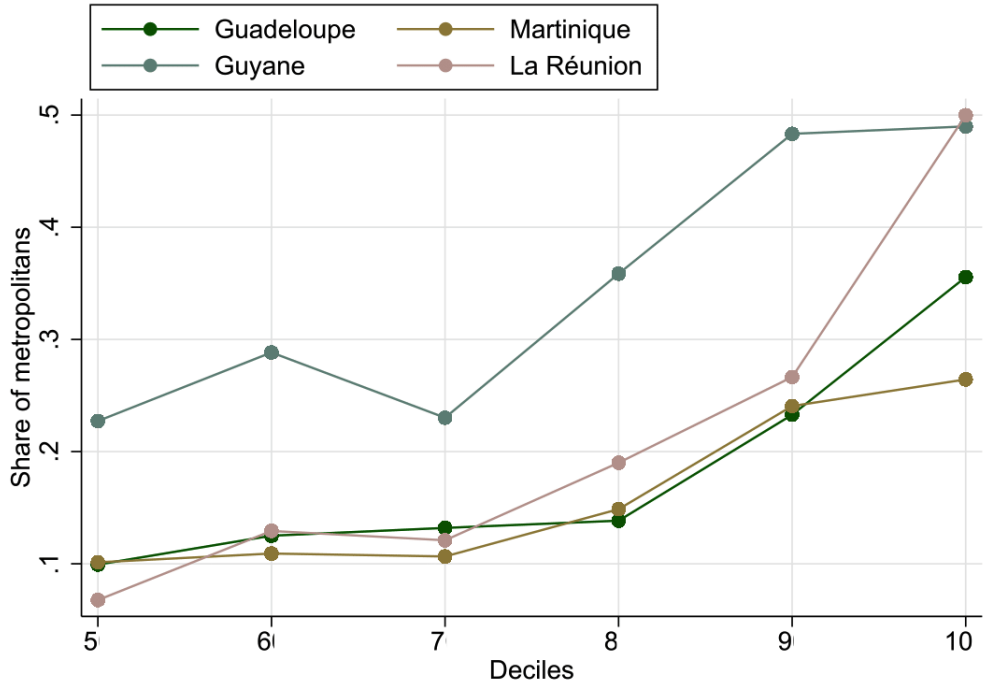


Figure 2.9: Share of metropolitans in each decile of labor income

Public-private divide

Knowing that metropolitans are more likely to occupy higher-ranks permanent positions in the public sector, it is interesting to have a look at the weight of the public sector in the overseas departments. Figure 2.10 shows the estimated share of different sectors of employment in each decile of labor income, restricting the sample to employed individuals. While in France, the private sector plays the major role in employment in each decile (stable at around 70% of total employment), the corresponding share in the overseas departments fluctuates across the different deciles. The public sector plays an increasingly important role higher up the distribution in all four departments, at around 50-60% for the top 20% labor income earners. The major role of the public sector in employment is an interesting peculiarity of the overseas departments as laid down in section 2.5. The data confirms that the difference between these territories and metropolitan France only exists in the public

sector, among those who are eligible for the premium[57].

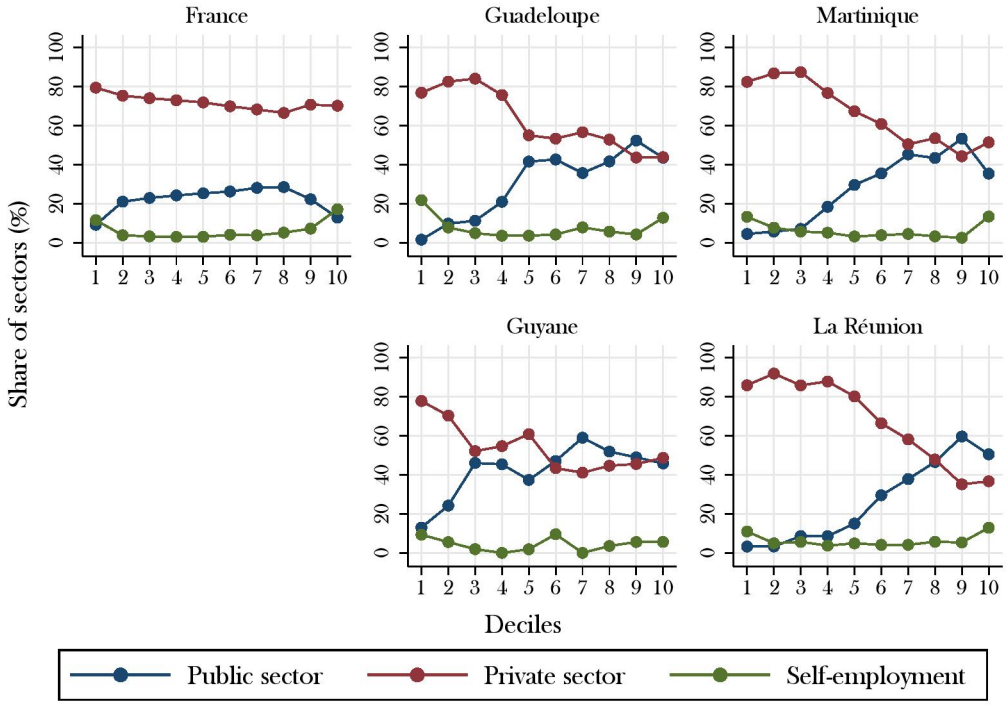


Figure 2.10: Share of different sectors in each decile of labor income

Regression Analysis

This section looks at the existence and extent of the metropolitan-native divide in the overseas departments. Table B.6.1 shows the descriptive statistics for adult[58] metropolitans and the native population. The metropolitan population is more educated, tend to be more active and employed, and earn a higher labor income than the native population on average. This is not very surprising given the very likely positive selection in the migration flow from the metropolis to the overseas departments. In the following analysis, I investigate this wage gap, controlling for observable characteristics. The model estimated is as follows:

$$Y_i = \alpha + \beta Metropolitan_i + \theta \text{Education}_i \ \gamma \text{Employment}_i + \delta Controls_i + \rho_d + \epsilon_i$$

Y refers to different definitions of annual labor income. *Metropolitan* is an indicator whether the individual is a native of a DOM or from the metropolis, based on their

[57] Public servants in permanent positions enjoy a wage premium in the overseas departments as explained in section 2.5

[58] The sample has been restricted to the population above 25 years to observe adult individuals who declare their income.

place of birth[59]. *Employment* refers to labor market characteristics of the individual. Depending on the specifications, these are dummies for being active, being employed, working full-time dummy, public sector employment, self-employment and type of contract (permanent or not). *Education* variable refers to the number of years of schooling and *Controls* include demographic characteristics such as age, gender and the matrimonial status (dummy for being married or not). All the specifications include the departmental fixed effects, ρ.

Different income definitions are used, and sample restrictions are applied depending on the specification at hand. Table B.6.2 in Appendix B.6 shows the regression results on a with a broader income definitions (salary only or including unemployment benefits or including retirement pensions). The dependent variable in the main analysis (table 2.2) refers to the annual salary, including income from self-employment and unemployment benefits of the working age population[60].

In order to grasp the full extent of the potential "discrimination" towards the native population, I estimate the above equation without any controls (Model 1), including demographic controls in Model 2 and controlling for number of years of education in Model 3. The first two columns of Table 2.2 shows that there is an important raw income gap between natives and metropolitans. The origin dummy in itself accounts for around 9% of the variations in annual labor income. Controlling for age, gender and matrimonial status, this "metropolitan premium" amounts to an average annual labor income of 12972 €. It decreases to 6764€ when controlling for the number of years of schooling, with an adjusted R^2 that more than doubles to 30%. This suggests that the level of education helps to explain a large share of the initial differences observed between the two groups of the population, since metropolitans are positively self-selected into the overseas departments.

Model 4 to 8 progressively includes labor market characteristics. β remains statistically significant even after controlling for all observable characteristics. The final model is able to explain 60% of the variations in annual labor income in the overseas departments and the "metropolitan premium", controlling for a set of observable characteristics, is around 5170€ per year. Overall, the results suggest that there exist an important gap in the overseas departments. As expected, there is a positive return to education. These results thus provides evidence of a native-metropolitan divide which might play a role in exacerbating an already tensed post-colonial society.

[59] Children born in metropolitan France of parents born in overseas departments are counted as metropolitans. Foreigners are excluded from this analysis, being a small minority in the population

[60] Between 25 and 65 years old.

Dependent Variable: Annual salary and self-employment earnings (inc. unemployment benefits)

	1	2	3	4	5	6	7	8
Metropolitan	14297.7***	12972.2***	6764.6***	7041.4***	6057.3***	5509.4***	5164.7***	5170.1***
	(448.8)	(445.7)	(417.5)	(390.1)	(345.5)	(316.2)	(314.2)	(313.6)
School Years			2093.3***	1764.3***	1274.7***	986.4***	947.7***	943.1***
			(38.84)	(37.11)	(33.87)	(31.52)	(31.50)	(31.45)
Active				14120.0***	2744.5***	2568.9***	2110.3***	1984.2***
				(332.0)	(351.8)	(321.7)	(322.9)	(322.9)
Employed					18098.2***	5485.5***	4170.4***	2471.7***
					(308.1)	(380.5)	(387.3)	(461.0)
Full-time						18338.6***	18719.2***	17272.9***
						(371.8)	(372.5)	(428.9)
Public Sector							2268.4***	2433.3***
							(282.8)	(283.4)
Self-employment							9407.9***	11979.5***
							(679.9)	(777.9)
Permanent								3363.0***
								(497.2)
Constant	13999.6***	22968.8***	-12141.6***	-27293.4***	-18151.1***	-14445.8***	-13590.5***	-13156.7***
	(340.6)	(921.5)	(1054.9)	(1048.1)	(940.2)	(863.1)	(860.7)	(861.5)
Observations	12438	12438	12438	12438	12438	12438	12438	12438
Adj R2	0.0854	0.132	0.296	0.386	0.519	0.598	0.605	0.606
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dept FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard Errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. The sample of 12 438 observations include all working-age population (between 25 - 65 years of age).

Table 2.2: Metropolitan Premium: Including and Excluding controls

Counterfactual Inequality Estimations

Given the existence of a "metropolitan premium" in the overseas department, it would be interesting to understand its role in the overall level of inequality observed. I first estimate the level of labor income inequality in each of these territories and France in 2014. Table 2.3 shows the actual top labor income shares. As expected from the results in the previous sections, La Réunion is the most unequal department in terms of labor income, with a top 10% labor income shares of 38% and a top 1% of 7%. We again observe that the top 10% income shares in the overseas departments are higher than the French level, while there does not seem to be much of a difference in the top 1% income shares (except perhaps for Guyane).

Actual Situation:	**Top 10%**	**Top 1%**
France	26%	6%
Guadeloupe	36.0%	7.2%
Martinique	34.9%	6.7%
Guyane	28.3%	5.0%
La Réunion	38.2%	7.2%

Table 2.3: Top labor income shares: Actual situation

In order to get a rough idea of the part played by the metropolitan-native divide, I estimate counterfactual levels of labor income inequality under two naïve scenarios. In the first scenario, I take the extreme setting of a total absence of metropolitans in the overseas departments. This would of course imply other consequences on the distribution of income, for instance if their positions were to be taken over by natives. For the sake of simplicity, I assume other things remain constant. Table 2.4 shows the result obtained from this simple exercise. In the "absence" of metropolitans, labor income inequality would have reduced in the Antilles and La Réunion, while slightly increasing in Guyane. The latter might be explained by the fact that metropolitans are present in all the deciles of income in Guyane and not only the top as seen in Figure 2.9. Thus, taking out all metropolitans from the Guyanese income distribution might lead to an increasing effect if there exist a larger disparity among natives. A back-of-the-envelope calculation would suggest that 4-12% of the difference in the actual level of inequality between the overseas departments[61] and France could be attributed to metropolitans.

Counterfactual I: No Metropolitans

	Top 10%	**Top 1%**
Guadeloupe	34.8%	5.9%
Martinique	34.5%	6.2%
Guyane	29.0%	4.9%
La Reunion	37.2%	6.3%

Table 2.4: Top labor Income Shares: Counterfactual I with no metropolitans

The second scenario consists of making the assumption that there is no income gap between natives and metropolitans. I impute a naïve counterfactual income for metropolitans based on the corresponding income of native with similar characteristics[62]. Panel B shows the resulting estimated level of top income shares. Under such a scenario, the top 10% labor income shares would reduce by around 2-3 percentage points in the overseas departments. Similar to before, a quick calculation suggests that 20 to 30% of the difference in labor income inequality between France and the overseas departments[63] could be attributed to the "metropolitan premium".

These naïve exercises allows to get a glimpse of the extent to which labor income inequality is driven by the presence of metropolitans and the metropolitan premium. These results have important implications in a post-colonial setting due to existing

[61] Excluding Guyane
[62] This exercise does not take into account the idiosyncratic features of the income variable.
[63] Excluding Guyane

Counterfactual II: No Metropolitan Premium

	Top 10%	Top 1%
Guadeloupe	33.0%	5.3%
Martinique	33.0%	6.0%
Guyane	25.5%	4.1%
La Reunion	34.7%	5.4%

Table 2.5: Top labor Income Shares: Counterfactual II with no metropolitan premium

and increasing racial tensions. For instance, it could take the form of quota for the native population in the public sector and raises questions about the fairness of the public sector premium paid to civil servants in the overseas departments, knowing that the high permanent positions are most likely to be occupied by metropolitans. These result from being attached to the metropolis.

2.7 Conclusion

This paper explores the post-colonial evolution of inequality in the four oldest colonies of France, which became part of the same country in 1946. These represent a peculiar post-colonial setting, as these territories have been assimilated to their ex-colonizers, rather than going through independence. Despite the widely acknowledged fact that these departments experience higher levels of inequality, no in-depth analysis has been devoted to this issue. In this paper, I fill this gap in the literature by estimating a consistent long-run series of income inequality at the regional level in France. Building a novel dataset based on fiscal data at the departmental-level, I estimate the income distribution in these territories since their departmentalization in 1946 until 2014. Results show that these territories have undergone various changes, leading to an initial increase in the top income shares until the 1960s, followed by a steep decline in inequality thereafter. While the top 1% income share in these territories has stabilized at the national level in the recent decades, the top 10% has remained consistently higher.

I then discuss some potential factors contributing to the level of inequality observed in these departments. The results provide suggestive evidences that the various policies put in place in the 20th century have been successful in reducing the extreme levels of inequality, though the gap between these departments and the metropolis has not completely disappeared. The difference in the level of inequality in the overseas departments compared to the metropolis might be explained by the larger

gap in wages in the public compared to the private sector in these territories. Civil servant wage premium coupled with low employment and low wages in the private sector distributed mostly around the minimum wage have led to a polarized labor market and thus labor income inequalities.

In the second part of the paper, I further investigate the labor income inequalities and, particularly, the metropolitan-native divide in the overseas departments. Using recent administrative fiscal data matched with the population census, I show the existence of a "metropolitan premium". The results suggest that metropolitans earn a higher income than the native population, even after controlling for observable characteristics. This adds a layer of complexity to the concern of the high level of inequality observed, given the social tensions stemming from their colonial past.

Bibliography

Alvaredo, Facundo, Lydia Assouad, and Thomas Piketty (2019). "Measuring lnequality in the Middle East 1990–2016: The World's Most Unequal Region?" In: *Review of Income and Wealth* 65.4, pp. 685–711.

Alvaredo, Facundo and Anthony B Atkinson (2010). "Colonial Rule, Apartheid and Natural Resources: Top Incomes in South Africa, 1903-2007". In:

Alvaredo, Facundo, Denis Cogneau, and Thomas Piketty (2021). "Income inequality under colonial rule evidence from French Algeria, Cameroon, Tunisia, and Vietnam and comparisons with British colonies 1920–1960". In: *Journal of Development Economics*, p. 102680.

Angeles, Luis (2007). "Income inequality and colonialism". In: *European Economic Review* 51.5, pp. 1155–1176.

Atkinson, AB (2010). "Top Incomes in a Rapidly Growing Economy". In: *Top Incomes*, p. 220.

Atkinson, Anthony B (2011). *Income distribution and taxation in Mauritius: A seventy-five year history of top incomes*.

– (2014). *The colonial legacy: Income inequality in former British African colonies*. Tech. rep. WIDER Working Paper.

Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez (2011). "Top incomes in the long run of history". In: *Journal of economic literature* 49.1, pp. 3–71.

Banerjee, Abhijit and Thomas Piketty (2005). "Top indian incomes, 1922–2000". In: *The World Bank Economic Review* 19.1, pp. 1–20.

Becker, Gary S (2010). *The economics of discrimination*. University of Chicago press.

Besson, D (2017). *L'impacte du spatial sur l'économie de la Guyane*. INSEE.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2017). "Generalized Pareto curves: theory and applications". In: *Review of Income and Wealth*.

Breton, Didier et al. (2009). "Les départements d'Outre-Mer face aux défis du vieillissement démographique et des migrations". In: *Population et sociétés* 460.

Drozin, G (2001). "Les politiques sociales en matiere d'action socialea la Réunion depuis 1946: un survol historique". In: *Communication aux Journées partenariales de l'action sociale, Institut Régional de Travail Social de la Réunion, Saint-Benoit.*

Engerman, Stanley L and Kenneth L Sokoloff (2002). *Factor endowments, inequality, and paths of development among new world economics.*

– (2005). *Colonialism, inequality, and long-run paths of development.*

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2018). "Income inequality in France, 1900–2014: evidence from distributional national accounts (DINA)". In: *Journal of Public Economics* 162, pp. 63–77.

Haddad, Marine (2018). "L'effet d'une politique d'État sur les migrations DOM-métropole. Les enseignements des recensements de 1962 à 1999". In: *Population* 73.2, pp. 191–224.

Maurin, L and M Bernier (2013). *Des départements d'Outre-mer marqués par les difficultés sociales et les inégalités, Compas Etudes n 9, octobre 2013; 8p.*

Michel, Christophe, Maël Theulière, and Nathalie Missègue (2010). "Les Inégalités de Revenus entre les DOM et la Métropole". In:

Nunn, Nathan (2008). "Slavery, inequality, and economic development in the americas". In: *Institutions and economic performance* 15, pp. 148–180.

Piketty, Thomas (2001). *Les hauts revenus en France au XXème siècle.* Grasset.

Williamson, Jeffrey G (2010). "Five centuries of Latin American income inequality". In: *Revista de Historia Economica-Journal of Iberian and Latin American Economic History* 28.2, pp. 227–252.

– (2015). "Latin American Inequality: Colonial Origins, Commodity Booms or a Missed Twentieth-Century Leveling?" In: *Journal of Human Development and Capabilities* 16.3, pp. 324–341.

Chapter 3

Land Inequality in the Developing World[1]

Abstract

Agricultural land is vital for three out of every four of the poorest billion individuals in the world, yet little is known about its distribution. Existing cross-country estimates of land inequality, based on agricultural census data, measure the size distribution of agricultural holdings. In practice, these neither reflect land ownership inequality nor value inequality and often do not account for the landless population. In this paper, we tackle these issues and provide novel and consistent estimates of land inequality across countries, based on household surveys. We show that (i) inequality in land value can differ significantly from inequality in land area, (ii) the proportion of landless households across countries varies substantially, markedly affecting estimates of inequality, and (iii) regional patterns in inequality according to our benchmark metric (land value inequality including the landless) contradict existing estimates from agricultural censuses. Overall, South Asia and Latin America exhibit the highest levels of inequality, with the top 10% of landowners capturing up to 75% of agricultural land, followed by Africa and "Communist" Asia (China and Vietnam) at levels of around 55–60%.

[1] Joint-work with Luis Bauluz, Filip Novokmet & Daniel Ordoñez Sanchez

Agricultural land is vital for three out of every four of the poorest billion individuals in the world, who depend on land and related activities for their subsistence (FAO, 2016). Over the past three decades, developing countries have gone through a profound economic transformation as they embark on a catch-up process with the advanced economies (Bourguignon, 2017). However, this process of convergence has been very unequal, with only a few countries (foremost China) undergoing a significant process of industrialization (Rodrik, 2016), while vast parts of the developing world (notably South Asia and Africa) proceed at a much slower pace, and follow a more fragile path (Lakner and Milanovic, 2015). As a result, the number of agricultural workers worldwide today is largely the same as it was 30 years ago (Figure 3.1). Despite the importance of land for the world's poor people, we know almost nothing about its value and distribution, since the existing estimates reflect neither inequality of landownership nor inequality of value and do not account for the landless population. Consequently, both policy-makers and academic researchers lack basic information to evaluate the economic conditions and the ownership structure which dictate the lives of the world's poorest people.

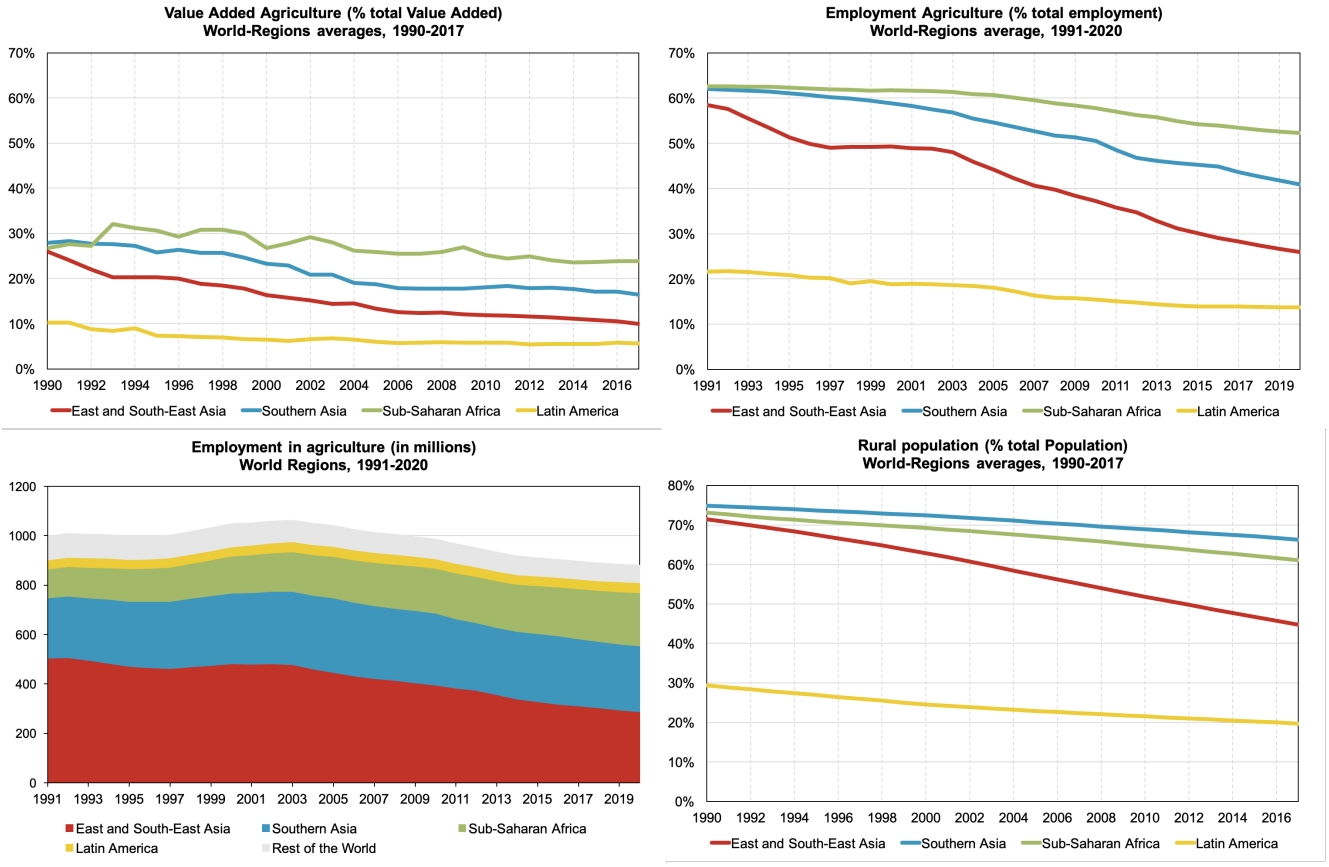
Precisely estimating land inequality is crucial, given its relevance to debates that range from institutions and human capital accumulation to food security and poverty alleviation. Existing cross-country estimates of land inequality, and the ensuing literature which analyses its effect on economic outcomes, are based on the distribution of the size of operational holdings, constructed using agricultural census data (Taylor and Jodice, 1983[2]; Deininger and Squire, 1998[3]; Frankema, 2010[4]). Operational holdings are defined as economic units of agricultural production under single *management* (FAO, 2018). Hence, these estimates pose serious conceptual challenges for measuring inequality in landownership since they do not capture the *ownership* of land holdings nor do they account for differences in the *value* of land (e.g. due to soil quality or location). It is thus unclear whether the distribution of the area of holdings from agricultural census data reliably captures overall land inequality. There is a need to assess the validity of this link and to better delineate the concept of land inequality, which is most pertinent in the context of developing countries.

The contribution of this paper is to provide consistent estimates of inequality in landownership across countries and regions of the world, both in terms of area and

[2]For 54 countries in the 1960s, based on FAO World Census Agriculture

[3]261 observations for 103 countries, based on FAO World Census Agriculture

[4]186 observations for 105 countries, based on census data from the International Institute of Agriculture (IIA) and FAO



Note: The upper left panel shows trends in the share of agricultural value added of total value added. The upper right panel depicts the share of agricultural employment of total employment. The bottom left panel corresponds to employment in agriculture in absolute numbers, and the bottom right panel shows trends in the share of the rural population of total population. These estimates are regional-weighted averages by country population sizes. Source: ILO and FAOStat.

Figure 3.1: Global trends in agriculture since the 1990s

value, as well as accounting for the landless population. Departing from the use of agricultural censuses, we exploit household survey data, which allows us to focus on land privately owned by a household rather than solely its holdings, the former being more appropriate when analyzing landownership inequality. Furthermore, while land area inequality provides an idea of the distribution of land, accounting for the differential value of land owned by households might give a different picture. To the best of our knowledge, this is the first paper to present and explore the relationship between inequality in land area and inequality in land value. Finally, since agricultural censuses do not capture landless individuals, this part of the population has been substantially disregarded in the literature and at best only roughly proxied. This is a significant shortcoming, as variations in ownership rates across countries are unaccounted for. Moreover, landless families are precisely the most vulnerable as they are at the bottom of the distribution but still rely on working and using land. In this paper, we provide and compare estimates of land inequality including and excluding the landless population.

This paper thus makes an important contribution in showing the need for more critical use of existing estimates based on census data. It is also the first to provide comparable estimates of land inequality, under different definitions, in various countries across the world. We provide a novel perspective on international patterns of land inequality. Our benchmark metric of agricultural land inequality (i.e. inequality of land value when including the landless population) reveals regional patterns that show South Asia and Latin America to be the most unequal regions of the world, followed by relatively more egalitarian African countries, and finally "Communist" Asia (China and Vietnam) as the least unequal world region.

Related literature.

This paper is related to several strands of the literature. The first strand looks at land inequality and poverty. The livelihood of the poor in developing countries critically depends on agricultural land (e.g. Deininger, 2003, Banerjee and Duflo, 2007, etc.). Correspondingly, access to land has often been seen as an important policy tool for relieving poverty, underpinning the demands for the land redistribution reforms (Banerjee, 1999; Besley and Burgess, 2000; Griffin, Khan, and Ickowitz, 2002; Lipton, 2009). At the same time, sustained growth of agricultural productivity is a critical source of economic development (and poverty alleviation in the long-run; e.g. Eswaran and Kotwal, 2006, Gollin, Parente, and Rogerson, 2007, Restuccia, D. Yang, and Zhu, 2008, D. Yang and Zhu, 2013 etc), bringing up concerns about the effects of land redistribution on production efficiency and agricultural output. For example, the nature of the relationship between farm size and productivity is at the heart

of the debate on desirability of land redistribution from the efficiency perspective (Sen, 1966; Berry and Cline, 1979; Vollrath, 2007; Foster and Rosenzweig, 2017, etc.). Given the central importance of these questions for poverty and development, there is a need for a consistent international database on agricultural production and distribution at the household level. By considering value and landless, our new dataset will be most relevant when studying land inequality's link with poverty.

Closely related strand in the literature looks at the relationship between land inequality and economic development. The development literature has often emphasized the positive relationship between the initial equitable distribution of land and subsequent growth rates (Rodrik, 1995, Deininger and Squire, 1998, Kanbur, 2000, etc.). In contrast, unequal distribution of land adversely affects development, as it results in institutions that preserve the distributive status quo (Engerman and Sokoloff, 1997; Sokoloff and Engerman, 2000). Land concentration hampers investment in education, as this goes against the interest of landed elites (Galor, Moav, and Vollrath, 2009). It can also affect and be reinforced by poor financial development (Binswanger and Deininger, 1999). Land concentration restricts the poor's access to credit and hence their access to land markets. The resulting poor institutions, inadequate investment in education and provision of public goods, underdevelopment of the financial market, and the increased incidence of conflicts are some of the channels through which land inequality affects growth and development (Deininger and Squire, 1998; Easterly, 2007; Midlarsky, 1988; Guereña and Wegerif, 2019).

The paper is organized as follows: In the next section, we discuss different concepts related to land inequality. We then describe the data and methodology used in this paper in section 3, followed by the main results in section 4. We first examine the link between the area of inequality in holdings from censuses and our estimate of inequality in landownership area from household surveys. We then provide estimates of both land area and value among the landowning class. Our results show that land value inequality differs significantly from land area inequality, and confirms the need to take both into account. Our results further show that accounting for the landless population increases land inequality unequally between countries due to differences in ownership rates.

3.1 Agricultural censuses vs. household surveys

The literature on land distribution has long relied on estimates of Gini coefficients for land using agricultural censuses, which provide tabulated data on the number of holdings and the total area of holding by size class (Deininger and Squire, 1998;

Frankema, 2010). However, in this paper, we turn to household surveys instead. In what follows, we discuss what the measurement of land distribution captures when using the two data sources and why we prioritize surveys in this study.

First, land distribution calculated using an agricultural census captures the distribution of operational holdings, which are "an economic unit of agricultural production under single management comprising all livestock kept, and all land used wholly or partly for agricultural production purposes, without regard to title, legal form or size" (FAO, 2018). Hence, censuses do not directly capture the distribution of landownership, since landholders may or may not be owners of the operational holding. Moreover, the census does not necessarily account for multiple landholdings per owner[5], thus failing to capture the full extent of land concentration. On the other hand, household surveys often include an agricultural module, which collects detailed information on land at the household level. The advantages of this source of data are numerous. Surveys provide a better idea of inequality in landownership since each plot of land is linked to the household owning it, unlike census data. Surveys also allow us to distinguish between privately-owned land and operated land, i.e. land that is merely utilized by the household, for instance through renting or sharecropping.

The literature relying on censuses has focused on the distribution of operated land. While this gives an idea of the extent of access to land in terms of utilization, it is not equivalent to landownership. In fact, households which operate land that they do not own will need to compensate the landowner for its use through rent payments or sharecropping. Moreover, land that is owned can also be used as collateral to gain access to credit, or can be rented out or sold in case of a need for liquidity – hence the need to distinguish between land that is merely operated and land that is effectively owned (Doss et al., 2015).

Second, census data typically cover not only household farms but also government lands, and holdings by private corporations.[6] Communal land is not considered an operational holding and is, therefore, not captured by census statistics. The

[5] "The holding's land may consist of one or more parcels, located in one or more separate areas or in one or more territorial or administrative divisions, providing that they all share such means of production as labor, farm buildings, machinery or draught animals. Several economic agricultural production units under the same ownership, or under the same general management, may be considered as separate holdings if they are operated by different persons." (FAO, 1999). As explained by Vollrath, 2007, p. 204, the distribution of operational holdings does not capture the distribution of landownership. The distribution of land holdings is relevant if "we are interested in efficiency, not equity".

[6] The sector coverage, however, varies across countries and over time. For example, most African countries only cover land operated by the household sector (excluding corporate land).

ample coverage of farms, regardless of the landholder being a household or not, is an advantage of the census to analyze the productive structure of the agricultural sector. It is less clear how to interpret statistics of land inequality which mix household holdings with public and corporate ones (note, also, that corporations can be owned by numerous households and by both the government and foreigners).

Household surveys, by contrast, do only capture land owned by households. Like censuses, surveys do not include communal land. Another caveat of the survey is that it misses land that is not owned directly by households but indirectly through businesses, potentially underestimating the actual level of inequality (since equity ownership is generally concentrates at the top of the distribution). Lowder, Skoet, and Raney (2016) estimate that family farms operate around 75% of total agricultural land, indicating that households are responsible for most of the world's agricultural and food production. To the extent that land operated and land owned are tightly connected, this would suggest a large role for households in agricultural land. In the results section, we show that for the sub-sample of countries included in our study for which this information is available, this figure is above 80 percent in most cases.

Third, differences in the value and quality of land are not measured in agricultural censuses. Unlike agricultural censuses, household surveys often provide information on the area (e.g., GPS measurements and farmers' estimates) as well as the market value of land at the household level. Both are valuable information, as the distribution of land in terms of area might not be equivalent to the distribution in terms of value. For instance, if larger landowners have disproportionately more valuable land, then land area inequality would not reflect the full extent of the unequal distribution. This paper bridges the gap in the literature when it comes to land value inequality and provides consistent estimates across countries.

Additionally, census data, by definition, do not account for landless households. This omission is crucial as the share of the landless population vary in different countries and thus existing estimates may not be fully comparable across countries. For instance, based on estimates of inequality between landowning households, a country where land is equally distributed among only a handful of landowners will have a lower level of inequality compared to another country that has a more disparate distribution of landownership among a larger share of landowning households. There is thus a need to include landless households to account for the full picture. In fact, Erickson and Vollrath (2004) show that the established effect of land inequality on institutions and financial development is sensitive to the inclusion of the landless population.

Erickson and Vollrath (2004) propose a complementary measure of inequality, which is the ratio between agricultural population and the number of holdings; this aims to capture the extent to which holdings are spread across the relevant population, using FAO data. However, the implicit assumption behind such a proxy for landless households is that each agricultural holding has a single owner. Although it is an improvement vis-à-vis the existing literature, this raises concerns similar to those around the existing literature on land inequality.

Finally, as argued by Lowder, Skoet, and Raney (2016), the coverage and methodology of agricultural censuses are not uniform between countries and over time, especially in developing countries, despite efforts by FAO to encourage uniformity. Agricultural censuses in different countries do not systematically distinguish between different forms of legal ownership and can also have different minimum thresholds for recording holdings, which further reduces comparability. Household surveys, on the other hand, provide the flexibility required to make them more comparable across countries and over time. Some papers in the literature have turned to household surveys to assess land distribution in different countries (see Doss et al. (2015) for a review on gendered land outcome in Africa based on surveys).

These factors suggest that agricultural census data do not allow to grasp the full extent of land inequality. Despite the caveats of survey data, we believe that surveys remain a substantially better source when estimating inequality of landownership. Surveys are mostly nationally representative and hence effectively designed to capture all types of households, whether landowning or landless. They provide detailed data on the land owned by a household, which allows for an in-depth analysis of inequality in landownership in terms of both area and value, while accounting for the landless population. To the best of our knowledge, this paper represents the first attempt to provide a comprehensive estimation of the distribution of landownership inequality by both area and value that is comparable across countries, spanning different continents, and exploiting household surveys.

3.2 Data & Methodology

In this paper, we start by revisiting and updating estimates of land area inequality based on agricultural census data. This data source is centralized and overseen by the UN Food and Agriculture Organization (FAO) and is published at the country level every decade under the World Programme for the Census of Agriculture (WCA). Agricultural census reports provide a tabulated distribution of operational holdings by

size bracket.[7] Previous estimates of land distribution based on this source cover most of the twentieth century, but with only a few estimates in the early 2000s (Deininger and Squire, 1998; Frankema, 2010). In this paper, we reassess and update estimates of land inequality based on census data, up to the most recently available year (which generally correspond with the 2010 decennial round). Given the tabulated format of the data, we use the generalized Pareto interpolation method (Blanchet, Fournier, and Piketty, 2017) to update census-based estimates of inequality. The goal is to examine census-based estimates relative to survey-based ones.

As explained previously, we then use household surveys to provide estimates of land area and value distribution, as well as including the landless population in different countries across the world. Two main types of surveys are used in this paper to do so: the World Bank's Living Standards Measurement Study (LSMS) household surveys and official household surveys conducted by different countries.[8] These generally consist of an agricultural module which collects information about the fields or plots owned by the household. The relevant information for estimating landownership inequality is the land area, reported value and an indication of ownership.

The choice of countries in this paper is based on the availability of household surveys capturing the ownership of land (Table C.1.1). In some countries the quality of the data was not sufficiently good, and they were therefore excluded from the analysis. Most surveys available in the different countries have a very short temporal dimension (e.g. in various cases only one year of data is available). For this reason, we restrict our analysis to a single observation per country and do not analyze trends in the concentration of agricultural land.

Our object of analysis is to measure the distribution of landownership. In this paper, landownership is defined as any agricultural land over which a household has private property rights. This is defined fairly consistently across countries. China and Vietnam are special cases, as in these countries private property is less clearly defined, but rural households are given extensive rights over land e.g. rights to control, profit from, and inherit land (McKinley and Griffin, 1993; Li and Zhao, 2007; Do and Iyer, 2003; Piketty, L. Yang, and Zucman, 2019). For further discussion on this, please refer to Appendix D.

[7] Some countries further provide decomposition by tenure, gender, land use and crops.

[8] In addition, we use the Demographic and Health Surveys (DHS) of the DHS Program for robustness purposes. DHS provides two key pieces of information regarding land, a direct question on ownership and the area of land owned, but it makes no estimation of value. We thus estimate landownership inequality and obtain estimates close to our main estimates, using the LSMS and other household surveys (see Figure C1).

As in previous studies that rely on census data, household surveys unfortunately also do not account for access to communal land. Communal land is often large plots of land that can be used equally by the members of the community to which it belongs. Members typically do not have exclusive or formal rights on the plot of land: they normally have use rights and sometimes common control rights, and very rarely transfer rights. Beyond the conceptual challenges of landownership when it comes to communal land, developing countries, especially in Africa, also face a more generalized issue of lack of land registration. This poses challenges in terms of securing land rights, and access to credit. Deininger and Feder (2009) document the impact of land administration interventions and conclude that land registration is beneficial to ensure a secure land tenure system and for the development of a formal land market, which is associated with higher investment, productivity, and overall efficiency. In addition, the formalization of these land provides better data that is crucially needed to analyze the importance of communal land in different countries[9].

In this paper, we focus on two ways of measuring the agricultural land owned by a household. The first is in terms of area of agricultural land (i.e. the size of the land holdings owned by a household).[10] The second is in terms of the value of agricultural land. The latter is our preferred measure, since it accounts for the large heterogeneity of land types within a country and captures the value of land as an asset. Values reported by surveys are based on the concept of current market value, where agricultural land is valued at prevailing market prices.[11]

To describe the distribution of agricultural land, we use standard measures of inequality such as the Gini coefficient and land shares (i.e. the percentage of land owned by a population group such as the top 10%, middle 40%, or bottom 50%). Although the Gini coefficient has been used predominantly in land inequality studies based on census data, we prefer to use land shares. The Gini index is a synthetic measure of inequality which summarizes the entire distribution into a single number, and it is thus less informative about where the important changes in distribution

[9]The success of Mexico in its registration of communally-owned land provides a case study of the possibility of formalizing these lands and the potential gains in doing so. It also entails the availability of better data to study communal land (Morett-Sánchez and Cosıo-Ruiz, 2017).

[10]Note that agricultural land area is reported in both agricultural censuses and in surveys. The difference is that surveys measure landownership at the household level, while agricultural censuses measure the land area of operational holdings.

[11]The valuation practice in surveys is generally based on a subjective assessment of respondents (surveys generally ask a question along the lines of: "What would be the amount received if the land was sold today?"), but this is often complemented by external assessments based on administrative data. In certain instances, in particular in the absence of well functioning agricultural land markets, the survey design evaluates the market value of land using alternative approaches, such as by capitalizing agricultural income (for example, this is an approach adopted by the China Family Panel Studies (CFPS)).

take place.

We measure inequality of landownership within two population groups. The first group consists of landowners (i.e. those households owning land). Our second group consists of landowners plus landless households. The latter is our benchmark unit, since it is important to account for landless households to obtain a complete picture of land inequality. Surveys are extremely useful in examining the landless population, since they capture both the population of households living in rural areas and the working activities of each member of a household, including agriculture. This information, together with the number of households who are landowners, allows us to identify the population of 'landless households'. We define landless households as those where at least one of its members is employed in agriculture but does not report owning any agricultural land.

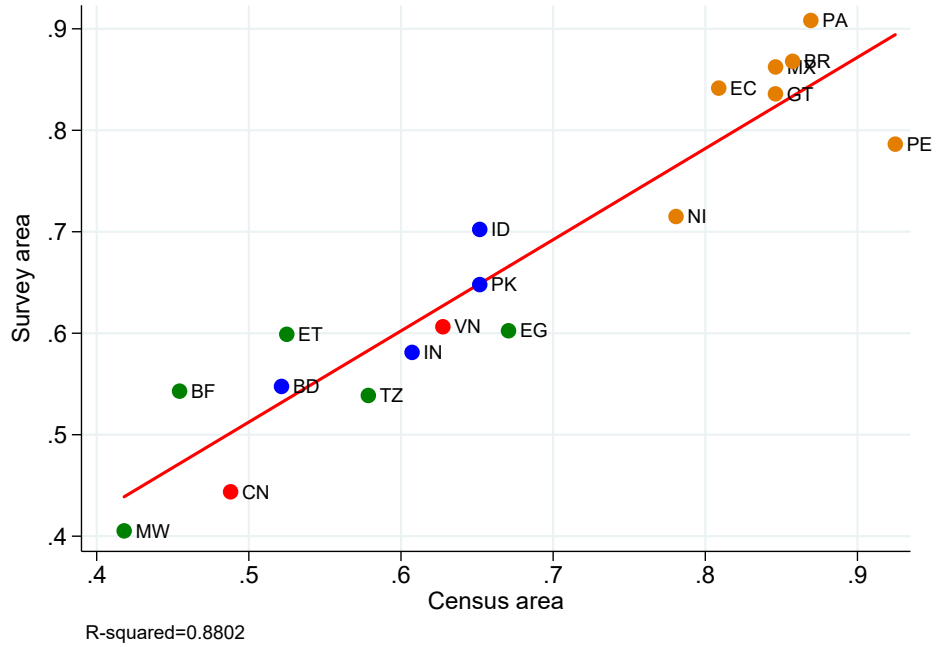
3.3 Results

3.3.1 Land area inequality: census vs. survey

Previous work on the measurement of the distribution of landownership across countries has been based on information contained in agricultural censuses. However, as argued above, censuses do not record agricultural ownership units but rather operational (or production) units. The implicit assumption behind this is that the size distribution of operational holdings provided in censuses serves as a proxy for the distribution of landownership. Moreover, the use of census data has restricted analysis to inequality of land area and not of land value, as well as to inequality between landowners, excluding landless households.

Given the wide use of census-based estimates in the literature, as a first step, it is useful to examine the extent to which the size distribution of farms reflects the distribution of land area ownership. Figure 3.2 compares agricultural land inequality estimated from survey and census data. More precisely, it shows the Gini index for the distribution of land area: (i) among households owning land from surveys (x-axis), and (ii) among land holdings from agricultural censuses (y-axis). In order to ensure comparability, we select rounds of survey data that are the closest to the census year of the countries included.

Interestingly, the Gini index is broadly comparable according to the different definitions in the two data sources (the regression line is almost equivalent to the 45-degree line). We find, according to both sources, that land inequality is highest in Latin America, at an intermediate level in Asia, and lowest in Africa.



Note: This figure includes Mexico, Peru, Egypt, and Burkina Faso, for which we have land area estimates from surveys but no information on value. They are hence not part of the sections that follow. Conversely, the Gambia, Nigeria, and Niger do not appear in this figure as there is no census information on the distribution of holdings. In order to ensure comparability, we select rounds of survey data that are the closest to the census year of the respective country.
The countries and the year of survey are as follows: BD – Bangladesh (2011); BF – Burkina Faso (2014); BR – Brazil (1996); CN – China (2002); EC – Ecuador (2014); EG – Egypt (2018); ET – Ethiopia (2011); GT – Guatemala (2000); ID – Indonesia (2014); IN – India (2012); MW – Malawi (2010); MX – Mexico (2009); NI – Nicaragua (2014); PA – Panama (2008); PE – Peru (2007); PK – Pakistan (2010); TZ – Tanzania (2018); VN – Vietnam (2014). For sources of data, see Appendix A.

Figure 3.2: Gini index based on census and survey data in selected countries

Given that the two estimates of land inequality tend to coincide, this suggests that inequality of landholding area can be an appropriate proxy for inequality of land area ownership. The high correlation between the two sources should, therefore, reflect the following three points: (i) households own the vast majority of the existing agricultural land; (ii) most of the landholders are also landowners; (iii) landowners do not own a significant amount of land through multiple operational holdings.

Making use of the information available in the latest two rounds of FAO censuses (2000 and 2010), we can provide evidence supporting the first two points (unfortunately, we could not find data on the third point). First, we can observe the total amount of land in operational holdings that is owned by landholders. This information is available for many countries, both in rich and developing regions (table appendix E1).

Interestingly, we find that on average, over 80% of the operated land in Asia, Latin America and Africa is also owned by the landholder (this is in contrast with Europe, North America and Oceania, where this figure is between 50 and 65%). Second, in a smaller set of countries, censuses also report the legal status of the landholder, distinguishing between civil and juridical landholders (the first status largely captures households[12]). Within the sample of countries included in this study, in eight cases this information was available (figure appendix E1). In seven countries, the amount of land operated by households is over 80 percent. The only exception is Peru, where the share is closer to 40 percent. This could potentially explain why results for Peru are slightly more distant from the regression line than other countries in Figure 3.2.

However, the various caveats associated with census data, such as inconsistencies in terms of coverage (household, corporate or government sector included or not, in an unsystematic way) should be kept in mind. Additionally, while census data could be seen as a first approximation for inequality in land area, it does not reflect inequality in land value. The following sections expand on this by including different dimensions of inequality to arrive at our benchmark concept of inequality, which is that of the distribution of agricultural land value among agricultural households (including both landowners and landless households).

[12]The classification of the legal status of the landholder varies across countries. The most common categories are civil vs. juridical, but some countries use different legal classifications. For example, in Brazil, the legal status is decomposed into seven categories: (i) Individual, (ii) Condominium, consortium or partnership, (iii) Cooperative, (iv) Limited liability company, (v) Institution, (vi) Government, and (vii) Others.

3.3.2 Distribution of land area vs land values

The value of one hectare of agricultural land can vary widely within a country, with numerous factors explaining the differences: diversity in soil quality (Benjamin, 1995), type of agricultural cultivation (e.g. cropland vs pastures), access to irrigation and agricultural capital, area of the agricultural holding (Barrett, 1996; Martinelli, 2016), land markets regulation (Restuccia and Santaeulalia-Llopis, 2017), factor market imperfections (Sen, 1966), etc. It is clear that agricultural land is not a homogeneous asset, and that estimates of land-area inequality fail to capture the diversity of values across land holdings.

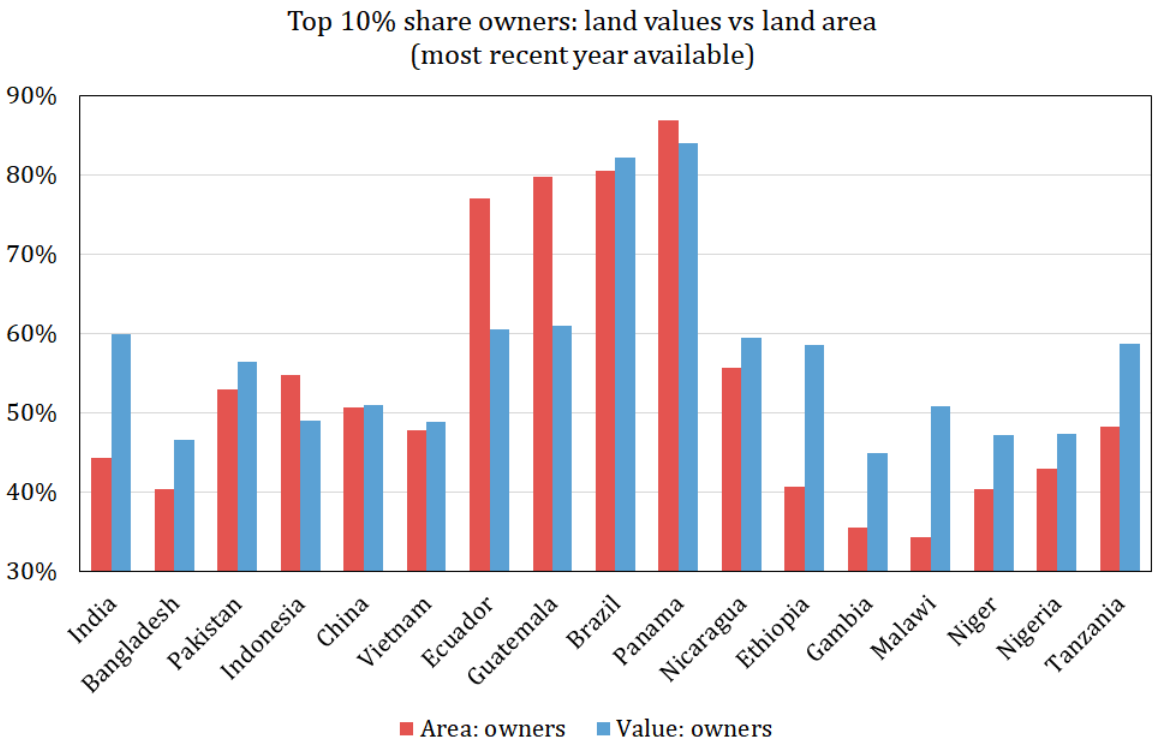
We go one step further than previous studies and compare inequality of land area with that of land values in household surveys, whenever this information is available. Figure 3.3 shows the distribution of agricultural land within landowning groups for the most recent year for which data are available, using two measures of agricultural land: land area (red bars) and land values (blue bars). In other words, the figure shows the share of total agricultural land owned by the top 10% of landowners, according to the two measures.

Importantly, the results indicate that inequality of land value can be significantly different from inequality of land area. For example, a comparison of India and Ethiopia with Ecuador and Guatemala is particularly informative. The first two countries have relatively low levels of land area concentration when compared with the second two: in fact, the share of the top 10% of landowners in Ecuador and Guatemala is twice that of the same groups in India and Ethiopia.[13] From this perspective, inequality in these two pairs of countries is remarkably different. Based solely on land area estimates, Ethiopia and India would be assessed as being relatively egalitarian countries compared with Ecuador and Guatemala, which are extremely unequal by any standard. However, results for land value inequality, as opposed to land area inequality, completely change this comparison. Under the land value metric, differences between the four countries all but disappear, as the top 10% of landowners own around 60% of total agricultural land value in all four countries.

Generally, our results point to important differences between land value inequality and land area inequality. In Appendix Figure C.5.2, we repeat the comparison between the Gini index in the census and the survey, but this time between land area in census and land value in surveys. While we still find a positive correlation between the two measures, it is much lower than in Figure 3.2. Concretely, the R-squared

[13] More precisely, 80% of total agricultural land is owned by the top 10% of landowners in Ecuador and Guatemala, compared with 40% in India and Ethiopia.

drops from 0.88 to 0.55, confirming that the distribution of land area from censuses only mildly predicts the distribution of land values.[14]



Note: This graph provides estimates of the top 10% share of area and value among the owning class, both from the urban and rural area.

Figure 3.3: Share of top 10% of landowners in selected countries, by land area and land value

3.3.3 Accounting for the landless population

As explained in previous sections, a meaningful measurement of the distribution of agricultural land should not be restricted solely to landowners. While inequality within the landowning class provides useful insights into the structure of inequality, a comprehensive assessment of the phenomenon needs to include the landless population (i.e. those directly involved in agriculture but who do not own land).

Figures 3.4a and 3.4b show the land shares of the top 10% and bottom 50%, respectively, for land value among landowners (blue bars) and landowners and landless households (green bars). In addition, Table 3.1 shows the share of landless households within the population of landowners plus landless households. Figures 3.4a and 3.4b show that including landless households is important in establishing

[14]The sample of countries in Figure 3.2 and appendix figure C.5.2 is not exactly the same. In appendix Figure C.5.3, we replicate Figure 3.2 using only the countries that are also included in Figure appendix C.5.2. In this case, the drop in the R-squared is even larger: from 0.92 to 0.55.

levels of inequality. More specifically, regions with the highest shares of landless households show larger increases in levels of inequality.

Three patterns are worth mentioning. First, South Asia (India, Bangladesh, Pakistan and Indonesia) and Latin America (Nicaragua, Ecuador, Guatemala, Brazil and Panama) are the most unequal regions, with the share of the top 10% rising to up to 70-80%, and that of the bottom 50% falling from 7–10% to 0–2%. In both regions, landless households account for more than one-third of the reference population.

In contrast, inequality in China and Vietnam is not significantly affected by the inclusion of the landless population, with an increase of a few percentage points in the share of the top 10% (and a decrease in the share of the bottom 50%). This is driven by the very low proportion of landless households, around 3–12%. This is explained by historical land reforms carried out in these countries under their communist regimes, which still provide widespread access today to agricultural land for most households in rural areas.

Finally, African countries have proportions of landless households that are somewhere in between. Hence, changes in levels of inequality when switching from one population concept to the other are in the middle compared with the two groups of countries as well.

Landless households as a percentage of agricultural households

World Regions	**Individual Countries**					
South Asia	India	Bangladesh	Pakistan	Indonesia		
36%	39%	40%	36%	28%		
China-Vietnam	China	Vietnam				
7%	3%	12%				
Sub-Saharan Africa	Ethiopia	Gambia	Malawi	Niger	Nigeria	Tanzania
29%	40%	33%	25%	27%	27%	21%
Latin America	Ecuador	Guatemala	Brazil	Nicaragua	Panama	
35%	36%	56%	29%	40%	12%	

Note: This table provides the proportion of landless household out of the landowning and landless households. The household is defined as landless i) if it does not own any piece of land and ii) if at least one household member participates in the agriculturally-related activities. We include Brazil and Peru in this table since we observe the percentage of landless households, despite surveys not covering the value of land. Hence, they are not included in subsequent analyses.

Table 3.1: Proportion of landless households

Overall, it is clear that any assessment of land inequality that excludes the landless population will result in an incomplete understanding of the complex structure of

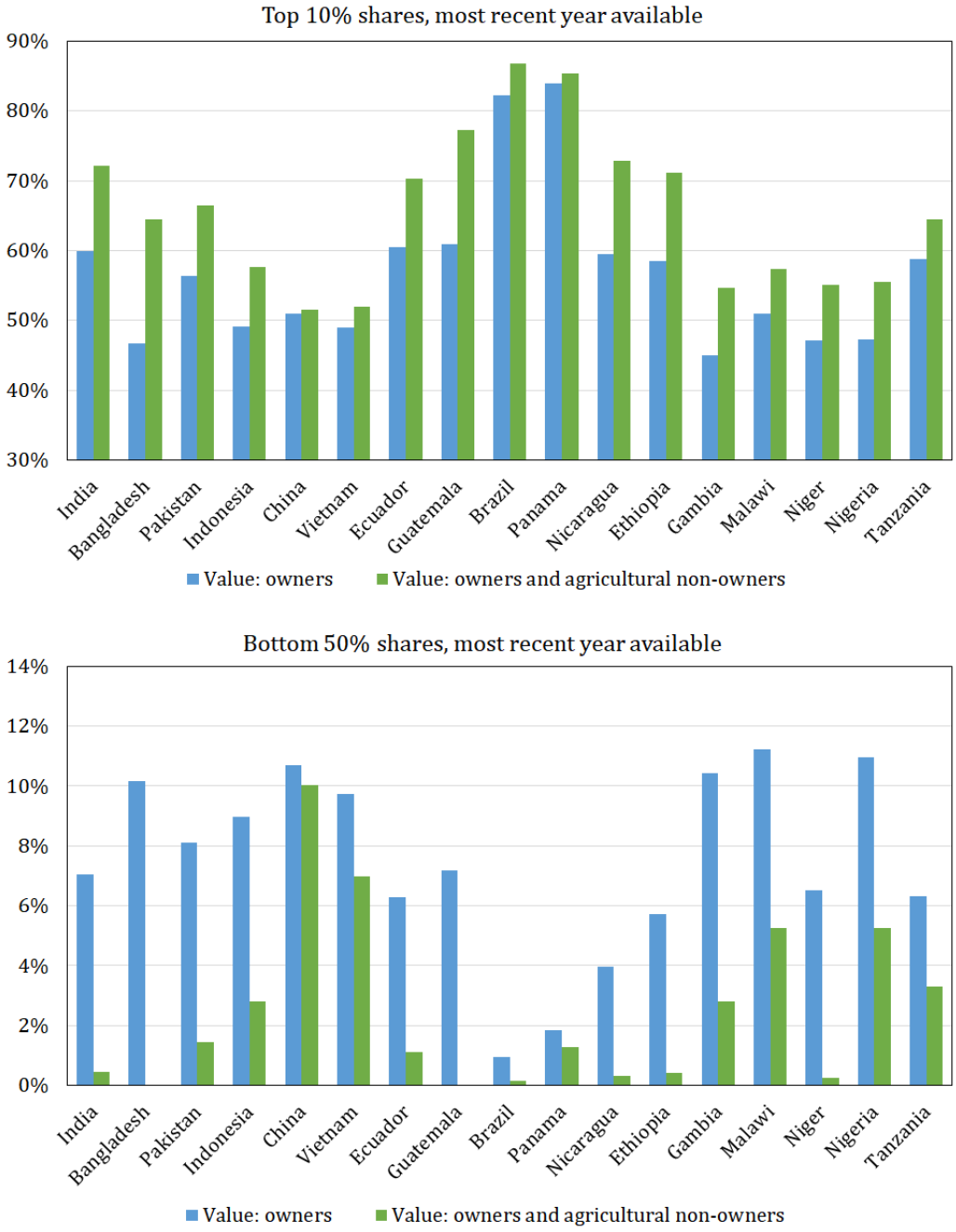


Figure 3.4: a (upper panel) and b (lower panel) : Shares of land value in selected countries of top 10% and bottom 50% of landowners, and landowners plus the landless

inequality present in different countries.

3.3.4 From land area inequality between landowners to accounting for landless people and land values

Figures 3.5a and 3.5b summarize the main results of the paper. They show the agricultural land shares of the top 10% and the bottom 50% for the three concepts examined in this paper: (i) land area inequality between landowners; (ii) land value inequality between landowners; and (iii) land value inequality within the population of landowners plus landless households. Rather than presenting results at the country level (as done in previous sections), Figures 3.5a and 3.5b show the unweighted country averages for four world regions: (i) South Asia – Bangladesh, India, Pakistan, Indonesia; (ii) China and Vietnam; (iii) Latin America – Nicaragua, Ecuador, Guatemala, Brazil, Panama; (iv) Africa: Ethiopia, the Gambia, Malawi, Niger, Nigeria, and Tanzania. The country groupings are based not only on geographical location but also on common patterns in the ownership of agricultural land and in macroeconomic trends (e.g. proportion of employment and value added in agriculture; share of rural population, etc.).

Figures 3.5a and 3.5b condense the main patterns documented in the paper. First, countries in the South Asia region appear to be moderately equal when looking at the distribution of land area between landowners. However, these countries have some of the highest levels of inequality when land values and the landless population are included. China and Vietnam, by contrast, display higher levels of land area inequality between landowners than both South Asia and Africa, but land concentration is only slightly higher when land values and landless households are taken into account. Overall, China and Vietnam appear to be the least unequal world region in our sample, according to our benchmark indicator of inequality.

Latin America (at least as reflected by Nicaragua, Ecuador, Guatemala, Brazil, and Panama) displays the most unequal distribution of agricultural land area between landowners. This also applies to Mexico and Peru (Figure 3.2), and is a finding that has been documented in most Latin American countries based on agricultural censuses (Frankema, 2010). Unlike the other world regions, inequality between landowners is somewhat lower in land value than in land area, a trend driven by three of the five countries: Ecuador, Guatemala, and Panama. When the landless population is included, similar patterns of land inequality are observed, with land value inequality also at one of the highest levels.

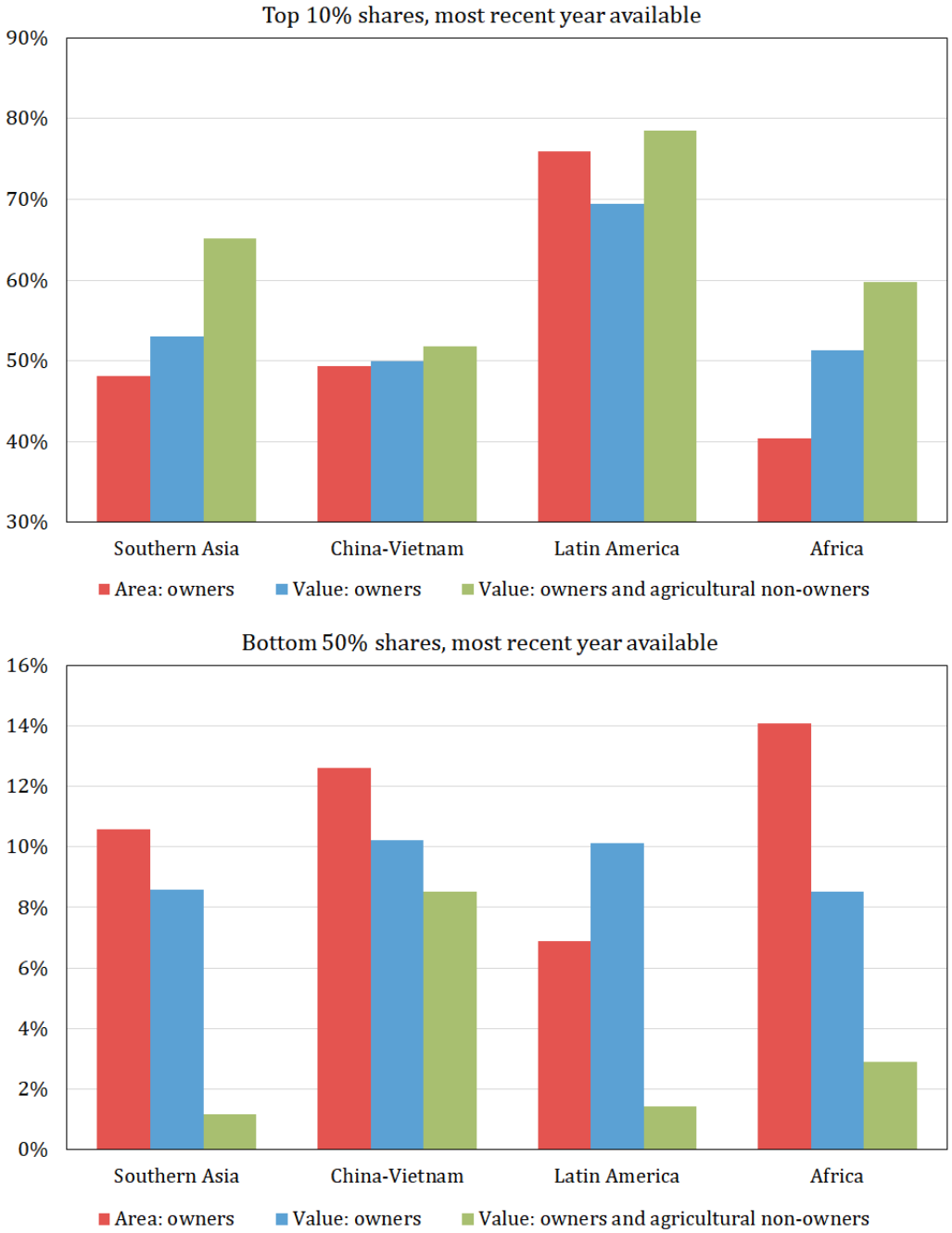


Figure 3.5: a (upper panel) and b (lower panel): Shares of land area and land value held by top 10% and bottom 50% of landowning class and including landless households, in selected countries

Finally, the African countries selected for this analysis occupy an intermediate position between China and Vietnam on the one hand and South Asia and Latin America on the other. Africa has the lowest levels of land area inequality between landowners, but inequality rises gradually when land values and the landless population are included.

3.4 Conclusion

This paper provides the first consistent estimates of agricultural land inequality in developing countries. As such, it presents the most comprehensive overview of the different dimensions of inequality in agricultural land and emphasises the importance of using well defined concepts and clear methodology for measurement. Notably, it shows that we need to go beyond existing studies looking at the size distribution of agricultural holdings based on agricultural censuses. Existing estimates do not reflect either landownership inequality or land value inequality and do not account for the landless population. We advocate instead the use of household surveys as the most appropriate data source to estimate landownership inequality across countries, both in terms of area and value, and to account for non-owners.

Our new estimates provide a novel perspective on international patterns of agricultural land inequality. According to our benchmark metric (i.e. land value inequality including the landless population), South Asia and Latin America show the highest levels of inequality, with the top 10% of landowners capturing up to 75% of agricultural land and the bottom 50% owning less than 2%. The African countries selected display relatively less unequal landownership patterns, while "Communist" Asia (China and Vietnam) represent the region with the lowest levels of inequality.

Having said this, we need to stress that the current results represent a first attempt at assessing agricultural land distribution in developing countries. Although we have included the most populated countries in the analysis, we intend to cover more developing countries to obtain a more complete picture. To do this, we are also developing robust approaches to estimate land values in countries for which surveys provide information on land area only (e.g. Mexico, Egypt, Russia).

To finish, we indicate several methodological extensions of the current work. First, we need to critically assess the role of different forms of landownership, especially those for which distinctions from private property are not clear-cut. Related to this, we need to better understand the importance of corporate land and public land and its impact on distributional patterns.

Finally, given the importance of land for the world's poorest people, we stress the need for governments and international organizations to invest more in collecting more detailed and systematic information on agricultural land in household surveys, especially in countries where data are currently not available.

Bibliography

Banerjee, Abhijit V (1999). "Land reforms: prospects and strategies". In: *Available at SSRN 183711*.

Banerjee, Abhijit V and Esther Duflo (2007). "The economic lives of the poor". In: *Journal of economic perspectives* 21.1, pp. 141–168.

Barrett, Christopher (1996). "On price risk and the inverse farm size-productivity relationship". In: *Journal of Development Economics* 51.2, pp. 193–215.

Benjamin, Dwayne (1995). "Can unobserved land quality explain the inverse productivity relationship?" In: *Journal of Development Economics* 46.1, pp. 51–84.

Berry, Albert and William Cline (1979). *Agrarian Structure and Productivity in Developing Countriess*. John Hopkins University Press.

Besley, Timothy and Robin Burgess (2000). "Land reform, poverty reduction, and growth: Evidence from India". In: *The Quarterly Journal of Economics* 115.2, pp. 389–430.

Binswanger, Hans and Klaus Deininger (1999). *Explaining agricultural and agrarian policies in developing countries*. The World Bank.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2017). "Generalized Pareto curves: theory and applications". In: *Review of Income and Wealth*.

Bourguignon, François (2017). *The globalization of inequality*. Princeton University Press.

Deininger, Klaus (2003). *Land policies for growth and poverty reduction*. World Bank Publications.

Deininger, Klaus and Gershon Feder (2009). "Land registration, governance, and development: Evidence and implications for policy". In: *The World Bank Research Observer* 24.2, pp. 233–266.

Deininger, Klaus and Lyn Squire (1998). "New ways of looking at old issues: inequality and growth". In: *Journal of development economics* 57.2, pp. 259–287.

Do, Quy-Toan and Lakshmi Iyer (2003). "Land Rights and Economic Development: Evidence from Vietnam". In: *World Bank, Policy Research Working Papers*.

Doss, Cheryl et al. (2015). "Gender inequalities in ownership and control of land in Africa: myth and reality". In: *Agricultural Economics* 46.3, pp. 403–434.

Easterly, William (2007). "Inequality does cause underdevelopment: Insights from a new instrument". In: *Journal of development economics* 84.2, pp. 755–776.

Engerman, Stanley and Kenneth Sokoloff (1997). *Factor Endowments, Institutions and Differential Paths of Growth among the New World Economies," in Stephen Haber, ed., How Latin America Fell Behind, Stanford: Stanford University Press.*

Erickson, Lennart and Dietrich Vollrath (2004). *Dimensions of land inequality and economic development*. 4-158. International Monetary Fund.

Eswaran, Mukesh and Ashok Kotwal (2006). "The role of agriculture in development". In: *Understanding poverty*, pp. 111–123.

FAO (1999). *Agricultural Censuses and Gender Considerations - Concept and Methodology*. Rome, Italy: Food and Agriculture Organization of the United Nations.

– (2016). *The State of Food and Agriculture: Climate Change, Agriculture and Food Security*. Rome, Italy: Food and Agriculture Organization of the United Nations.

– (2018). "World programme for the census of agriculture 2020". In: *FAO statistical development series* 2.16.

Foster, Andrew and Mark Rosenzweig (2017). *Are there too many farms in the world? Labor-market transaction costs, machine capacities and optimal farm size*. Tech. rep. National Bureau of Economic Research.

Frankema, Ewout (2010). "The colonial roots of land inequality: geography, factor endowments, or institutions?" In: *The Economic History Review* 63.2, pp. 418–451.

Galor, Oded, Omer Moav, and Dietrich Vollrath (2009). "Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence". In: *The Review of economic studies* 76.1, pp. 143–179.

Gollin, Douglas, Stephen Parente, and Richard Rogerson (2007). "The food problem and the evolution of international income levels". In: *Journal of Monetary Economics* 54.4, pp. 1230–1255.

Griffin, Keith, Azizur Rahman Khan, and Amy Ickowitz (2002). "Poverty and the Distribution of Land". In: *Journal of Agrarian change* 2.3, pp. 279–330.

Guereña, Arantxa and Marc Wegerif (2019). "Land Inequality Framing Document". In: *International Land Coalition*.

Kanbur, Ravi (2000). "Income distribution and development". In: *Handbook of income distribution* 1, pp. 791–841.

Lakner, Christoph and Branko Milanovic (2015). *Global income distribution: from the fall of the Berlin Wall to the Great Recession*. The World Bank Economic Review.

Li, Shi and Renwei Zhao (2007). "Changes in the Distribution of Wealth in China, 1995-2002". In: *WIDER Research Paper* 2007/03.

Lipton, Michael (2009). *Land reform in developing countries: property rights and property wrongs*. Routledge.

Lowder, Sarah, Jakob Skoet, and Terri Raney (2016). "The number, size, and distribution of farms, smallholder farms, and family farms worldwide". In: *World Development* 87, pp. 16–29.

Martinelli, Pablo (2016). "Land Inequality in Italy in 1940: The New Picture". In: *Rivista di storia economica* 32.3, pp. 303–350.

McKinley, Terry and Keith Griffin (1993). "The distribution of land in rural China". In: *The Journal of Peasant Studies* 21.1, pp. 71–84.

Midlarsky, Manus I (1988). "Rulers and the ruled: Patterned inequality and the onset of mass political violence". In: *American political science Review* 82.2, pp. 491–509.

Morett-Sánchez, J Carlos and Celsa Cosıo-Ruiz (2017). "Panorama de los ejidos y comunidades agrarias en México". In: *Agricultura, sociedad y desarrollo* 14.1, pp. 125–152.

Piketty, Thomas, Li Yang, and Gabriel Zucman (2019). "Capital accumulation, private property, and rising inequality in China, 1978–2015". In: *American Economic Review* 109.7, pp. 2469–96.

Restuccia, Diego and Raul Santaeulalia-Llopis (2017). *Land misallocation and productivity*. Tech. rep. National Bureau of Economic Research.

Restuccia, Diego, Dennis Yang, and Xiaodong Zhu (2008). "Agriculture and aggregate productivity: A quantitative cross-country analysis". In: *Journal of monetary economics* 55.2, pp. 234–250.

Rodrik, Dani (1995). "Getting interventions right: how South Korea and Taiwan grew rich". In: *Economic policy* 10.20, pp. 53–107.

– (2016). "Premature deindustrialization". In: *Journal of economic growth* 21.1, pp. 1–33.

Sen, Amartya (1966). "Peasants and Dualism with or without Surplus Labor". In: *Journal of political Economy* 74.5, pp. 425–450.

Sokoloff, Kenneth and Stanley Engerman (2000). "Institutions, factor endowments, and paths of development in the new world". In: *Journal of Economic perspectives* 14.3, pp. 217–232.

Taylor, Charles Lewis and David Jodice (1983). *World handbook of political and social indicators*. Yale University Press,

Vollrath, Dietrich (2007). "Land distribution and international agricultural productivity". In: *American Journal of Agricultural Economics* 89.1, pp. 202–216.

Yang, Dennis and Xiaodong Zhu (2013). "Modernization of agriculture and long-term growth". In: *Journal of Monetary Economics* 60.3, pp. 367–382.

Appendix A

Appendix to "Is Naturalization a Passport for a Better Labor Market Integration?"

A.1 Design

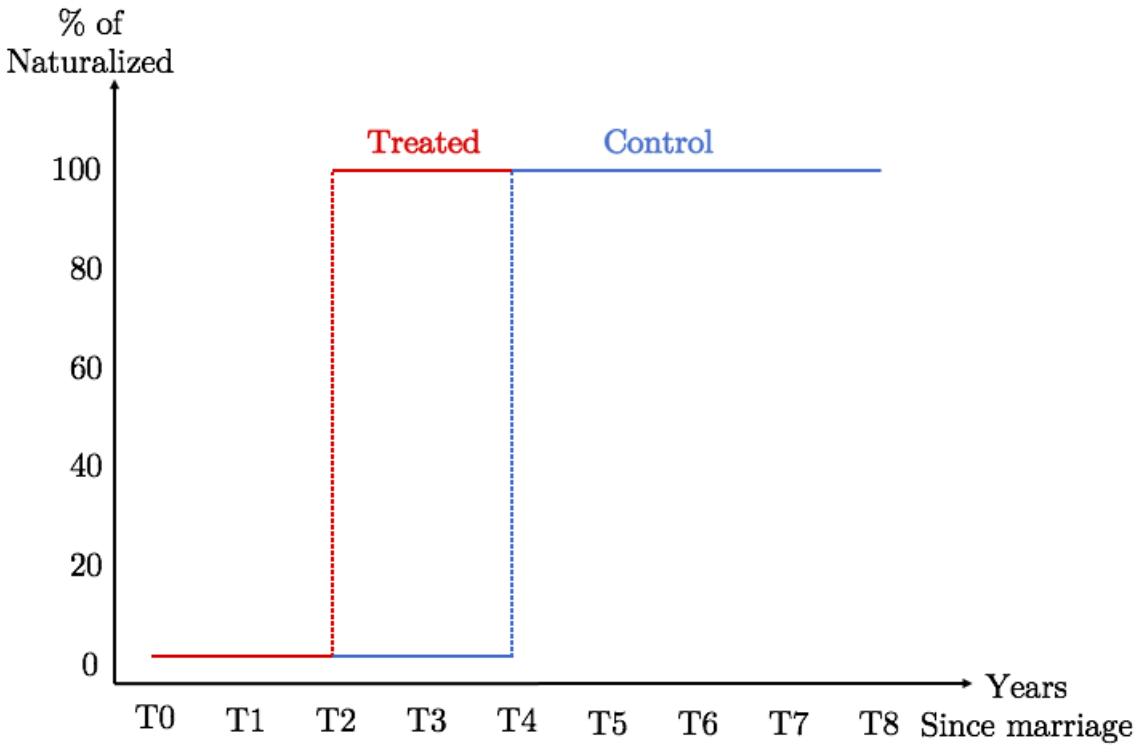


Figure A.1.1: Expected share of naturalization in treated and control group under full compliance

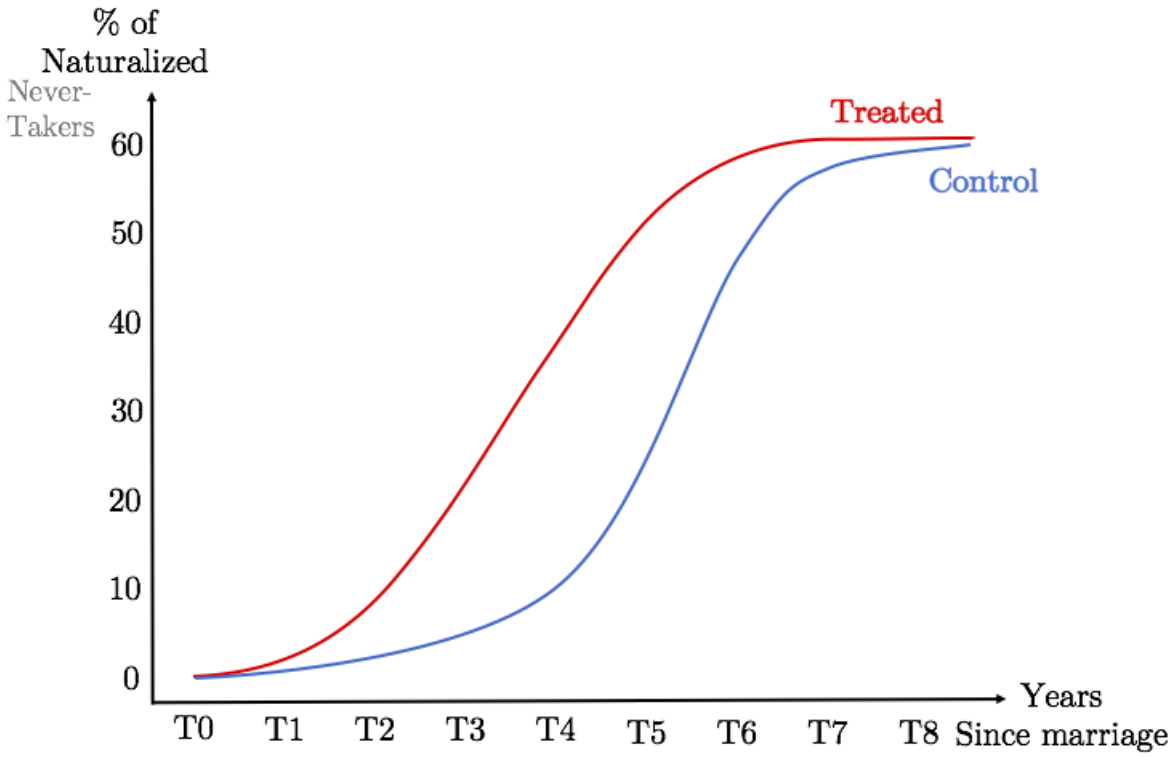
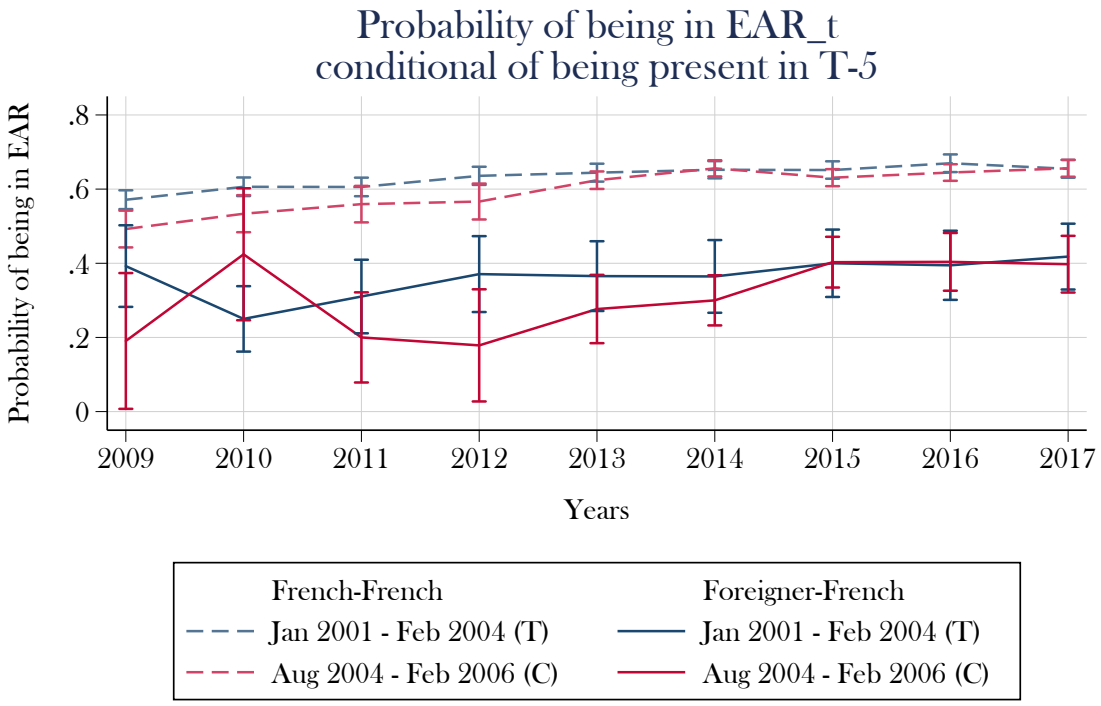


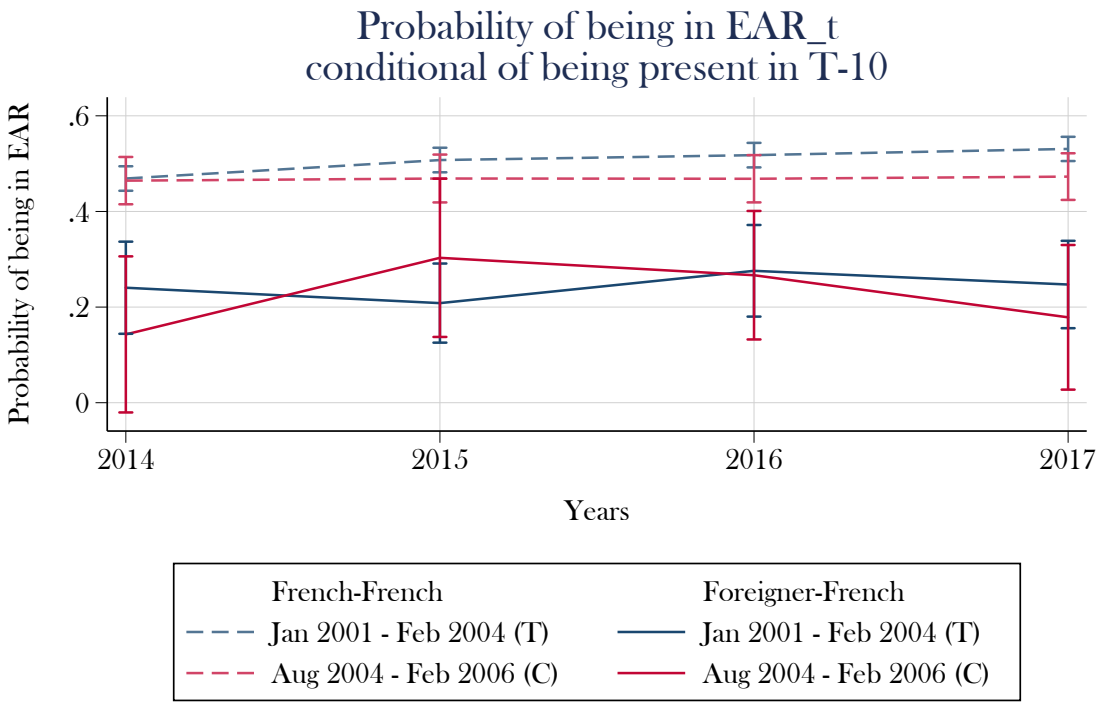
Figure A.1.2: Share of naturalized among treated and control group with non-compliance

A.2 Probability of attrition between teh treated and control group

Figure A.2.1 seems to suggest that there is no differential attrition rate between the early-treated and late-treated group. The upper (lower) panel, shows that the probability of being in the population census 5 (10) years after being observed in a given year T is similar for the early-treated and late-treated group. The following figures show the average number of years in France and the level of education over the years, showing that there is no major change within the period of interest.



E.g Probability of being observed in 2009 for Fr-Fr marriage in treated and control group is around 50% conditional of being in the 2004 census. The same for For-Fr marriage is around 20% in the treated group and 38% in the control group. The differences are not statistically sig.



E.g Probability of being observed in 2014 for Fr-Fr marriage in treated and control group is around 40% conditional of being in the 2004 census. The same for For-Fr marriage is around 20% in the treated group and the control group. The differences are not statistically sig.

Figure A.2.1: Sample Composition

A.3 Descriptive Statistics

	Foreigners to French	All foreign-born	Difference
Age at Arrival	24.11	17.19	6.92***
	(10.19)	(11.60)	(0.18)
Undergraduate or above	0.41	0.36	0.05***
	(0.49)	(0.48)	(0.01)
Manual worker	0.28	0.22	0.06***
	(0.45)	(0.41)	(0.01)
Employees	0.28	0.29	-0.01*
	(0.45)	(0.45)	(0.01)
Intermediate Professions	0.19	0.21	-0.02***
	(0.39)	(0.41)	(0.01)
Executives	0.17	0.19	-0.02**
	(0.38)	(0.39)	(0.01)
Origin from Maghreb	0.42	0.39	0.03***
	(0.49)	(0.49)	(0.01)
Francophone	0.61	0.56	0.05***
	(0.49)	(0.50)	(0.01)
Observations	7,385	10,226	18,061

Table A.1: Sample selection based on observable characteristics

	Foreigner-French			Foreigner-Foreigner			
	(1) Early-Treated	(2) Late-treated	(3) Diff	(4) Early-Treated	(5) Late-treated	(6) Diff	(7) Diff of Diffs
Age	31.30	33.14	-1.84***	34.80	34.52	0.28	-2.13***
	(6.14)	(7.71)	(0.41)	(8.21)	(7.67)	(0.54)	(0.66)
Age Diff.	5.66	5.27	0.38	6.95	5.52	1.43***	-1.05**
	(5.26)	(5.15)	(0.30)	(6.07)	(4.79)	(0.36)	(0.47)
% of women	0.33	0.33	0.00	0.48	0.44	0.04	-0.04
	(0.47)	(0.47)	(0.03)	(0.50)	(0.50)	(0.03)	(0.04)
Prob(Panel)	0.75	0.73	0.02	0.77	0.69	0.08***	-0.05
	(0.43)	(0.44)	(0.02)	(0.42)	(0.46)	(0.03)	(0.04)
Full-time	0.65	0.69	-0.04	0.63	0.68	-0.05	0.01
	(0.48)	(0.46)	(0.03)	(0.48)	(0.47)	(0.04)	(0.05)
No of hours	1140.5	1178.6	-38.1	1205.2	1278.5	-73.3	35.3
	(690.6)	(688.9)	(45.2)	(665.6)	(725.0)	(55.7)	(71.5)
Annual earnings	12264.5	13337.2	-1072.7	13459.0	14470	-1011	-61.8
	(10310.7)	(10008.5)	(664)	(10445.2)	(11329.1)	(872.2)	(1079.9)
Observations	531	768	1,299	342	588	930	2,229

Table A.2: Balancing Test between treated and control groups among the group of interest and never-treated group

A.4 Placebo Analysis

Figure A.4.1 shows the differential naturalization rate between early-treated and late-treated groups when changing the reform timing. The top left panel corresponds to the actual date of the reform, July 2006 and is exactly the same as Figure 1.3. The top-right panel of Figure A.4.1 shows the differential rates under the assumption that the reform occured in July 2008. In the bottom left and right panels, the reform date is assumed to be in July 2010 and 2012 respectively[1]. There seems to be no significant differential naturalization rates under the three placebo scenarios.

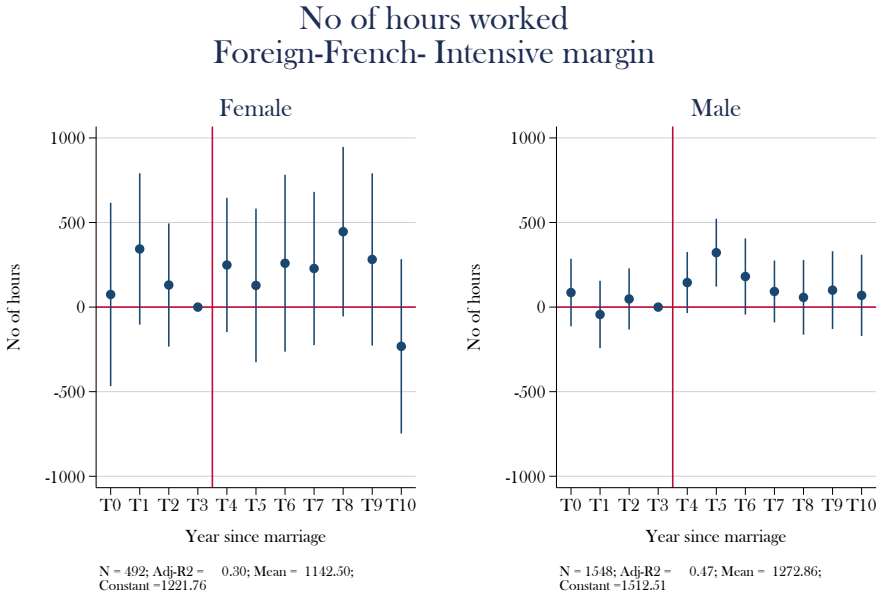
Excluding Europeans

Figure A.4.1: Placeb test: Difference in naturalization rate based on different reform dates

[1]Choosing a more recent reform date restricts the number of periods after marriage that can be observed in the data, knowing that the latest year for which population census data is available is 2016.

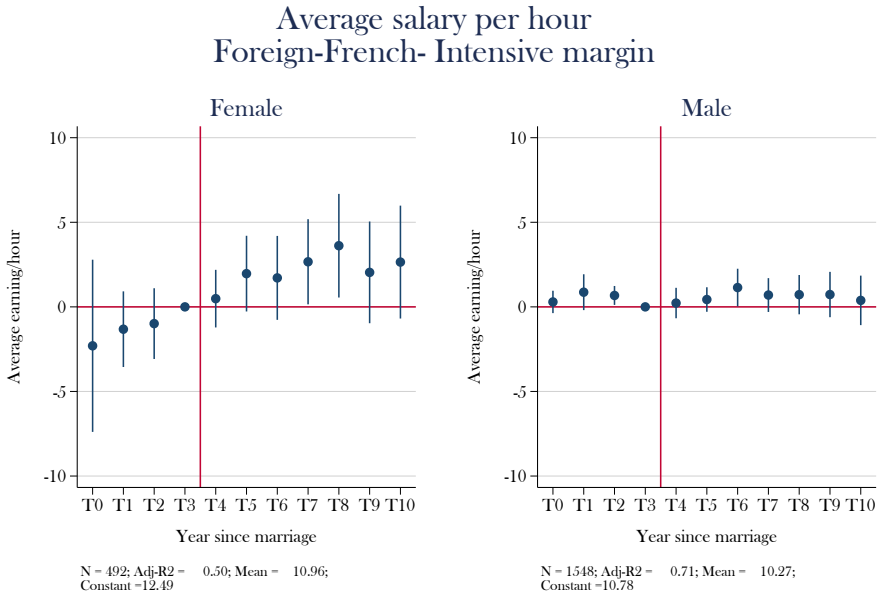
A.5 Heterogeneity Analysis & Mechanisms

E.1. Decomposition by Gender



Simple regression with only foreign-french marriage. With Individual FE.
Clustered SE. Missing values not taken into account

Figure A.5.1: Heterogeneous analysis by gender: Dynamic effect on the number of hours worked



Simple regression with only foreign-french marriage. With Individual FE.
Clustered SE. Missing values not taken into account

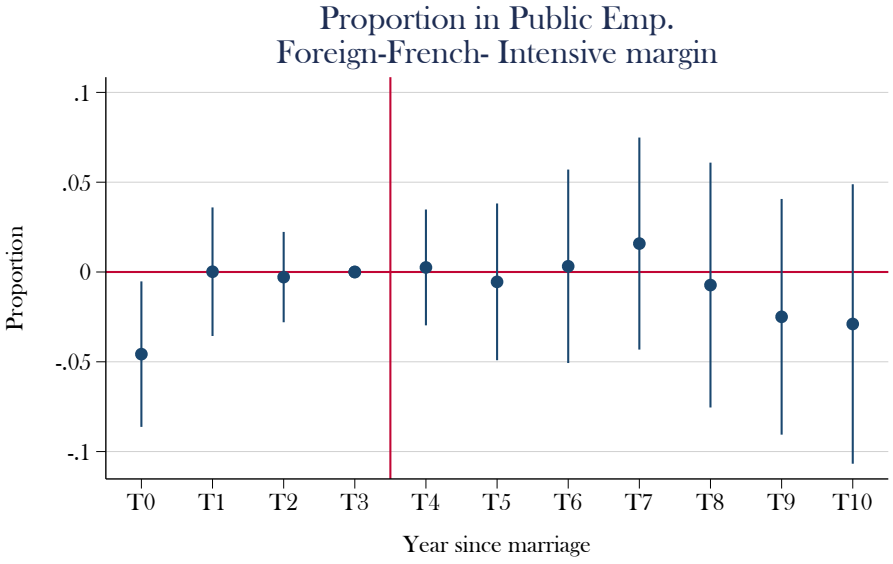
Figure A.5.2: Heterogeneous analysis by gender: Dynamic effect on the hourly wages

E.2. Labor market access

	(1) Full-time emp.	(2) Public sector emp.
Post x Treat	0.05	0.03
	(0.07)	(0.03)
Observations	3,238	3,238
Adj R-squared	0.34	0.59
Ind. FE	Yes	Yes
Mean	0.740	0.0386

The table present the difference-in-differences coefficient. The pre-period consist of the first three years of marriage (T0 - T3) and the post-period is defined as time periods beyond the third year of marriage (T4 - T10). Results are conditional on working in the pre-treatment period. Standard errors clustered at the individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Effect on full-time employment and public sector employment



Simple regression with only foreign-french marriage. With Individual FE.
Clustered SE Missing values not taken into account
N = 2237; Adj-R2 = 0.62; Mean = 0.04; Constant =.04

Figure A.5.3: Dynamic effect on public sector employment

E.2. Informal sector

	(1)	(2)	(3)	(4)	(5)	(6)
	Construction			Non-Construction		
VARIABLES	Net annual earnings	No of hours worked	Hourly wages	Net annual earnings	No of hours worked	Hourly wages
Post x Construction	7,140.29***	496.13***	1.03	1,607.99	49.82	0.94**
	(2,220.83)	(141.67)	(0.84)	(1,187.22)	(69.17)	(0.47)
Observations	220	220	220	1813	1813	1813
Adj R-squared	0.63	0.51	0.56	0.65	0.44	0.62
Ind. FE	Yes	Yes	Yes	Yes	Yes	Yes

Clustered standard errors at the individual-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
The pre-period consist of T0 - T3 and the post-period is defined as T4 - T10.

Table A.4: Heterogeneous analysis by sector: construction v/s non-construction

A.6 Divorce

In this section, I will look at the effect of the 2006 reform on divorce. There different factors that might come into play when analyzing the effect on divorce rates. First, the reform aimed at fighting against fraud marriage. These marriages are considered as marriages for the sole purpose of obtaining the nationality, and by definition, these are more likely to end in divorce once the nationality is obtained. This would result in an increase the divorce rate due to a higher separation rate among the treated group after obtaining the nationality.

Additionally, the effect of the reform on the late-treated group couples' incentive to remain or separate can be two-fold. First, the additional "burden" of a longer waiting period might induce the couples in the late-treated group to divorce more. However, as explained in section 2.2, given that the naturalization process through the marriage channel remains the relatively easiest channel despite the longer waiting period, these couples might instead have the incentive to stay longer in the couple. This would lead to a reduction in divorces among the late-treated group and a net positive effect of naturalization on divorce rates. Finally, better labor market outcomes due to naturalization might int themselves lead to higher divorce rates. In all, one expects to find a positive effect on divorce, not necessarily attributed to fraud marriages.

In order to understand the general effect of naturalization on the divorce rate, equation 2 and 2b are estimate with Y is a dummy for being divorced for each individual i at time t. Figure A.6.1 shows the dynamic results of a difference-in-differences analysis on the probability of being divorced. This result tend to confirm the positive effect of the reform on divorce rates.

Accounting for year trends in this case is more complex than in the main analysis since divorce is a joint decision in the couple. Hence, it is not clear whether the relevant never-treated group should be french married to french citizens or instead, foreigners married to foreigners. Figure **??** shows the dynamic results for the triple difference estimation for both never-treated group. The left panel shows the results when compared to french married to french couples and the right panel shows the equivalent when the never-treated group are foreigners married to foreigners. These suggest that the result on divorces are sensitive to the choice of the comparison group. Taken together, it is unclear if the reform had a positive effect on divorce rates.

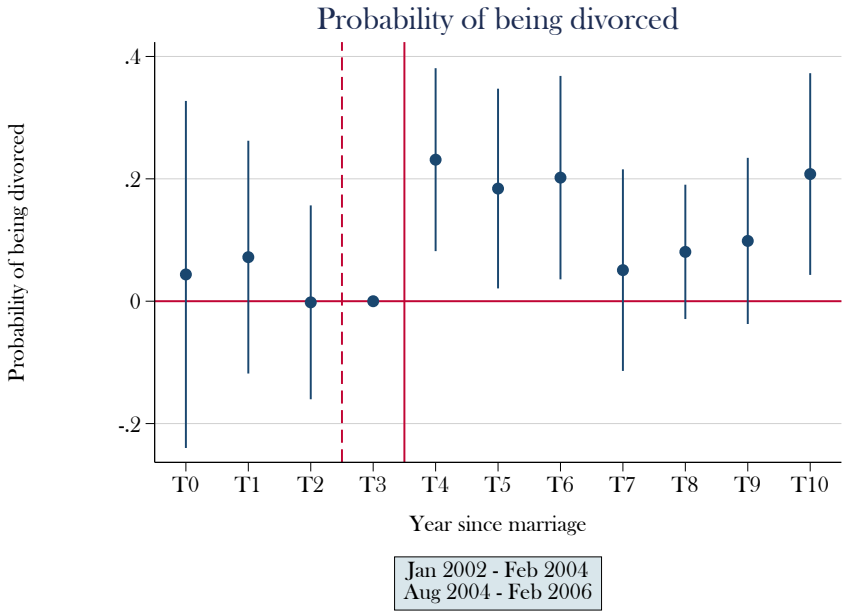
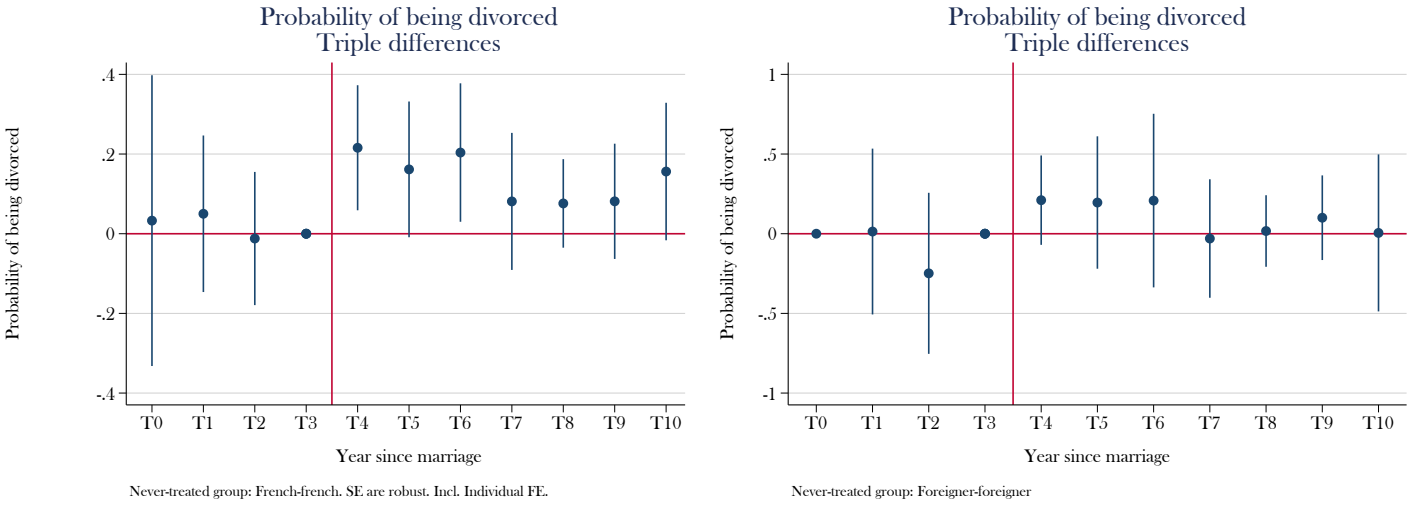


Figure A.6.1: Dynamic effect on the probability of divorce



(a) French married to french

(b) Foreigners married to foreigners

Figure A.6.2: Dynamic effect on the probability of divorce. Triple differences with a never-treated group

Appendix B

Appendix to "Post-colonial Trends of Income Inequality"

B.1 Demographics

As seen in Figure B.1.6 to B.1.9, slaves constituted around 60-80% of the population, the rest being white population in the mid-18th century. By 1842, the white population constituted only around 6-10% of the population in the Antilles and Guyane compared to around 20% of the population in La Réunion.

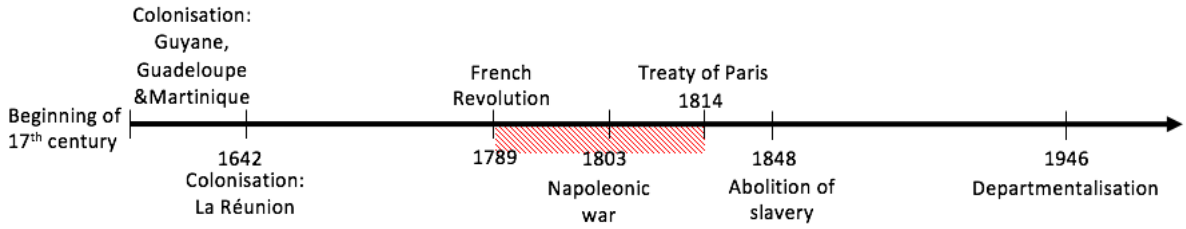


Figure B.1.1: Main events of the overseas departments' history

The growing share of freed colored population together with the emancipation of slavery in neighboring British colonies has led to mounting pressures on the local colonial forces to give in to the abolition of slavery in the four "old colonies". In 1848, the ex-slaves in these colonies were all emancipated and acceded a pseudo-citizenship status. The constant need for cheap labor led to the immigration of Africans and Indians to these territories. It is only a century later, in 1946 that these territories were fully transformed into French departments. This rather rare form of decolonization process was thought in a logic of institutional, judicial and cultural assimilation. Three centuries of colonial domination was deemed ample to instill

French values in the population.

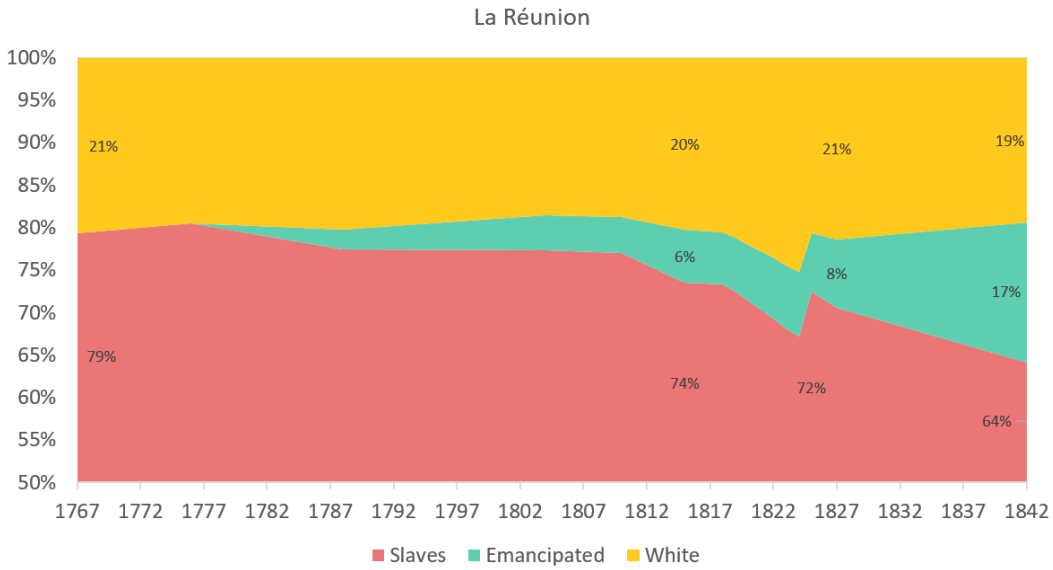


Figure B.1.2: Colonial Population Composition: La Réunion

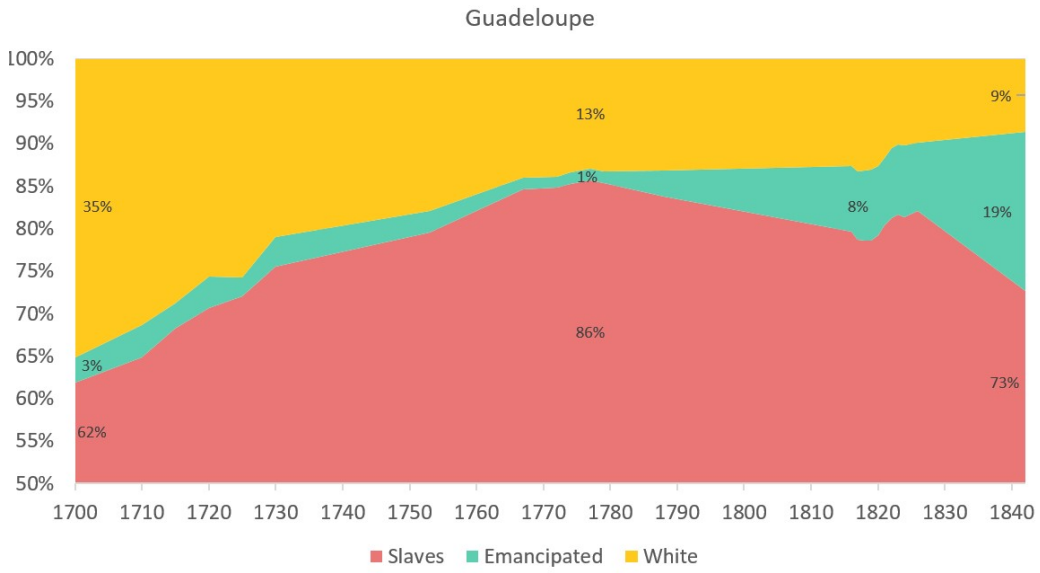


Figure B.1.3: Colonial Population Composition: Guadeloupe

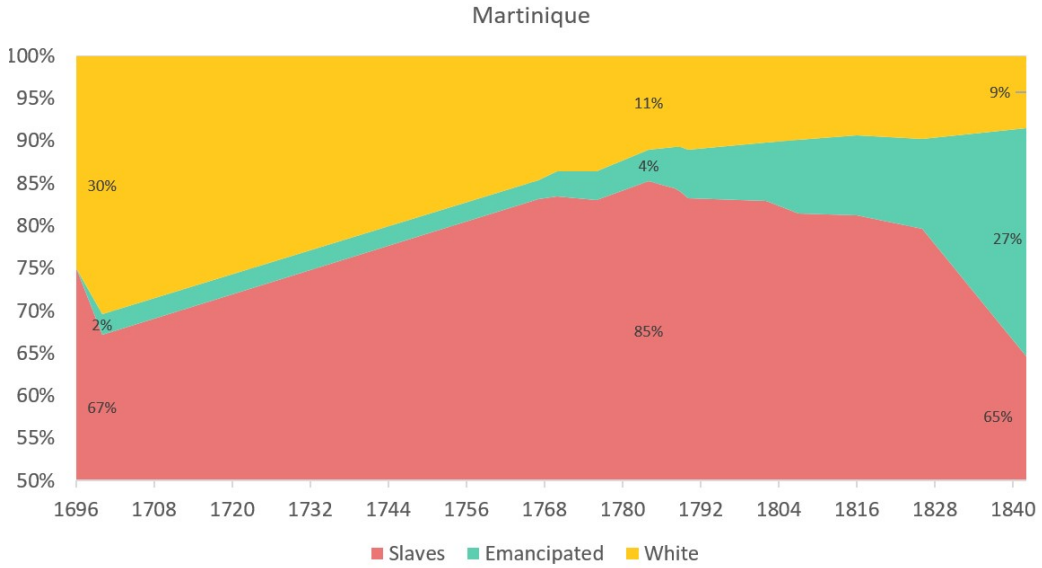


Figure B.1.4: Colonial Population Composition: Martinique

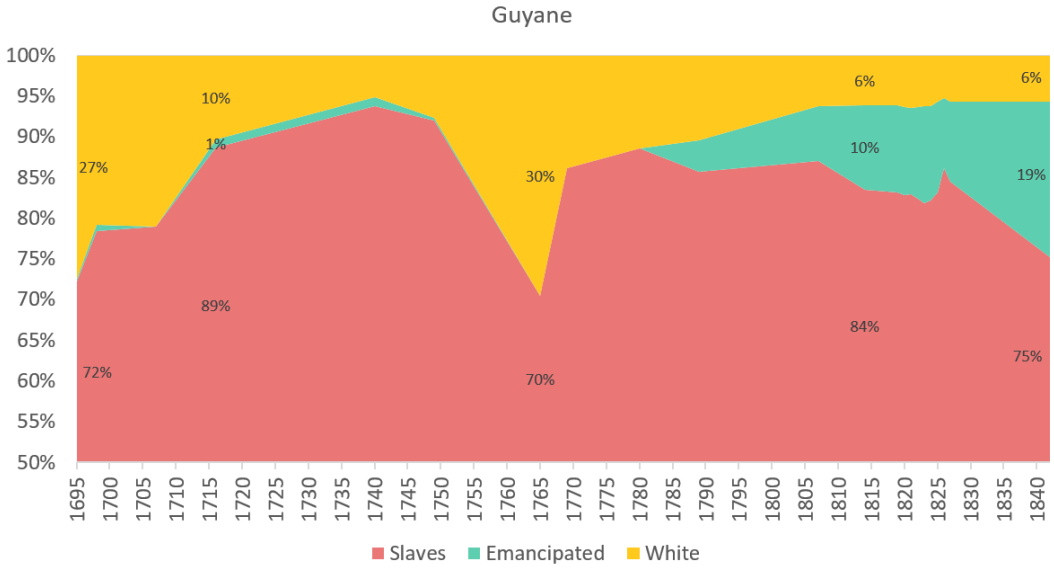


Figure B.1.5: Colonial Population Composition: Guyane

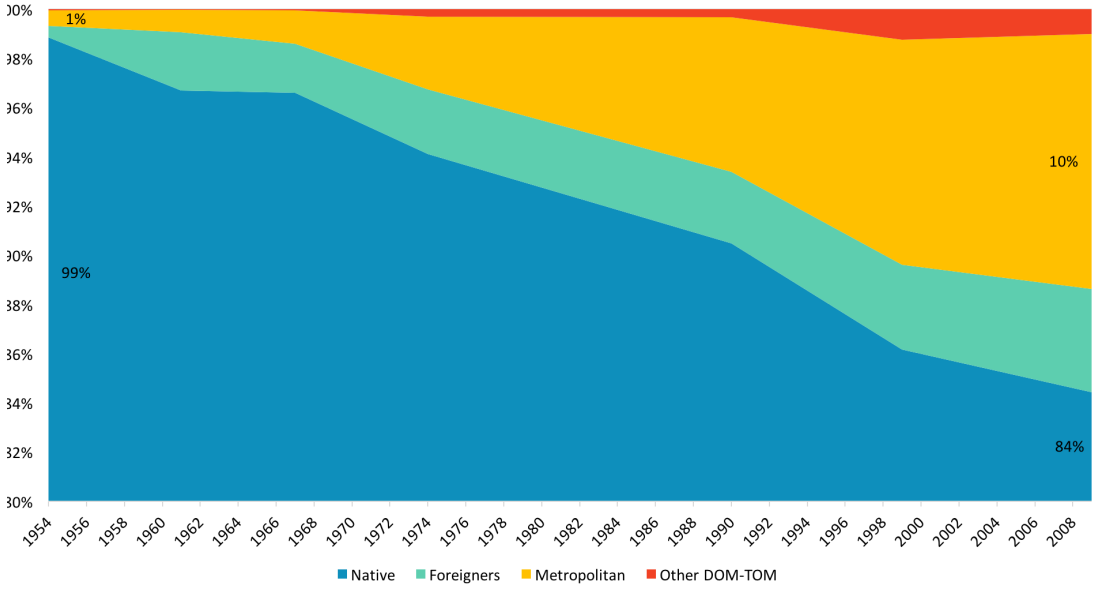


Figure B.1.6: Post-colonial Population Composition: La Réunion

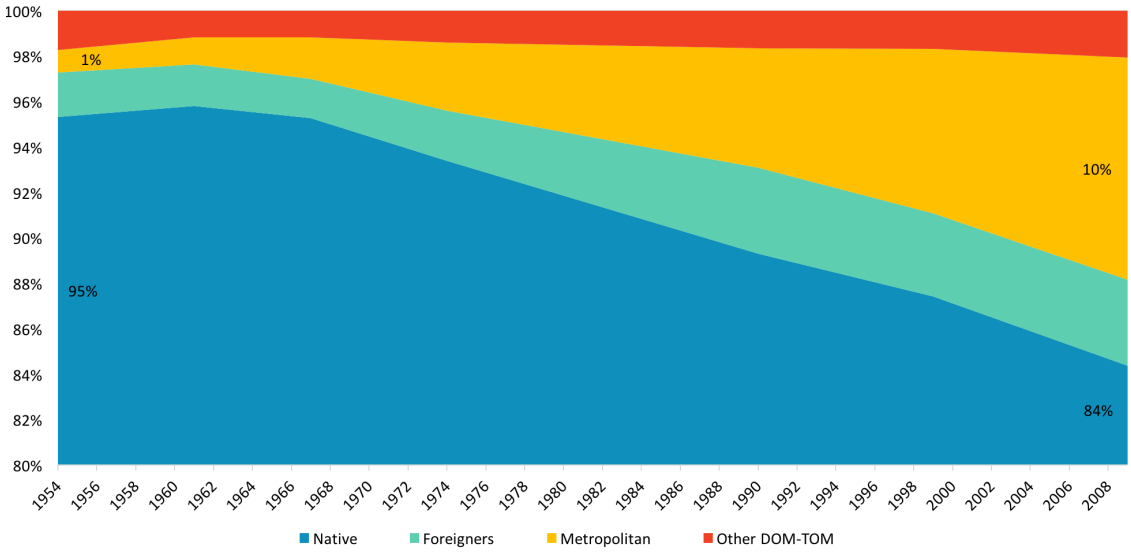


Figure B.1.7: Post-colonial Population Composition: Guadeloupe

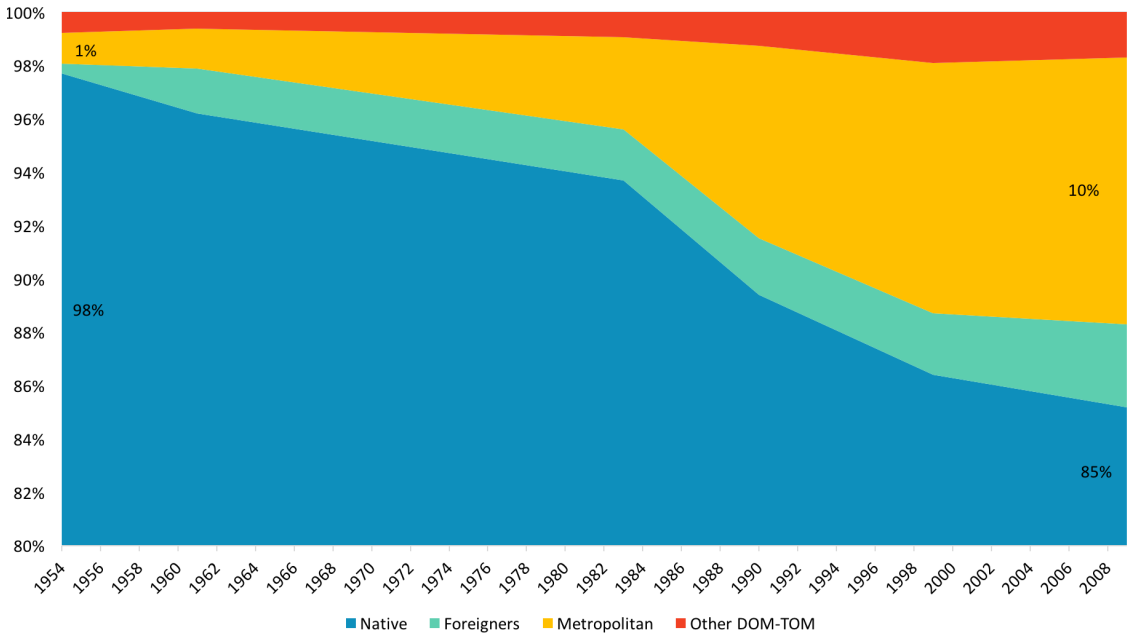


Figure B.1.8: Post-colonial Population Composition: Martinique

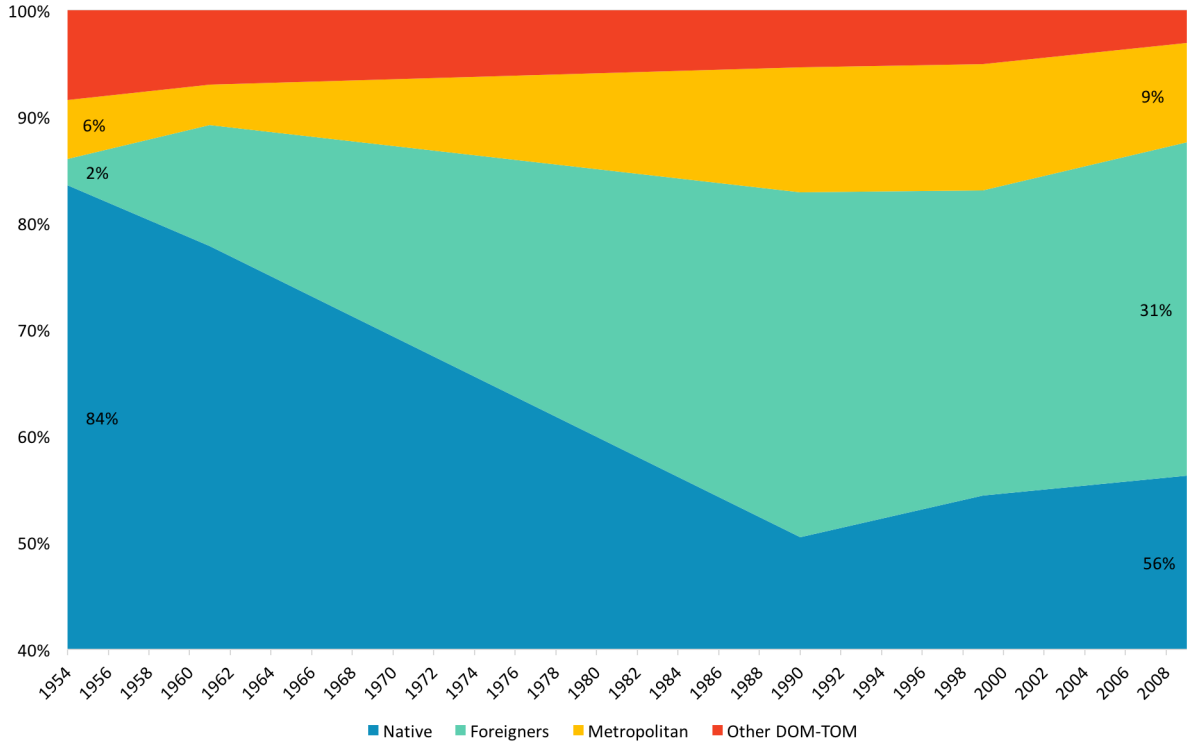


Figure B.1.9: Post-colonial Population Composition: Guyane

B.2 Income Tax Data

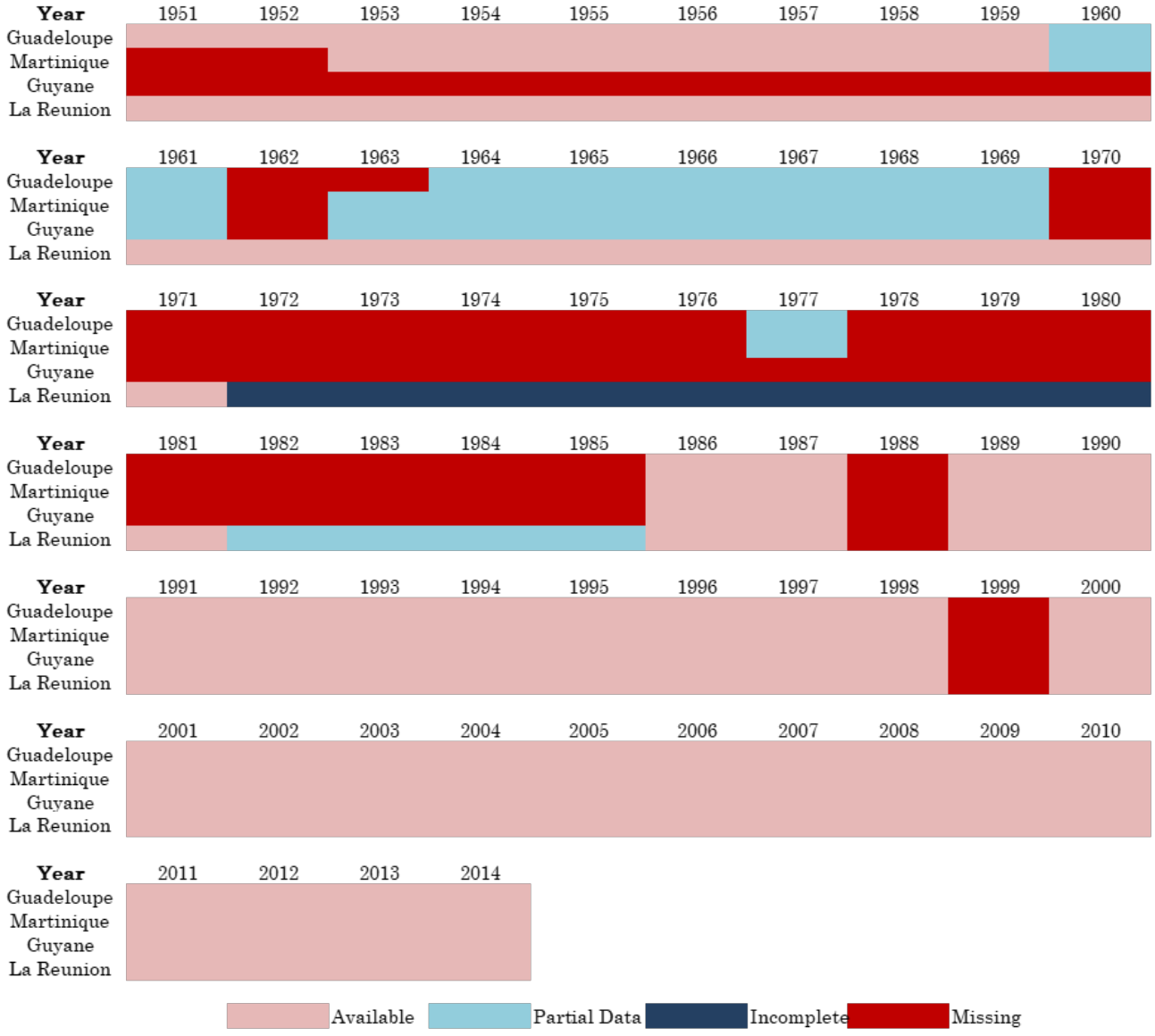


Figure B.2.1: Availability of Tax Data

B.3 Control Total for Population

In order to estimate the distribution of income, there is a need to estimate the total tax units that should have been observed in the income tax data, had every tax unit been required to fill a tax form. It should be noted that a person living in France can detach from his/her parents' tax unit and thus declare his/her income separately as from the age of 18. However, it is not mandatory to do so until the age of 21, except for 24 years old for unmarried students or in a liberal profession. In addition, married people (including PACSed couples) are required to fill a unique tax declaration. Given this setting, the control total for population (TU_{it}) is estimated as the number of adult population (A_{it}) deducting the number of married couple (M_{it}) in order to avoid double counting married couple.

$$TU_{it} = A_{it} - M_{it} \tag{B.1}$$

These data are obtained from the Population Census in the overseas departments for the following years of census: 1954, 1961, 1967, 1974, 1982, 1990, 1999, 2009 and 2014. It is linearly interpolated for the years for which we don't have this information. The age threshold at which we define the adult population can be set in different ways, namely at 18, 19 or 20 years old. In this paper, the definition of adult population is taken as the population above 20 years, as is widely done in this literature, for two main reasons:

i) given that the estimate of control population based on the definition of adult population above 20 years seem to provide a good enough approximation of the total number of tax declarations (See figure B.3.6);

ii) given that the population census reports, which dates back to the 1950s, report population by pre-defined age groups. The age group is typically as follows: 15 - 19 years old, 20 - 24 years old and so on. Hence, a threshold of 18 or 19 years would require further hypotheses on the distribution of the population within the age group 15 - 19 years to estimate the adult population of interest. Hence, to have the most consistent method in estimating the control population, 20 years old is the threshold taken for defining adult population.

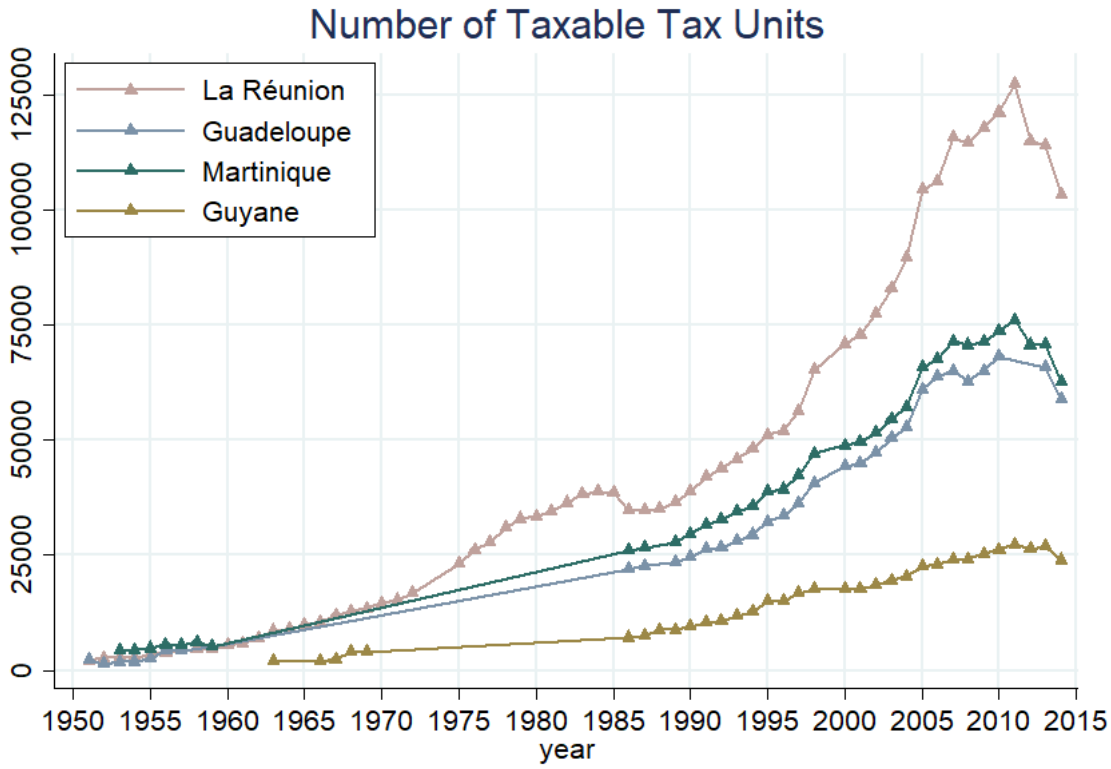


Figure B.3.1: Number of taxable tax units (1954 - 2014)

Similarly, the number of married couples is estimated from the population census data by taking the average number of married individuals divided by 2. While this is not a perfect count of the number of joint fillings, it should nevertheless give a more or less precise estimate of the latter.

Figure B.3.1 and B.3.2 show the number of taxable tax units since the 1950s and the number of non-taxable tax units since the mid-1980s respectively. There is a clear upward trend in both the number of taxable and non-taxable units since the beginning of the period, with a slight downward turn at the end of the period for the number of taxable. Figure B.3.3 shows the total number of declarations to the tax office and the total estimated tax units (using equation B.1 over the years in La Réunion. We observe a very small number of declarations in the years prior to 1986. Thereafter, with the requirement for nontaxable to declare their income, there is a steady rise in the total number of declarations reaching the number of estimated tax units in the early 2000s. Given this trend, we make the hypothesis that we start to observe all the tax units in the income tax data as from the year 2003. The total number of tax units estimated from the definition above is presented in figure B.3.4.

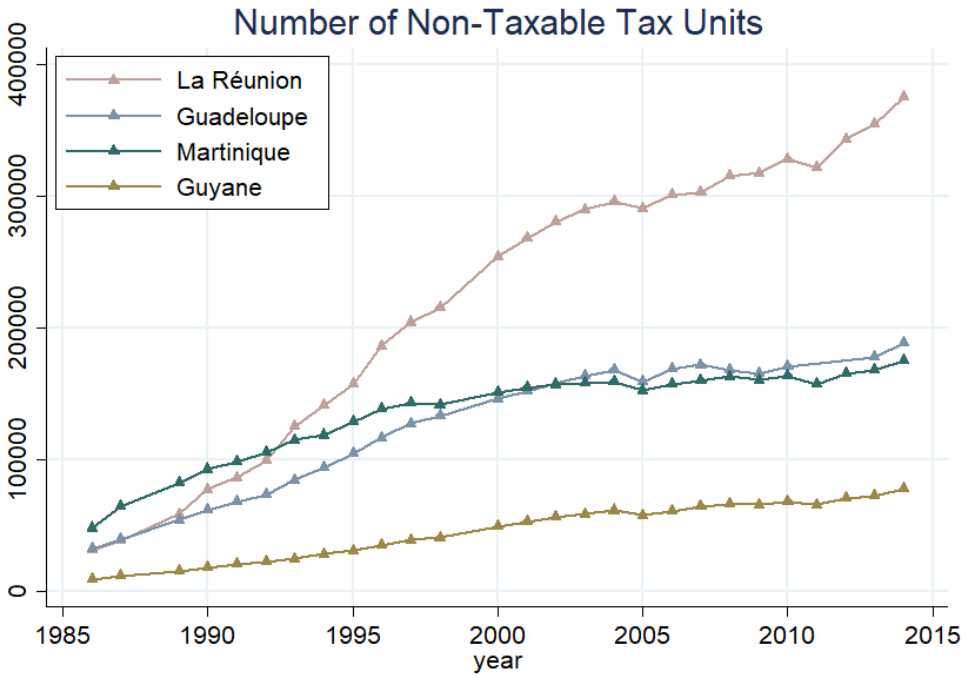


Figure B.3.2: Number of non-taxable tax units (1954 - 2014)

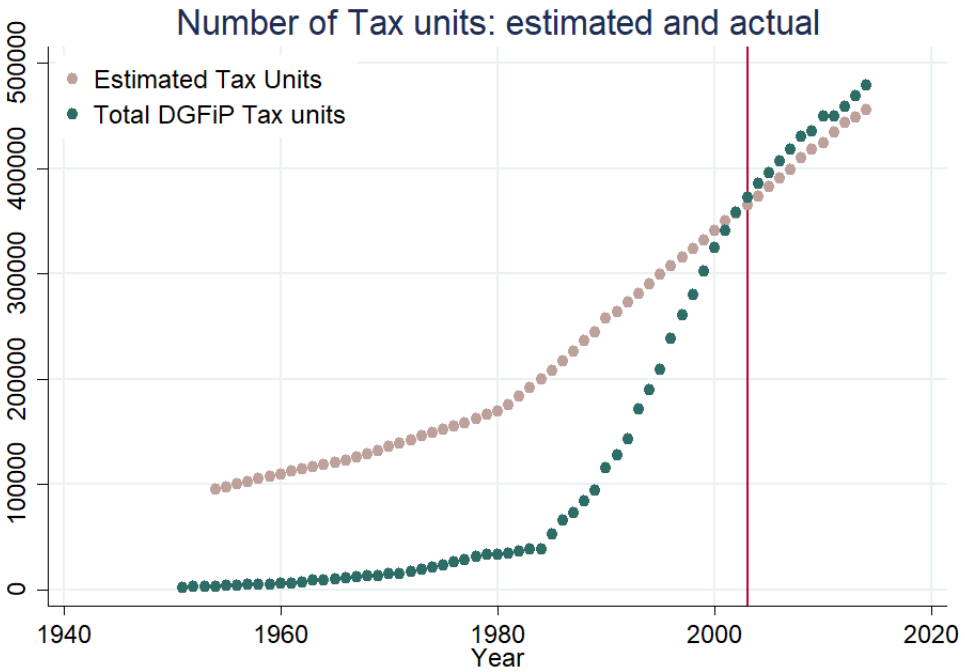


Figure B.3.3: Total number of declarations and total estimated tax units in La Réunion

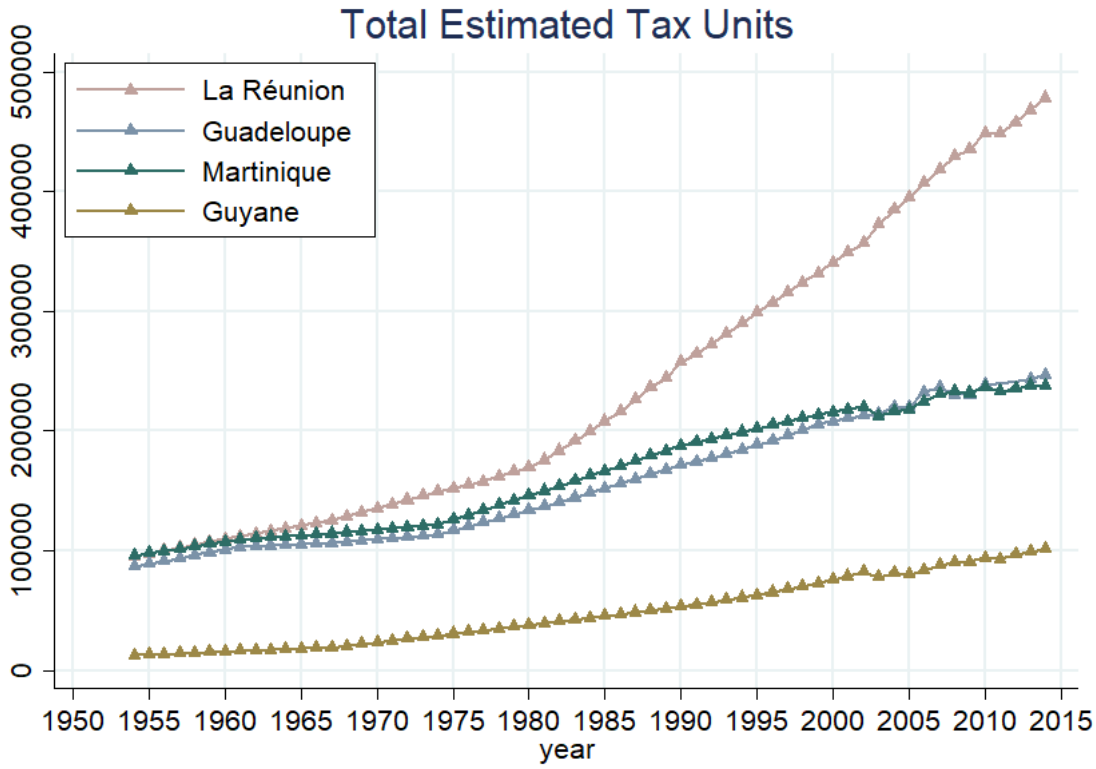


Figure B.3.4: Total taxable tax units (1954 - 2014)

In order to get a better understanding of the evolution of declarations over time, the proportion of declarations (P) is estimated. P is simply the number of declarations (D) divided by the total number of tax units (Tot):

$$P_{it} = \frac{D_{it}}{Tot_{it}} \tag{B.2}$$

The numerator in equation B.2 refers to the total number of tax units reported by the tax authorities for an overseas department i at time t, while the denominator is our estimate of tax units obtained from equation B.1. We observe a general increase in the proportion of declarations from the mid-20$^{\text{th}}$ century until recent years, partly due to the non-declaration of non-taxable tax units at the beginning of the period. As from the mid-80s, both taxable and non-taxable tax units are required to declare their income, and we observe a steady increase in the proportion of declaration from the mid-1980s until the early 2000s and a stabilization thereafter. This suggests that, as from the beginning of the 21$^{\text{st}}$ century, we observe more or less everyone in the tax data. In effect, we should be observing a proportion of declaration of 100% in the recent years.

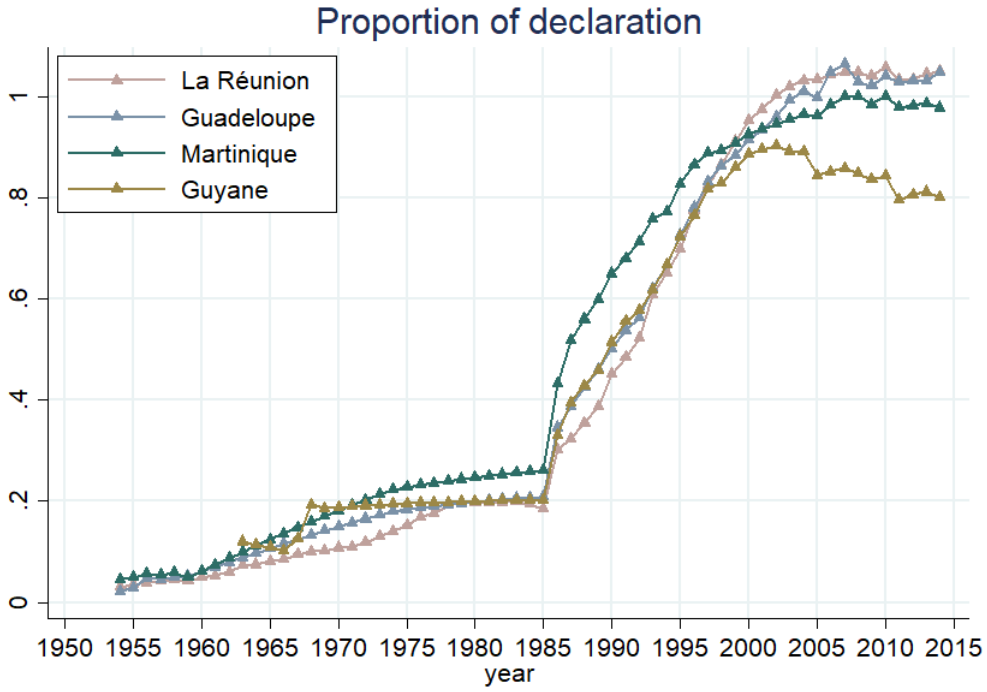


Figure B.3.5: Proportion of tax declaration (1954 - 2014)

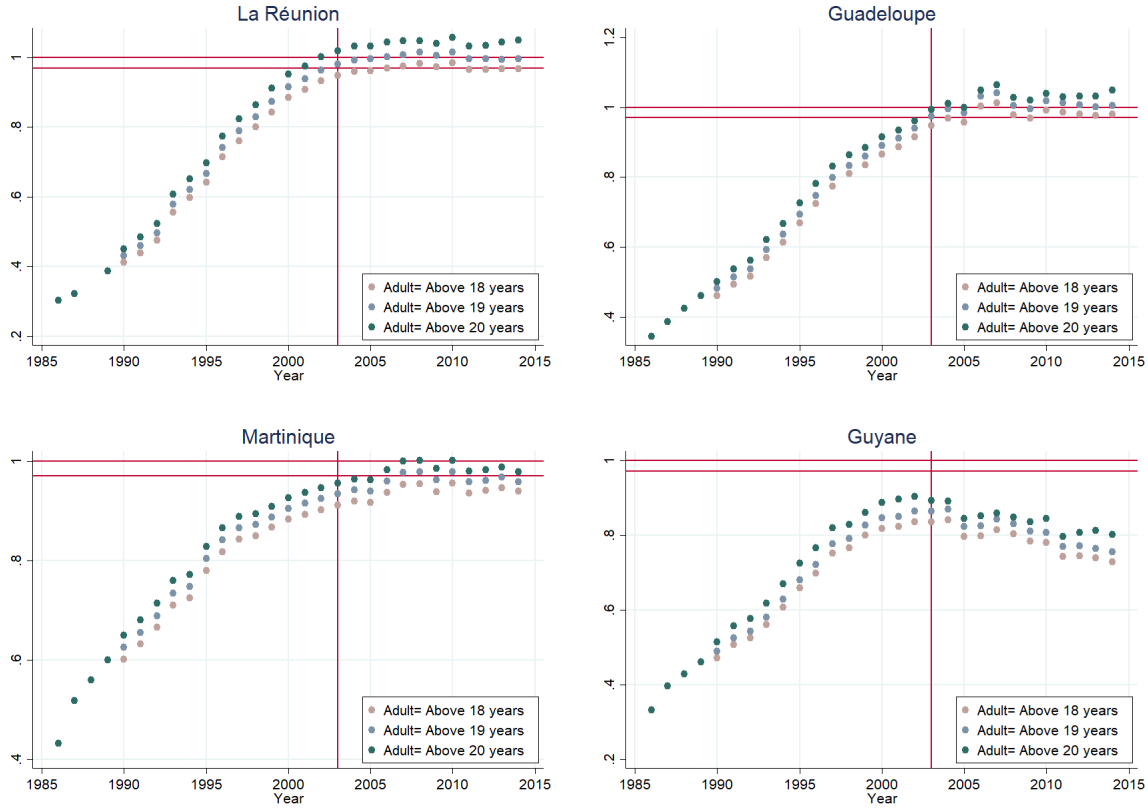


Figure B.3.6: Tax units subject to tax declarations the overseas departments (1986 - 2014)

Figure B.3.6 shows a trend in the proportion of declaration from the year 1986 in all four overseas departments. The following are estimated based on the three alternative adult population: above 18 years old, above 19 years old and above 20 years. Note that the estimates before 1990 with the alternative definition of above 18 and 19 years old are not presented here as censuses before 1990 do not provide the population by age but rather age groups (e.g. 15 to 19 years old) as explained above and would require further hypotheses to estimate the population of interest. We notice that irrespective of the definition used, there seems to be a stabilization in the proportion of declaration as from the early 2000s. During that period, approximately 100% proportion of declaration, depending on the definition chosen, is reached in La Réunion, Martinique and Guadeloupe, while Guyane reaches a maximum of 90% during that period.

La Réunion, Guadeloupe and Martinique depict more or less similar situations as far as the proportion of income tax declarations are concerned. However, the case of

Guyane seems to be a very peculiar one, as seen in figure B.3.6. There is a steady increase in the proportion of declaration, reaching around 90% at its peak at the beginning of the 2000s and there seems to be a slight decline thereafter. We argue that we never reach 100% declaration in Guyane due to the nature of the data used in the construction of the control population (the denominator of equation B.2). Population census normally documents the population living in the territory at the time of the census, without differentiating between legal and illegal residents. Guyane has had a long history of illegal migration, mostly from poorer neighboring countries. However, since there are no estimates of the share of the population within the French Guyanese territory that is illegal, there is a need to make some hypotheses.

As seen in figure B.1.9, 30% of the population in 2014 are foreigners, mostly from Suriname, Brazil and Haiti. Only a minority of asylum seekers are granted this status, accounting for 2.3% of applicants in 2009 (Baranger, 2017). Hence, the non-negligible share of the illegal population in Guyane is part of the reason for the overestimation of the number of tax units, as that population is counted in the population census but do not declare their income to the tax authorities. Moreover, the increasing and stabilizing trend observed in figure B.3.6, similar to the other departments, suggests that there is an increase in the number of tax units detected by the tax office reaching almost full declaration as from the early 2000s. This could mean that the remaining 10% that we do not observe in the recent period in Guyane are either undocumented migrants[1] (captured in the census) or population living in remote areas of Guyane (not captured by the tax office). While we cannot entirely discard the latter, the former seems to be a more important share in the Guyane context.

Hence, the evolution in the proportion of declarations over time tends to confirm our hypotheses that:

i) The definition of adult population with a threshold of 20 years of age does a fair job in estimating the total number of tax units

ii) We observe all the tax units as from 2003

Given these estimates of the control population, we then need to estimate the associated control income. The step-by-step methodology employed to estimate this control income is laid down in the next section.

[1]The Interior Ministry estimates the number of people in irregular situation to be between 30000 and 60000 persons. "Les étrangers en France", Rapport du Comité interministériel de contrôle de l'immigration, April 2014.

B.4 Control Total for Income

To estimate the share of income that accrues to the top groups, there is a need to estimate the total income that would have been declared had all the tax units been required to declare their income. In other words, there is a need to estimate the income accruing to the tax units who did not declare their income and hence who are not counted in the tax data. As explained in Section 2.3.2, there exists different methods used in the literature to construct a control total for income. In this paper, a national income approach is adopted. This implies that the total taxable income is estimated by deducting all non-taxable income and irrelevant factors (such as depreciation) from the national income or GDP of the territory.

The estimates of GDP for the overseas departments are obtained from INSEE publications. More specifically, GDP of La Réunion is obtained from INSEE-La Réunion for the period 1950 - 2014, while these estimates are obtained from publications (See Besson, 1997 and INSEE website) for the other overseas departments, covering the period 1970 - 2014. In order to have an uninterrupted series from the 1950s to 2014 for Guadeloupe, Martinique and Guyane, a relationship between the different GDP per adult population is observed during the period 1970 (1975 in the case of Guyane) to 1990 as shown in figure B.4.1.

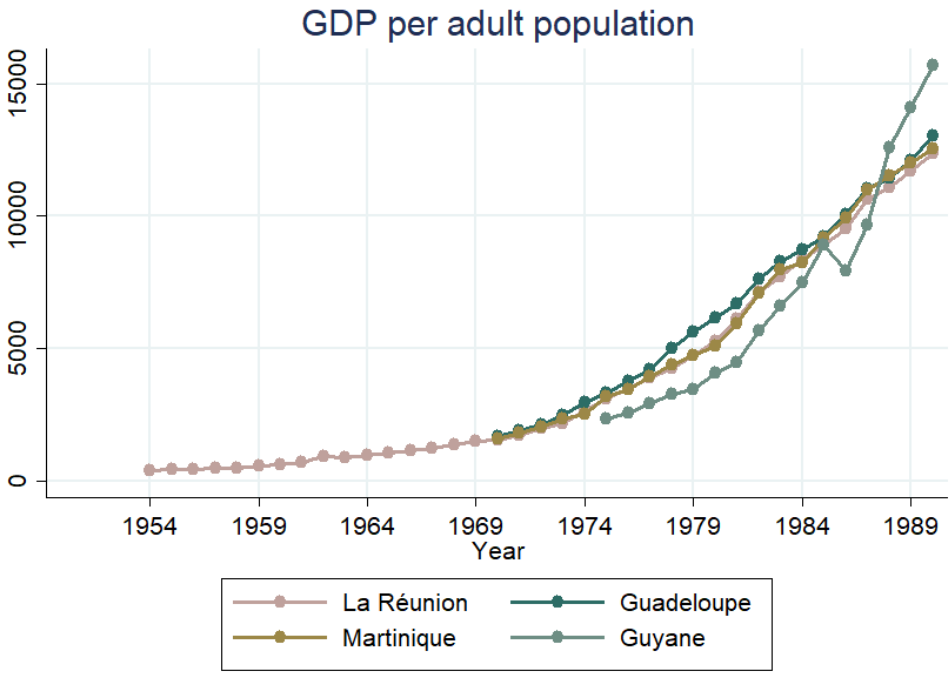


Figure B.4.1: GDP per capita (1954 - 1990)

As far as the former two departments are concerned, it seems reasonable to assume

that the GDP per adult population has been the same as La Réunion for the whole period. This assumption does not hold for Guyane, we assume a constant ratio[2] between the GDP per adult for La Réunion and that of Guyane throughout the period of 1954-1975. However, in order to estimate the non-taxable income as explained above, we also need a detailed breakdown of the national accounts. Since these are not available at the level of the overseas departments, we will rely on the taxable income series observed at the French national level to estimate its equivalent in the overseas departments.

Estimation of Taxable Income

First we establish the ratio R between average taxable income per tax unit in France, $T(tu)$ and GDP per adult population at the national level, $GDP(a)$ for the period 1950 - 2014 as follows (where i = France):

$$R_i = \frac{T(tu)_i}{GDP(a)_i} \tag{B.3}$$

From the previous section, based on the control total for population and the trends in the proportion of tax units subject to declaration, we make the assumption that everyone fills a tax form as from 2003. As a result, we can also assume that we observe the totality of the taxable income in the tax data as from the year 2003. Based on this logic, we can thus observe the ratio between taxable income per tax unit and GDP per adult population for the overseas departments for the period 2003- 2014, using equation B.4, where i = La Réunion, Guadeloupe, Martinique and Guyane. These estimations are presented in figure B.4.2.

We can observe in figure B.4.2 that on average there seems to be an approximately parallel trend between R_{fr} and the ratio for the overseas department. Note that there is a break in the series for Guadeloupe in 2007 due to the detachment of the two islands: Saint-Martin and Saint-Berthélémy. If we look at the pre-2007 and post-2007 trends separately, it is reasonable to say that they closely relate to the trend in the ratio for France. There are exceptions for some years in Guyane and Martinique but on average, it seems to fit relatively well. Based on this scenario, we assume a constant relationship between the two ratios for each overseas department, estimated as the average of the coefficient α_i over the period 2003 - 2014.[3]

[2] An average over the period.

[3] Note that for the case of Guadeloupe, we only take into account the period (t) 2003 - 2006 and for the case of Guyane, the year 2011 is excluded.

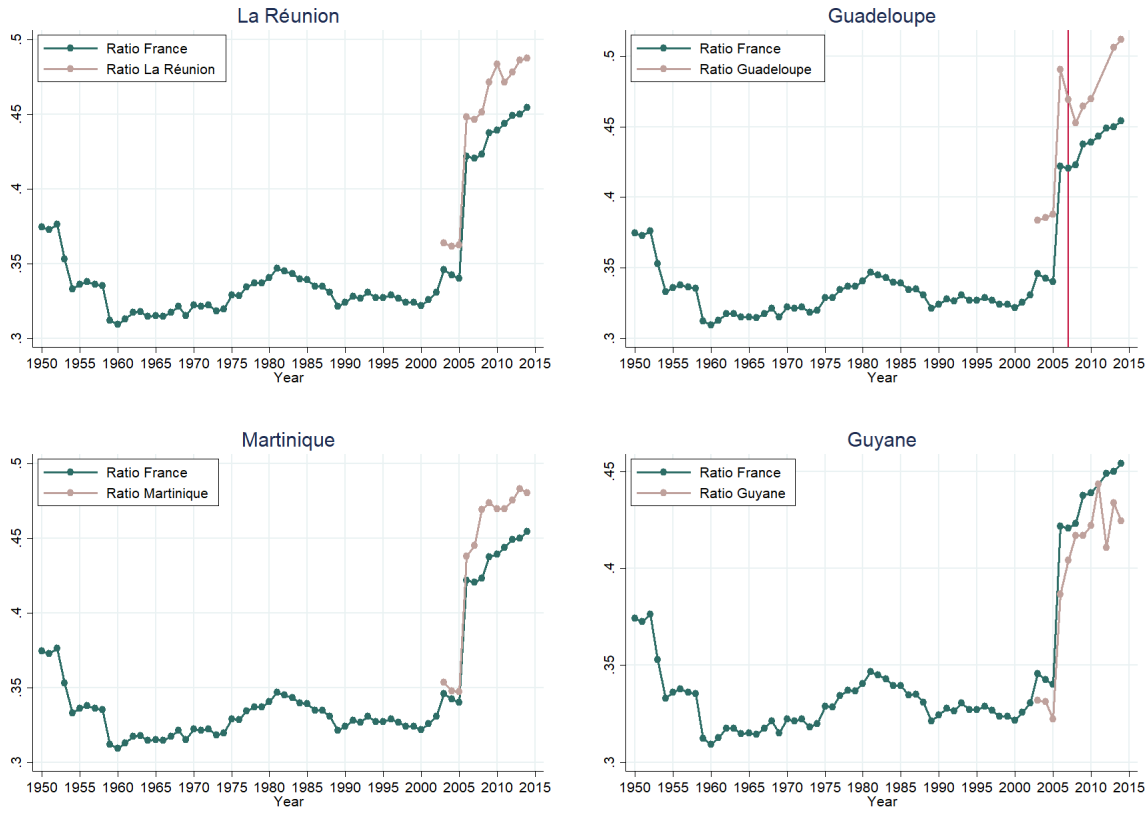


Figure B.4.2: Ratio of Taxable Income per Tax Unit to GDP per capita in France and the overseas departments (1950 - 2014)

$$\overline{\alpha_i} = \frac{\sum_{t=2003}^{2014} \frac{R_{fr,t}}{R_{i,t}}}{n} \tag{B.4}$$

i = La Réunion, Guadeloupe, Martinique and Guyane; t= 2003 - 2014 for La Réunion, Martinique and Guyane (excluding 2011) and t=2003-2006 for Guadeloupe and n = number of years.

Given α_i, we can estimate the ratio between average taxable income and GDP per adult for the period 1950 - 2002 for the overseas department based on the series of France, as follows:

$$R_{i,t} = \alpha_i \times R_{fr,t} \tag{B.5}$$

The estimation of this ratio for the entire period is presented in figure B.4.3.

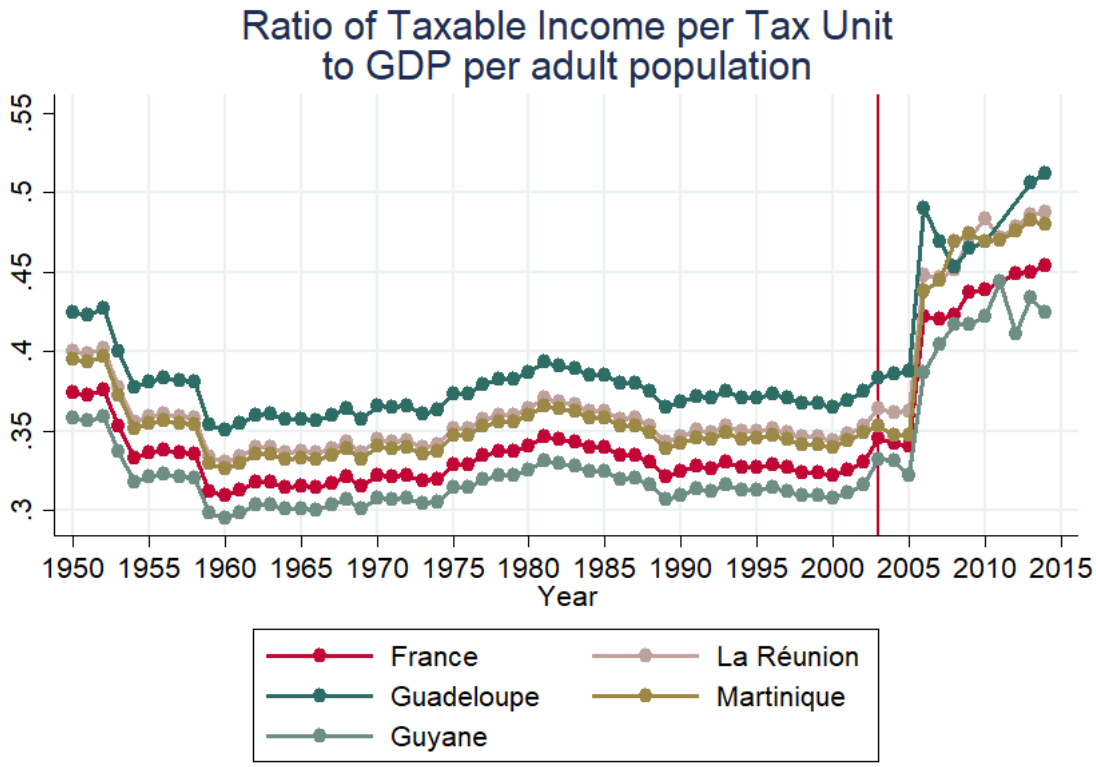


Figure B.4.3: Taxable Income per tax units to GDP per capita (1950 - 2014)

Having estimated this ratio, an uninterrupted series for total taxable income, and hence control total for income, can be computed for the whole period.The total taxable income and the average taxable income per tax units for the overseas departments are presented in figures B.4.4 and reffig:tot$_a$v$g_t$ax1.

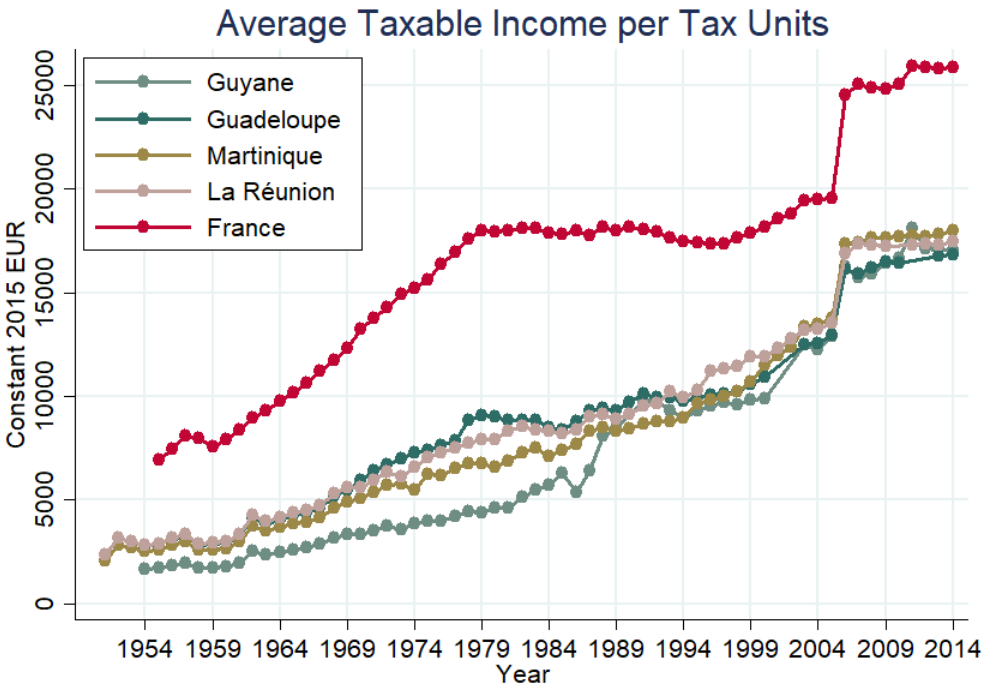


Figure B.4.5: Average Taxable income (1950 - 2014)

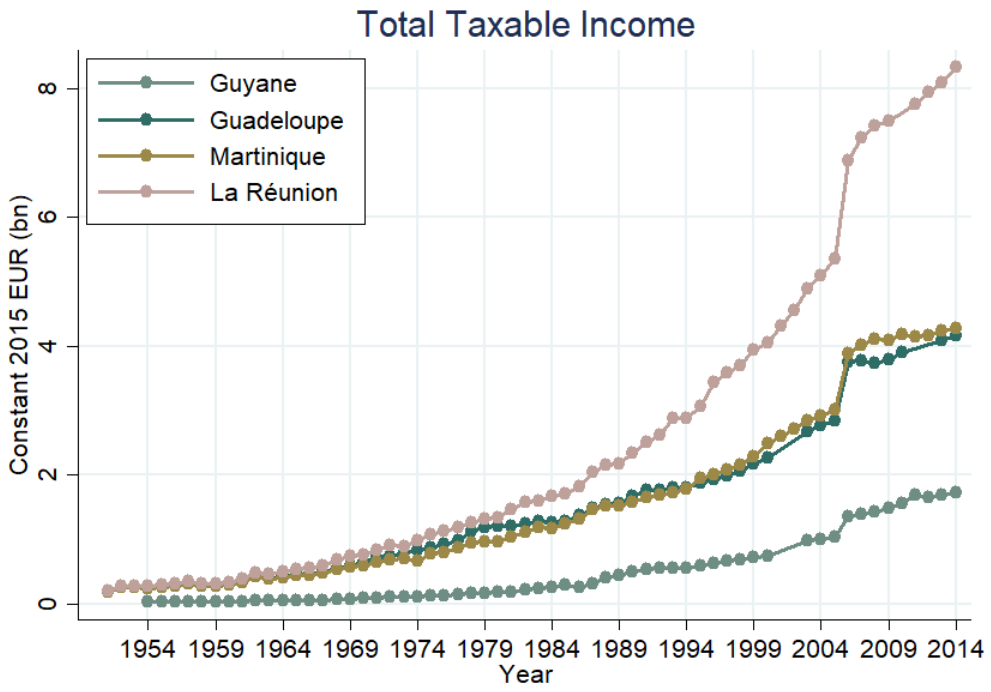


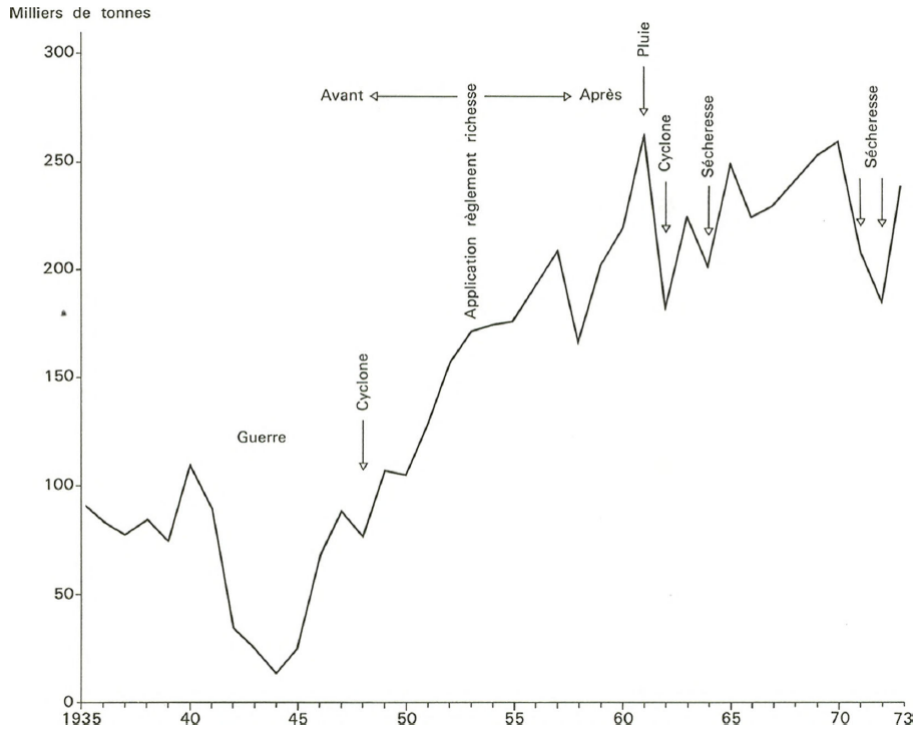
Figure B.4.4: Total taxable income (1950 - 2014)

Estimation of Fiscal Income

As explained in section 2.3.2, the income reported by the tax authorities are taxable income- which is fiscal income deducting allowances. As the rules for allowances changes over time, we would like to look at fiscal income instead. In order to go from taxable income to fiscal income, various corrections have to be made to the series. The corrections made here follow the ones in Garbinti, Goupille-Lebret, and Piketty (2018). For a more detailed explanation of these corrections, please refer to DINA Appendix D.2 of that paper and Piketty (2001). We apply the same correction factors as used in the series for France. These include an upgrade rate due to previous-year-tax deductions and other types of deductions, such as the lump sum deductions for wage earners. In 2006, the 20% deductions for additional professional expenses was repealed and is accounted for in the corrections factors. Similarly, we assume the same aggregate taxable income to fiscal income ratio as in Garbinti, Goupille-Lebret, and Piketty (2018).

B.5 Economic Situation

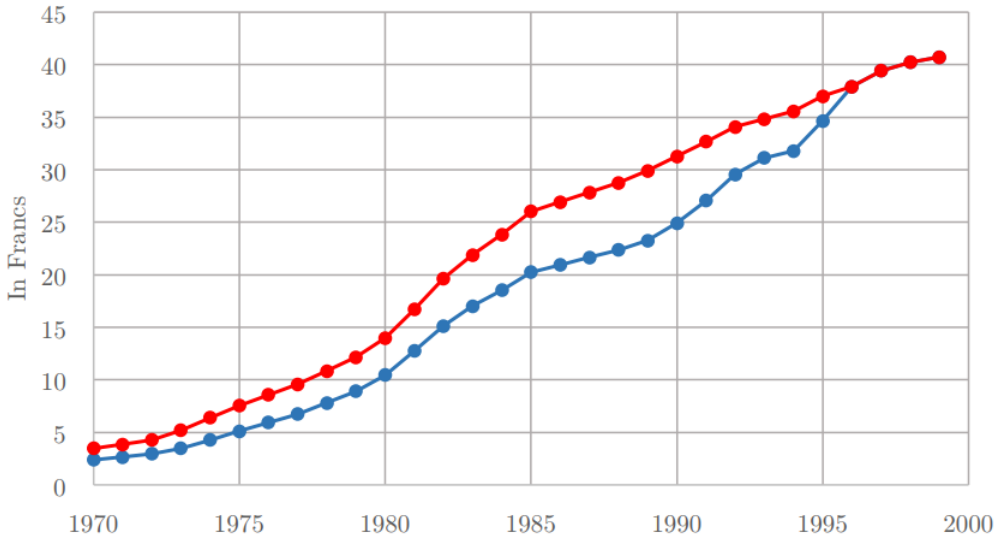
1. **Sugar production**



Source: INSEE

Figure B.5.1: Sugar production in La Réunion (1935 - 1973)

2. Minimum Wage



Source: INSEE. Note: The overseas departments are in blue and metropolitan France in red

Figure B.5.2: Evolution of hourly minimum wage (1970 - 2000)

3. **Wage density distribution in the public and private sector in La Réunion in 1988**

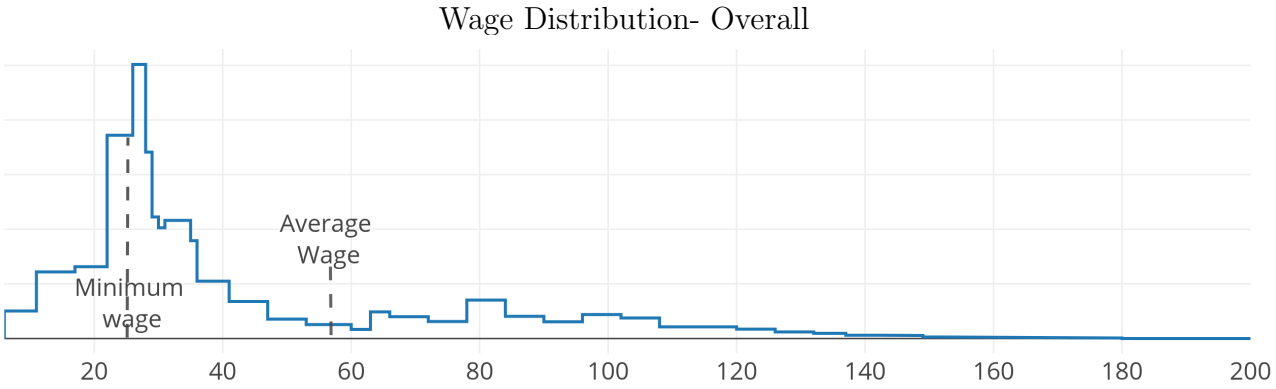


Figure B.5.3: Overall wage distribution in La Réunion (1988)

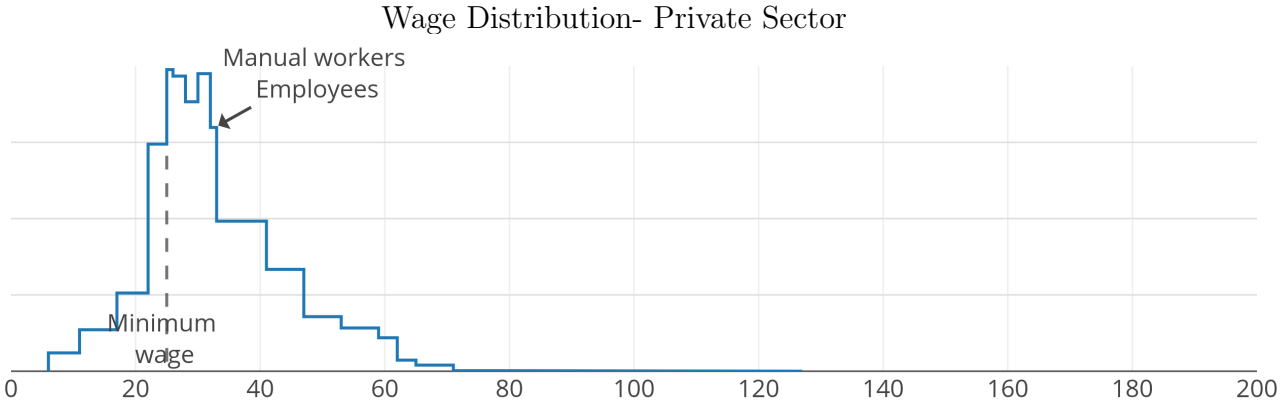
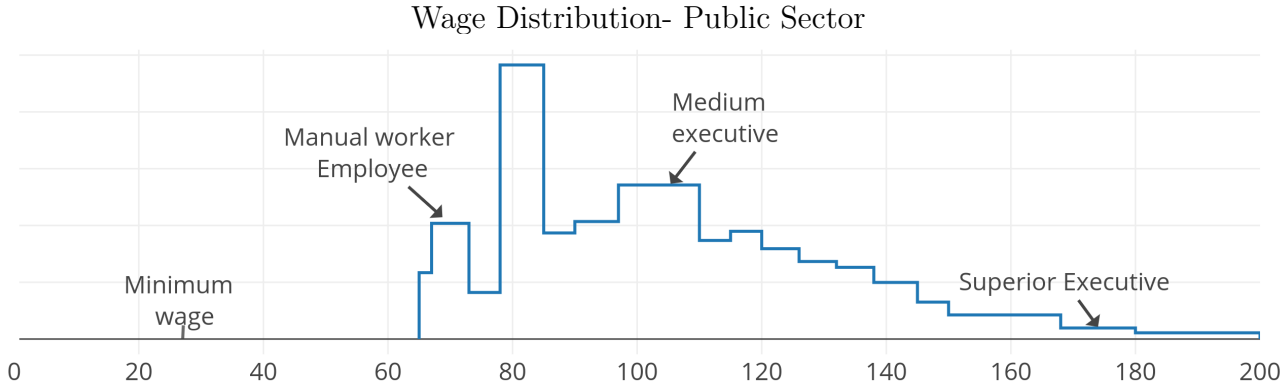


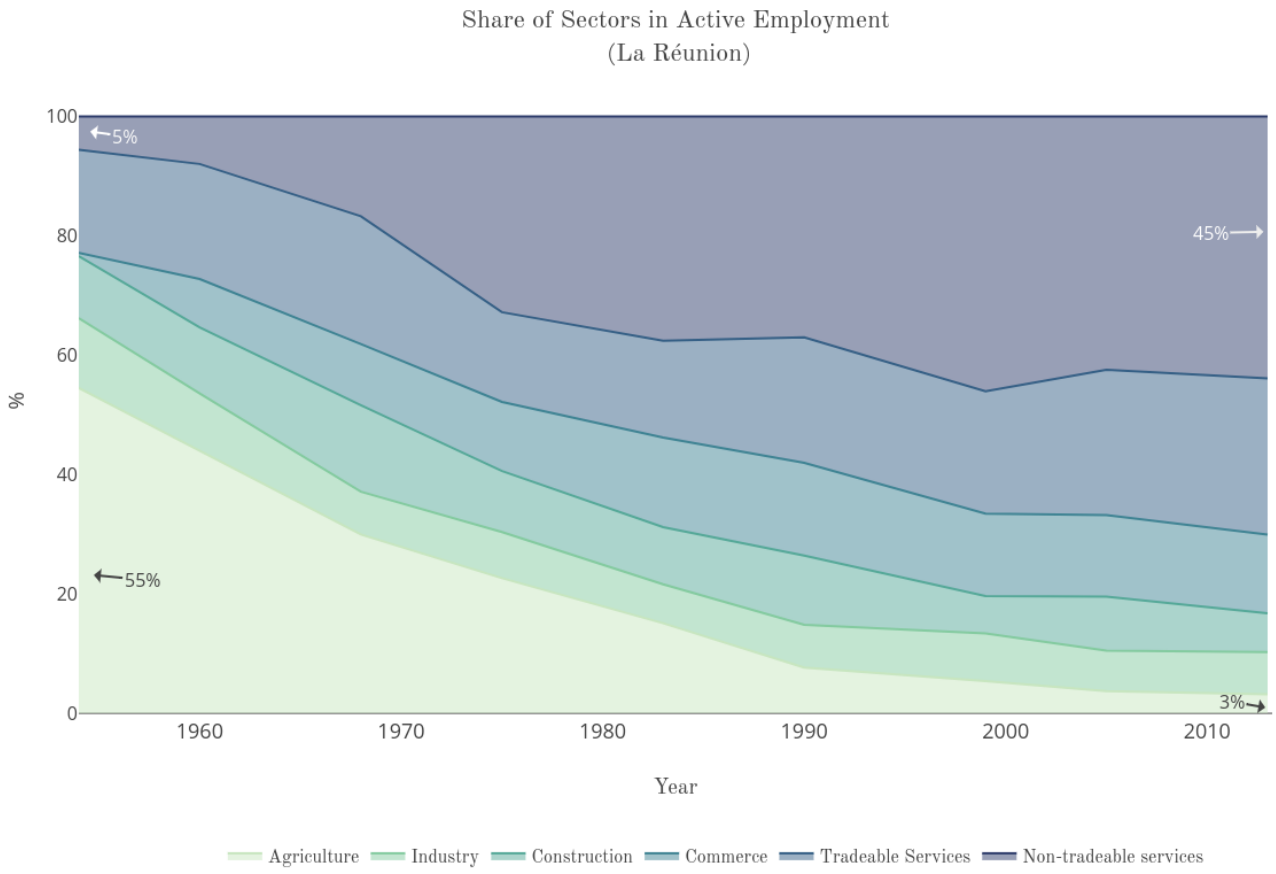
Figure B.5.4: Private sector wage distribution in La Réunion (1988)



Source: INSEE

Figure B.5.5: Public sector wage in La Réunion (1988)

4. **Share of sectors in Active Employment (La Réunion)**



Source: INSEE

Figure B.5.6: Share of sectors in Active Employment in La Réunion 1954 - 2014

5. **Illiteracy rate**

	1954	**1961**	**1967**
La Réunion	60,6%	49,4%	39,0%
Guadeloupe	34,5%	22,1%	15,4%
Martinique	25,5%	15,2%	

Source: Population Census

Table B.5.1: Illiteracy Rates in the overseas departments in 1954, 1961, and 1967

B.6 Native-Metropolitan Divide

In section 2.6, I use the Echantillon Démographique Permanent (EDP), a rich administrative dataset established in 1967 by INSEE. It comprises of historical census and registry office data on individuals born on certain dates of the year[4]. The overseas departments were added in this panel since 2004 and fiscal data were incorporated as from 2011. In this paper, I exploit the population census data matched with the fiscal data in the four overseas departments in 2014.

To begin with, I only keep "EDP individuals" who are fiscal residents of the overseas department[5]. This EDP sample represents around 4% of the total population. I opt for the fiscal data of 2014 and only keep the primary declaration of adults who declare their income to the tax office[6]. The annual french population census is carried out on a succession over a five-year period. Since the demographic variables of interests are mostly time-invariant, the fiscal data is matched over five years of population census: the two preceding and two subsequent years of 2014 (2012, 2013, 2014, 2015, 2016) in order to minimise the possibility of error. I further excludes individual of less than 25 years of age in 2014 as they are likely to have varying educational and professional status over the five-year period.

Table B.6.1 shows the descriptive statistics of the resulting sample and table **??** to B.6.2 are the regression results with different definitions of income and sample restrictions.

[4]As from 2004, individuals born on 16 dates of the year

[5]The database also have some information on the other individuals in the household where a person is born on the 16 dates of the year.

[6]This excludes individuals (mostly students) who only declare housing taxes.

	Native	Metropolitan	Overall
Observations	14103	2205	16308
	86%	14%	
Departments			
Guadeloupe	88%	12%	25%
Martinique	90%	10%	27%
Guyane	72%	28%	6%
La Réunion	85%	15%	41%
Demographic Characteristics			
Female (%)	57%	51%	56%
Married (%)	43%	33%	41%
Age	54.6	46.6	53.5
Years of schooling	9.63	13.86	10.2
Labour market Status			
Active (%)	54%	74%	56%
Employed (%)	62%	80%	66%
Sector of employment			
Public (%)	44%	57%	47%
Private (%)	51%	35%	48%
Self-employed(%)	5%	9%	6%
Average Income (euros)			
Salary	24 961.69	36 470.13	27 389.70
Salaried & self-emp	25 277.76	37 743.59	27 999.73
Salaried &Self-emp (incl. unemp)	17 659.83	32 435.61	20 294.80
Salaried, Selfemp & Retirement (incl. unemp)	14 271.48	29 330.07	16 307.98

Table B.6.1: Demographic and Economic characteristics of natives and metropolitans in the overseas departments

B.6.1 Regression Results

Dependent variable: Salaried Income, Self-employment earnings (incl. Unemployment benefits) and Retirement pensions

	1	2	3	4	5	6	7	8
Metropolitan	14886.1***	13391.8***	7228.6***	7412.1***	6423.5***	5900.5***	5582.5***	5591.0***
	(387.9)	(384.6)	(357.3)	(344.5)	(312.8)	(293.4)	(292.3)	(292.0)
School Years			1938.9***	1741.5***	1405.9***	1223.9***	1199.4***	1196.3***
			(30.37)	(29.81)	(27.61)	(26.16)	(26.15)	(26.13)
Active				10395.0***	-1220.9***	-1425.9***	-1892.7***	-2009.3***
				(295.3)	(331.6)	(310.8)	(312.9)	(313.2)
Employed					17457.6***	5304.3***	4013.9***	2548.6***
					(293.9)	(376.1)	(384.2)	(458.3)
Full-time						17256.2***	17595.6***	16353.2***
						(363.6)	(365.1)	(422.1)
Public Sector							2147.8***	2282.8***
							(279.8)	(280.5)
Self-employment							8997.4***	11196.5***
							(664.8)	(763.1)
Permanent								2882.6***
								(492.8)
Constant	14541.1***	24607.0***	-9147.3***	-23151.5***	-17035.5***	-14930.8***	-14334.0***	-14108.9***
	(265.8)	(631.0)	(773.3)	(845.0)	(773.0)	(726.0)	(724.2)	(724.5)
Observations	16308	16308	16308	16308	16308	16308	16308	16308
Adj R2	0.0920	0.146	0.317	0.365	0.478	0.541	0.547	0.548
Dept FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard Errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. The sample of 16308 observations include the adult population (more than 25 years of age)

Table B.6.2: Metropolitan premium: with and without controls

Appendix C

Appendix to "Land Inequality in the Developing World"

C.1 Sources of Data

Data sources census and surveys

Country	Type	Year	Description	Source
AFRICA				
Burkina Faso	Census	2010	FAO 2010	FAO
	Survey	2014	Enquete Multisectorielle Continue (EMC-BF) 2014	LSMS (World Bank)
Egypt	Census	2009	FAO World Census of Agriculture 2010	FAO
	Survey	2018	Egypt Labor Market Panel Survey	Egypt's Central Agency for Public Mobilization and Statistics
Ethiopia	Census	2000	FAO World Census of Agriculture 2000	FAO
	Survey	2015	Ethiopia Socioeconomic Survey (ESS)- Wave 3 (2015-16)	LSMS (World Bank)
Gambia	Survey	2015	Gambia Integrated Household Survey (IHS) 2015	The Gambia Bureau of Statistics
Malawi	Census	2006	FAO World Census of Agriculture 2010	FAO
	Survey	2016	Fourth Integrated Household Survey (IHS4) 2016	LSMS (World Bank)
Niger	Survey	2014	National Survey on Household Living Conditions and Agriculture 2014	LSMS (World Bank)
Nigeria	Survey	2015	General Household Survey, Panel 2015-2016	LSMS (World Bank)
Tanzania	Census	2007	FAO World Census of Agriculture 2010	FAO
	Survey	2014	National Panel Survey (NPS) 2014-15	LSMS (World Bank)
ASIA				
Bangladesh	Census	2008	FAO World Census of Agriculture 2010	FAO
	Survey	2011	Bangladesh Integrated Household Survey (BIHS) 2011	International Food Policy Research Institute (IFPRI)
China	Census	2002	ILC	Khan 2001
	Survey	2002	Chinese Household Income Project	Chinese Academy of Sciences and others
	Survey	2016	China Family Panel Studies (CFPS) 2016	Institute of Social Science Survey, Peking University
India	Census	2010	FAO World Census of Agriculture 2010	FAO
	Survey	2012	All India Debt and Investment Survey (AIDIS), 2012	Ministry of Statistics and Programme Implementation
Indonesia	Census	2013	FAO World Census of Agriculture 2010	FAO
	Survey	2014	Indonesia Family Life Survey	
Pakistan	Census	2010	FAO World Census of Agriculture 2010	FAO
	Survey	2010	Pakistan Household Integrated Survey (HIES), 2010-11	Pakistan Bureau of Statistics
Vietnam	Census	2011	FAO World Census of Agriculture 2010	FAO
	Survey	2014	Vietnam Household Living Standards Survey (VHLSS), 2014	General Statistical Office (GSO) of Vietnam
LATIN AMERICA				
Brazil	Census	1996	FAO World Census of Agriculture 2000	FAO
	Survey	1997	Pesquisa sobre padrões de vida 1996-1997	LSMS (World Bank)
Ecuador	Census	2000	FAO World Census of Agriculture 2000	FAO
	Survey	2014	Ecuador Living Conditions Survey	National Statistical Office of Ecuador
Guatemala	Census	2003	FAO World Census of Agriculture 2000	FAO
	Survey	2000	Encuesta Nacional sobre Condiciones de Vida 2000	LSMS (World Bank)
Mexico	Census	2007	FAO World Census of Agriculture 2010	FAO
	Survey	2009	Mexican Family Life Survey	UIA and CIDE
Nicaragua	Census	2011	FAO World Census of Agriculture 2010	FAO
	Survey	2014	Encuesta Nacional Sobre Medicion de Nivel de Vida	National Statistical Office of Nicaragua
Panama	Census	2011	FAO World Census of Agriculture 2010	FAO
	Survey	2008	Encuesta de Niveles de Vida	LSMS (World Bank)
Peru	Census	2012	FAO World Census of Agriculture 2010	FAO
	Survey	2007	Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza 2007	National Statistical Office of Peru

Table C.1.1: Land Inequality Data Sources

C.2 Results

Country	Bottom 50%	Middle 40%	Top 10%	Gini index
India (2012)	0,5%	27,4%	72,1%	0,84
Bangladesh (2015)	0,0%	31,5%	68,5%	0,84
Pakistan (2010)	1,4%	32,2%	66,4%	0,80
Indonesia (2014)	2,8%	39,5%	57,7%	0,74
China (2012)	10,0%	38,4%	51,5%	0,64
Vietnam (2014)	7,0%	41,1%	51,9%	0,68
Ecuador (2014)	1,1%	28,6%	70,3%	0,82
Brazil (1997)	0,1%	13,1%	86,7%	0,91
Guatemala (2000)	0,0%	22,7%	77,3%	0,88
Nicaragua (2014)	0,3%	26,8%	72,9%	0,85
Panama (2008)	1,3%	13,4%	85,4%	0,90
Ethiopia (2015)	0,4%	28,4%	71,2%	0,83
Gambia (2015)	2,8%	42,5%	54,7%	0,73
Malawi (2016)	5,3%	37,3%	57,4%	0,72
Niger (2014)	0,3%	44,6%	55,2%	0,75
Nigeria (2015)	5,3%	39,1%	55,6%	0,71
Tanzania (2015)	3,3%	32,3%	64,4%	0,77

Note: Distribution of agricultural land value, including the landless population

Table C.2.1: Agricultural land distribution

C.3 DHS and LSMS comparison

As part of the effort by the World Bank, the Living Standards Measurement Surveys (LSMS) has been implemented in a number of countries, aiming at providing nationally-representative household surveys and in some countries, with a panel component. The coverage of the LSMS is particularly wide in Africa compared to other world regions, providing detailed information on agricultural activities and land. Since the focus of these surveys often aim at capturing agricultural activities, they cover both land operated or owned by households. In surveys in which the distinction between the two are not straightforward, a proxy for ownership is defined as individuals who have inherited or purchased land. As a robustness check, the Demographic Health Surveys (DHS) are used. These are nationally-representative

household surveys that focus on health and nutrition aspects but also have basic information on land ownership since the 2000s, reporting whether a given household owns or not land, and the area of the land owned. Gini coefficients estimated from the LSMS and DHS being very similar, validating the ownership proxy of the LSMS (Figure C.3.1).

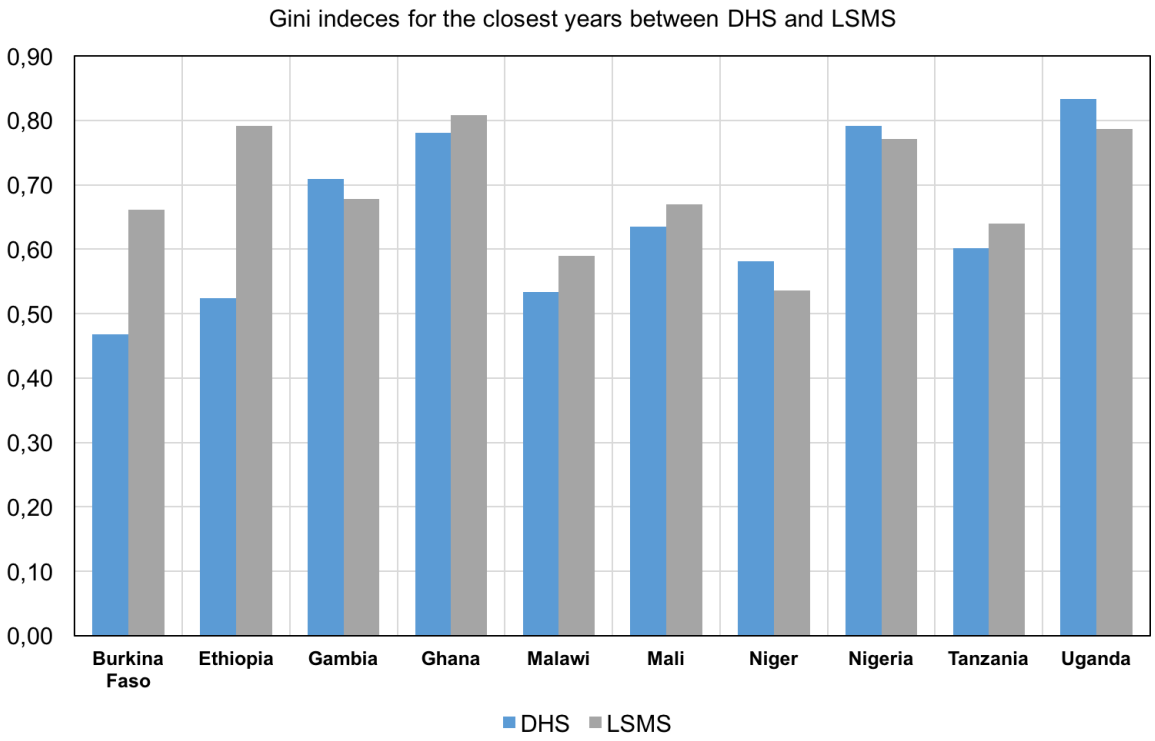


Figure C.3.1: Gini indices among owners based on DHS and LSMS

C.4 Ownership Regime

In order to construct commensurate land inequality estimates across various countries, and more importantly, regions of the world, differences in land ownership regimes must be taken into account. Ideally, all surveys would contain a question directly asking, "The household enjoys which rights over this plot of land?", followed by an extensive list of all possible property rights (i.e. right to keep profits from land, right to transfer land, right to sell land or use as collateral, etc.) which exist in each country. This would allow for the creation of sub-samples of households which have exactly equivalent ownership regimes across countries which may have different institutional setups, resulting in inequality measures which are systematically comparable. In the absence of such detailed data, it is necessary to understand the exact land rights that households used in our analysis actually possess.

With this goal, we identify the key rights associated with the economic benefits of land ownership. First, landownership involves the right to use the land for grazing animals, cultivating crops or other agricultural activities, what we call a *use right*. Second, a land-owning household has the right to alter the land and transform it, as well as profit directly from its productive capacities, what we call a *control right*. Third, ownership rights may include the right to transfer the land (and the corresponding ownership rights associated to it) to a third party, in any capacity. This right may be broken up into transfers within the household (through, for example, inheritance), what we call a *bequest transfer right*, and transfers between households (through sales, leases or as collateral), what we call a *full transfer right*. Fourth, and finally, a household may have the most secure property right, that of a fully transferable, socially recognized property title, securing all the previous rights in in order to provide protection from seizure of land asset by a third party (including the state), what we call *titled right*.

Having delineated these rights of land ownership, we move on to applying this framework to our results. We can use, as a first example, the use and control rights. We note that all landowning households in all of our surveys hold these rights, no matter the institutional set-up. A household with land in, say, Guatemala, just like a household with land in, say, China, enjoys the right to use the land they possess and profit from it. In this respect, our results are completely comparable across all countries under analysis. Stated differently, with respect to use and control right–ownership inequality, our inequality estimates are completely equivalent across regions. However, if we were to compare, for example, titled right-ownership inequality, our results, as they stand, are not immediately comparable. This is because, while households in Guatemala might have access or even possess, regularly defined property titles, which permit the sale of a land asset, households in China, due to its idiosyncratic institutional set-up, do not[1]. For this reason, with this strictest property right, we are only able to compare a sub-sample of the countries and households.

The baseline definition of ownership in this paper – that is – the definition which covers all countries is one where households with land enjoy three rights: use, con-

[1]For example, land in China has, for the around the last 70 years, been owned collectively, and administered by the state in a registration system denominated hukou. Farmers have long-term contracts that may be renewed and passed on within a household, but may not be sold to other farmers for profit. For example, see Vendryes, 2010 for more details on particularities of land tenure in China.

trol, and bequest land rights. This implies that all households which report being landholders in our analysis fully control and profit from their land, as well as possess the right to pass said land down within the household. [2]. Consequently, with respect to this definition, which we call *private household* ownership, all the inequality shares reported are comparable across all countries and regions. We can conceptually understand some stricter definitions compared to this baseline definition. *Marketable private household* ownership refers to land ownership including all previous rights as well as existence of and legal right to sell, lease, or use as collateral. Lastly, *titled private household* ownership designates all previous rights, as well as the possession of a legally recognized land title protecting from all expropriation of land by a third party. The coverage of countries and households for each definition can be found in Table C.4.1. In future versions of this paper, inequality shares will be estimated using the last two definitions discussed.

Definition of Ownership (Including which land property rights)	*Subsample considered*
Private Household (Benchmark Definition): • Use • Control • Bequest Transfer	All countries, all households reporting land ownership
Marketable Private Household: • Use • Control • Full Transfer	All countries except China and Vietnam. Coverage within certain African countries will decrease.
Titled Private Household: • Use • Control • Full Transfer • Titled	All countries except China and Vietnam. Coverage within all countries will decrease even more as we consider only households with full property titles.

Table C.4.1: Coverage and sample according to different ownership definitions

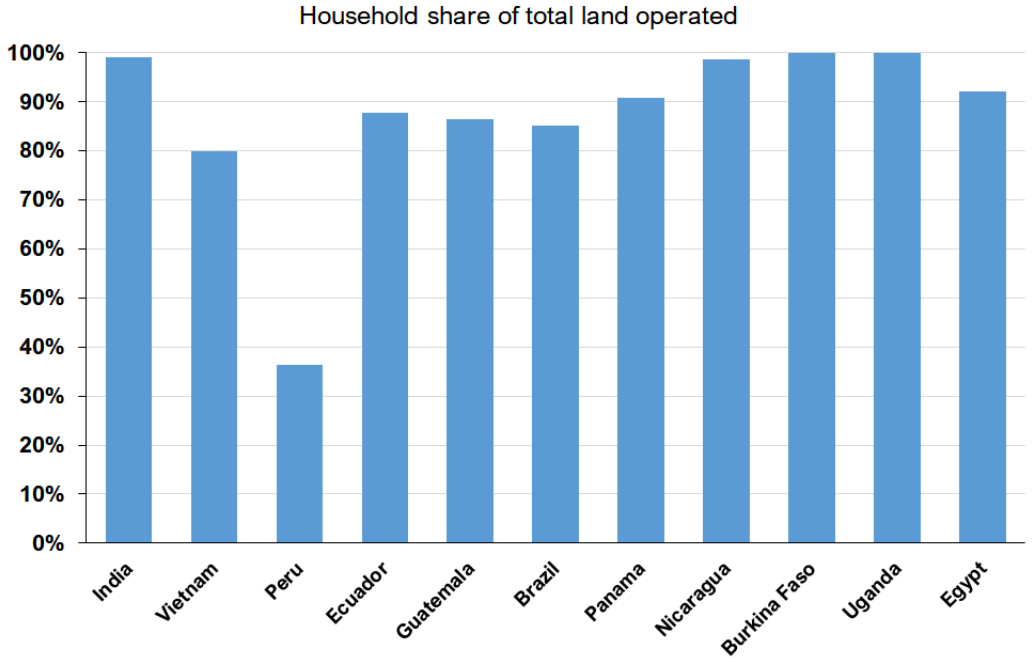
[2]Vietnam, for example, whose land is legally owned by the state, confers to all landholders the rights to fully exploit land as well as freely transfer it (Ravallion and Van de Walle, 2008). For less clear cases, such as for African countries with varying institutional contexts, the households designated as landowners only include households stating land as "purchased" or "inherited".

C.5 Additional figures and tables

Region	Asia	Africa	South America & Caribbean	Europe	North America & Oceania
Average	85%	89%	81%	53%	64%
Median	91%	89%	85%	57%	64%
Standard Deviation	12%	7%	15%	19%	2%
No of observations	11	6	13	34	3

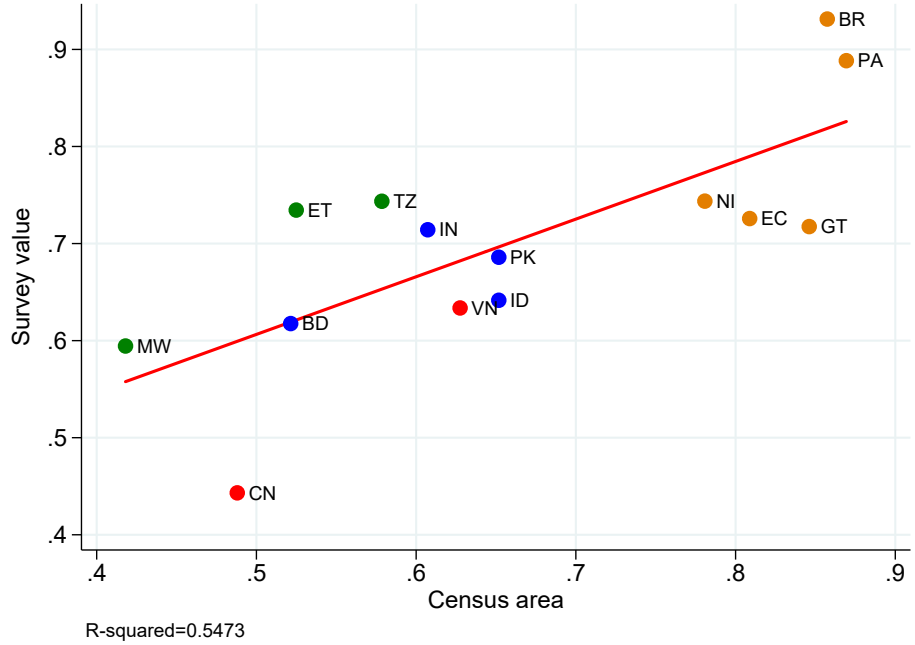
Note: This table shows the percentage of total agricultural land that is operated by households.
Source: FAO World Programme for the Census of Agriculture of 2010.

Table C.5.1: Share of operated land owned by landholders



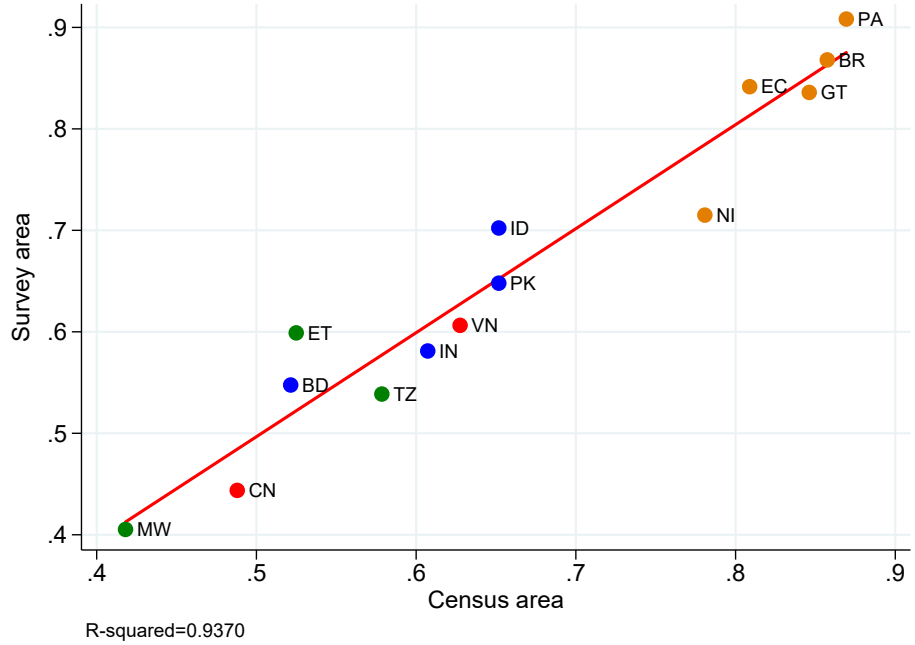
Note: This table shows the percentage of total agricultural land that is operated by households.
Source: FAO World Programme for the Census of Agriculture of 2010.

Figure C.5.1: Share of agricultural land operated by households



The countries and the year of survey are as follows: BD – Bangladesh (2011); BR – Brazil (1996); CN – China (2002); EC – Ecuador (2014); ET – Ethiopia (2011); GT – Guatemala (2000); ID – Indonesia (2014); IN – India (2012); MW – Malawi (2010); NI – Nicaragua (2014); PA – Panama (2008); PK – Pakistan (2010); TZ – Tanzania (2018); VN – Vietnam (2014). For sources of data, see Appendix A.

Figure C.5.2: Gini index in selected countries: land area from census data vs. land value from survey data



Note: this figure compares the Gini index for land area from census and survey data in the sample of countries included in figure appendix E2. The countries and the year of survey are as follows: BD – Bangladesh (2011); BR – Brazil (1996); CN – China (2002); EC – Ecuador (2014); ET – Ethiopia (2011); GT – Guatemala (2000); ID – Indonesia (2014); IN – India (2012); MW – Malawi (2010); NI – Nicaragua (2014); PA – Panama (2008); PK – Pakistan (2010); TZ – Tanzania (2018); VN – Vietnam (2014). For sources of data, see Appendix A.

Figure C.5.3: Gini index in selected countries: land area from census data vs. land area from survey data (in countries for which land value from surveys is also available)