UNIVERSITAT DE BARCELONA

Essays on Urban Economics

Ghizlen Ouasbaa Azzouani

PhD in Economics

Essays on Urban Economics

Ghizlen Ouasbaa Azzouani

PhD in Economics | **Ghizlen Ouasbaa Azzouani**

2023

UNIVERSITAT DE BARCELONA

PhD in Economics

Thesis title:

Essays on Urban Economics

PhD candidate:

Ghizlen Ouasbaa Azzouani

Advisors:

Albert Solé-Ollé

Elisabet Viladecans-Marsal

Date:

December 2023

UNIVERSITAT DE BARCELONA

Acknowledgments

I cannot imagine myself writing this dissertation without the assistance and support of many people. I want to express my gratitude to all of them for providing me with the best possible academic and personal support.

First and foremost, I would like to thank my supervisors, Albert Solé-Ollé and Elisabet Viladecans Marsal. I am deeply grateful to them for guiding me and generously providing me with their valuable advice during my doctoral studies. Their dedication and insightful comments made me capable of delivering this thesis successfully. I would also like to thank them for their patience and confidence in me. My gratitude is also extended to Marie Aurélie Lapierre with whom I had the pleasure to collaborate in the writing of the last chapter of this thesis. I have enjoyed working together, and have learned so much. Thank you for being so patient.

I am also very grateful to both the Institut d'Economia de Barcelona (IEB) and the Department d'Economia de la Universitat de Barcelona (UB). These institutions have provided me with an environment in which to conduct research that is friendly, cozy, and enjoyable. Within the IEB, I particularly want to thank Susana and M. Angels and within UB, Jordi Roca, for much more than their administrative assistance. They made my life easy at the most difficult times. I deeply appreciate their professional support at every step of the process.

The department provided me with wonderful professors with whom to share insights and learn. Pilar, Judit, Jordi, Daniel, Matteo, Javier, and Andreu, thank you all for your continuous support and genuine interest during these years. Thanks also to José and José Maria.

Part of the work carried out during these years has been done while visiting the Department of Economics at the University of Paris Saclay. I am deeply grateful to Miren, José, Alex and Théo for making my research stay truly worthwhile.

Since the day I started my PhD, I have had the opportunity to meet some amazing and hard-working colleagues, without whom these years would have been really tough. A special mention to Pierre, Rodrigo, Magdalena, Alessio, and Filippo, you welcomed me and made sure to help me find my path from the beginning. Thank you for all your time and advice.

A very special note of gratitude goes to Claudia, Lili, Marie, Gerard, Candan, Marianna, Kaan, Aléx and the rest of the guys who have walked this path with me every step of the way. Thank you for being a lifesaver. Your friendship is one of the best things I take with me over these years.

Finally, some extraordinary people need to be mentioned here. First, my parents. Mom and Dad, you have given me the resources and the opportunity to develop into the person I am now. Thank you for your endless support and encouragement, and for that, I will be always grateful. Then, my warm-hearted sisters, I feel incredibly fortunate to have you.

Last but not least I want to thank Reda. Since day one you have been a huge support, you have helped me to get my priorities straight and filled my days with joy and laughter. Thank you for being my safe space.

And finally, to Inara, your presence in my life has enriched it in ways you may never fully comprehend.

Barcelona, December 18th, 2023.

Contents

1. Introduction

Urban agglomerations lie at the heart of modern economies, with cities accounting for less than 1% of land use globally but the vast majority of economic activity. Cities are home to a wide range of economic activities and social groups creating a heterogeneous environment. As Glaeser (2011) argues in his book "Triumph of the City," cities are the engines of economic growth and innovation. Cities provide access to a wider range of jobs, educational institutions, and cultural attractions. They offer residents the opportunity to connect with people from all walks of life and to learn from their diverse experiences. According to the World Economic Forum, since 1950, world's population living in cities has risen almost six-fold.[1] This unprecedented surge in urbanization reflects not only the allure of urban centers as hubs of opportunity and innovation but also the immense challenges that come with the concentration of people and activities.

Housing and tourism are two key factors that shape city development and significantly impact the lives of residents. Firstly, housing, as a household's primary asset and a major determinant of financial stability, has seen a sharp increase in prices, raising concerns about access to affordable housing and its broader economic implications (Causa et al., 2019). Over the past two decades, house prices in OECD countries have surged by approximately 46%, exacerbating the difficulty of finding affordable housing in desirable neighborhoods and contributing to the widening gap in inequality levels in developed countries (Piketty, 2018).

Housing affordability is a major societal challenge with significant impacts on households and the economy. Households struggling to afford housing may have difficulty accessing essential needs such as food, healthcare, and education (Wetzstein, 2017). At the macro level, housing affordability can stifle economic growth by misallocating labor and capital (Brueckner and Sridhar, 2012; Hsieh and Moretti, 2019; Parkhomenko, 2020). This has been evidenced in the United States (Herkenhoff et al., 2018; Hsieh and Moretti, 2019) and urban areas in Asia (Helble et al., 2021). Therefore, understanding the causes and consequences of housing affordability is crucial for addressing this pressing societal issue.

[1] https://www.weforum.org/agenda/2019/09/mapped-the-dramatic-global-rise-of-urbanization-1950-2020/

In parallel, cities and their development have undoubtedly been profoundly influenced by tourism over the past decade. Reports from the World Travel Monitor indicate that city tourism is one of the most rapidly expanding segments in the travel industry, with its evolving nature having increasingly noticeable effects in numerous cities.[2] As a critical sector of the global economy, tourism makes a substantial contribution to both GDP and employment. The World Travel and Tourism Council (WTTC) reports that tourism is responsible for creating one in every four new jobs globally. This translates to 10.3% of all jobs, which amounts to 333 million positions, and also contributes to 10.3% of the global GDP.

Despite its importance, evidence regarding tourism's impact on economic development is limited and inconclusive. While tourism can stimulate structural economic changes through sectoral redistribution of economic activities, creation of employment opportunities, and attracting foreign exchange and local investments (Song et al., 2012), other studies have pointed out its negative consequences. These include reduced productivity, potential displacement of technologically advancing industries, prevalence of temporary and low-skilled employment, and the proliferation of low-quality jobs (Parrilla et al., 2007; Sequeira and Maças Nunes, 2008; Arezki, 2009; Brida et al., 2016).

Furthermore, it is well-recognized that the expansion of tourism comes with other drawbacks. Issues such as higher resource consumption, noise pollution, increased crime, air and water pollution, and a more significant role in climate change are among the problematic effects associated with tourism (Gössling and Peeters, 2015). Additionally, tourism can exert tremendous pressure on local land use, leading to issues like soil erosion, loss of natural habitats, and increased threats to endangered species (Sunlu, 2003). Given these mixed potential impacts of tourism on urban development, it is crucial to conduct a thorough exploration and analysis of its influence on cities.

In a world where housing costs are rising rapidly and tourism is booming, it is essential to analyze the impact of these factors on local economies. This Ph.D. dissertation synthesizes three carefully conducted studies that explore these interconnected areas, each focusing on a specific aspect and providing nuanced insights into their respective effects.

Who is responsible for the high cost of housing? The second chapter of this thesis, 'The Power of Developers: Evidence from California,'[3] addresses California's current housing shortage. Over time, local land-use regulations in the United States have evolved into a stringent framework (Glaeser et al., 2005). This trend has been

[2]https://www.aptservizi.com/wp-content/uploads/2013/03/ITB_World_Travel_Trends_Report_2015_2016.pdf

[3]Co-authored with Albert Solé-Ollé and Elisabet Viladecans-Marsal

accompanied by a notable decline in housing permits, which aligns with a substantial surge in housing prices, pointing to a constrained housing supply (Gyourko et al., 2013). This phenomenon has been particularly pronounced in coastal regions like California, where the scarcity of new housing units, exacerbated by the rising demand for desirable locations, is fueling a severe housing affordability crisis (Quigley and Raphael, 2005).

This chapter delves into the influence wielded by politicians with affiliations to the real estate sector, referred to as "developers," over local housing policies. While the existing literature primarily emphasizes the role of homeowner voters, there is scant evidence to substantiate significant disparities in their preferences compared to renters (Hankinson, 2018; Marble and Nall, 2021). The findings demonstrate that real estate interests are remarkably present in California's local elections, with developers accounting for 11% of local council candidates. Moreover, 22% of elections feature at least one developer candidate running for office, and 11% of open seats are ultimately won by developers. These figures underscore the substantial involvement of this industry in city politics within California. This study delves into the motivations underlying developers' aspirations to hold local office, which are inextricably linked to the impact of local policy on their endeavors and their potential to wield influence upon election. While the chapter explores several aspects pertaining to the participation of developers in local politics, its primary objective is to provide concrete evidence regarding the extent to which developers can influence local housing policies upon assuming city council positions. To test the hypothesis, we use data encompassing all California city council elections from 1995 to 2017, obtained from the California Electoral Data Archive (CEDA). Notably, California's nonpartisan local elections mandate that candidates must display their occupation on the ballot, enabling us to distinguish between developer and non-developer candidates. To establish causality, a close-elections regression discontinuity design is used, focusing on races where a developer candidate faces off against a non-developer opponent and their vote shares are closely matched.

The study's findings indicate that developers possess substantial influence over local housing policies. For instance, the election of a developer to the city council leads to a remarkable 65% increase in the number of housing units permitted within the city during their entire office term. Moreover, this research reveals that the effects of developer influence extend beyond their term of office. We observe that the number of permitted housing units continues to rise at a similar pace during the following term, even though only a small fraction of developer-backed candidates are re-elected. This suggests that the long-term impact of policy decisions is significant, even if the original decision-makers leave office. The results also demonstrate that the presence of a developer on the city council has a more pronounced impact on multi-family housing decisions compared to single-family housing matters. Furthermore, a substantial portion of the effect on multi-family housing appears to be driven by lifting restrictions on the construction of such housing units. Intriguingly, we find that a developer's influence on housing policy is not dependent on the presence of other developers on the council. This implies that even a single developer can wield significant influence through effective negotiation tactics or specialized knowledge.

The third chapter of this dissertation, titled "A Country of Waiters: The Economic Consequences of Tourism Specialization," analyzes the lasting impact of local tourism specialization on per capita income. In the 21st century, tourism has grown in an unprecedented way and, importantly, it increasingly takes place in cities. This study focuses on Spain, one of the most popular tourist destinations in the world, where the tourism sector accounts for a substantial share of total employment, especially in coastal regions. This chapter brings new evidence to the debate on the long-term impact of tourism on the local economies by providing new empirical evidence on the localized effects of tourism specialization on income levels in Spanish municipalities. Moreover, it explores the underlying mechanisms behind this impact including the effects on job stability, sector displacement, and education levels.

Spain presents a perfect setting for this study, given its dramatic increase in tourism since the 1960s. Tourism plays a vital role in the country's economy, accounting for 13% of employment in 2019. That year, Spain attracted over 83 million international tourists, positioning it among the world's premier tourist hotspots. The country's appealing climate and picturesque beaches are key attractions, leading to a higher concentration of tourism activities in its warmer, coastal areas. According to the National Statistics Institute, in 2019, coastal municipalities accounted for over 70% of the total tourism activity.

This study uses two different approaches to analyze the impact of tourism development on per capita income. The first approach examines tourism growth since the 1960s, using data on the local business tax base from tourism-related activities to measure per capita tourism growth from 1960 to 2019. To address the potential endogeneity between tourism growth and local economic development, the existence of a beach is used as an instrumental variable for tourism growth, focusing on the variation in tourism exposure across different areas. This method is partly inspired by the insights of Fabert and Gaubert (2019), who emphasize the role of natural amenities in influencing tourism. In Spain, where most coastal municipalities have beach access, my strategy involves comparing municipalities with sandy beaches (treated municipalities) to inland municipalities (control municipalities) within the same coastal provinces.

The second approach focuses on the later surge in tourism from the 1990s, influenced by European policy regulations that improved air accessibility and facilitated the entry of low-cost airlines, as well as by the introduction of platforms like Airbnb. These changes had a significant impact on tourism inflows to Spain and were arguably not driven by the tourist sector's situation in Spain, making them exogenous. To leverage this context, the study shows that municipalities with a higher number of residents from tourist-source countries in 1996 experienced a more substantial increase in the number of overnight stays in hotels over the following years. To address the potential endogeneity, we adopt an identification strategy that uses a shift-share analysis. The strategy used involves distributing the positive surge in tourism inflows by source country across municipalities based on their pre-established composition of residents by nationality. This method is similar to the shift-share analysis proposed by Bartik (1991).

Using both research designs, the results indicate that municipalities with the highest growth in tourism specialization now exhibit lower levels of per capita income. Specifically, as of 2019, a municipality experiencing a median increase in tourism per capita over the period has a per capita income around 20% lower as of 2019. The chapter also investigates the potential channels through which tourism negatively affects income per capita. Findings indicate that tourism led to a shift in the local economy, resulting in a greater specialization in the service sector, as well as an increase in demand-related industries such as construction, real estate, hotels, and foodservice. Tourism also resulted in a decrease in manufacturing employment, lower educational attainment, and ultimately, higher job instability, since the tourism sector is more reliant on part-time and temporary contracts.

The fourth chapter, titled "The Impact of Housing Costs on Labor Market Trajectories,"[4] presents empirical findings on how housing costs influence the labor market outcomes of high-skilled and low-skilled workers in urban areas of France. Over the past twenty years, housing prices in France have experienced a significant increase, rising over 59% (OECD, 2022). Classic economic theories, as put forward by Rosen (1974) and Roback (1982), propose that city wages should adjust to account for varying housing costs. In this view, higher wages in larger cities would counterbalance increased housing prices, a result of businesses gaining from economies of scale in urban agglomerations. These models are predicated on the idea of worker mobility and the equalization of real wages and overall utility across different cities. However, considering heterogenous skills and preferences, and income elasticity of demand for housing below one, the model implies that low-skilled workers need relatively more compensation in expensive areas. Nevertheless, this theoretical framework

[4]Co-authored with Marie Aurélie Lapierre

doesn't entirely hold true in the French context, where lower-skilled workers exhibit less mobility and receive lesser wage adjustments for higher housing costs compared to their higher-skilled counterparts. This observation is in line with recent findings in similar studies conducted in other countries, as evidenced by research from Schmutz et al. (2021), Diamond and Moretti (2021), and Albouy et al. (2016).

Existing empirical evidence suggests that as housing in city centers becomes expensive, workers are often forced to move to more affordable regions, typically located on the peripheries of cities (Ong and Leishman, 2020; Bailey and Minton, 2018; Hulchanski, 2011). This shift results in a spatial disconnection between where people live and where job opportunities are located, profoundly affecting labor market outcomes. This is because a person's location plays a critical role in determining the amenities they have access to, the social networks they can develop, and their chances for human capital development.

This chapter hypothesizes that the higher relative burden of housing costs on low-skilled workers might significantly impact their labor market outcomes compared to high-skilled workers. In this paper, we focus on French urban areas, investigating how housing prices contribute to labor market disparities between high-skilled and low-skilled workers. Using a between-city approach that relies on the variation in housing costs across cities, the study analyzes how housing costs influence the probability of having a long-term contract and the probability of promotion, with a focus on the disparities between these two skill groups.

Methodologically, the chapter adopts a two-step approach similar to that of Combes et al. (2008), addressing worker sorting. Then an inverse housing supply elasticities is used to instrument for housing costs. The findings indicate that escalating housing costs intensify the disparities in securing long-term employment contracts between high and low-skill workers. Specifically, findings indicate that a 10% rise in housing costs correlates with a 0.14 percentage point increase in this disparity. This underscores the impact of rising housing costs on the job stability of low-skilled workers in comparison to their high-skilled counterparts. Further, the sector-specific analysis reveals that low-skilled workers consistently face greater disadvantages in the service, industry, and commerce sectors. The effects are more pronounced for low-skilled women than high-skilled women concerning the impact of housing cost fluctuations on long-term employment contracts. Age also emerges as a factor, with low-skilled workers above 34 years experiencing increased job insecurity relative to high-skilled workers as housing costs rise. However, this study does not find a significant impact of housing prices on the probability of promotion, suggesting that other factors may be more influential in determining career progression for workers of different skill levels.

Finally, Chapter 5 provides concluding remarks. It reviews the main results of the previous chapters. Additionally, it addresses the main lessons drawn from this research in terms of policy design.

2. The Power of Developers: Evidence from California

2.1. Introduction

Government regulations are widely recognized as an important driver of housing supply (Gyourko and Molloy, 2015). In the United States, the local land-use regulatory environment has become increasingly strict over time (Glaeser et al., 2005). Prior to the 1970s, the number of new housing units permitted annually was sufficient to moderate housing price growth. In contrast, in later decades, there was a decrease in permitting that coincided with a significant increase in housing prices, suggesting supply constraints (Gyourko et al., 2013). This trend has been particularly pronounced in coastal markets such as California (see Figure A.1 in the appendix), where the shortage of new housing, combined with rising demand for desirable locations, is leading to a severe housing affordability problem (Quigley and Raphael, 2005).

Who is responsible for the high cost of housing? Local politics undoubtedly plays a part in this situation. In the United States, local officials are responsible for drafting and approving land-use plans and zoning ordinances, as well as managing the permitting process, giving them the power to selectively deny, delay or speed permits for new development. This practice aims to ensure their re-election by pleasing various voter and interest groups that care about development (Trounstine, 2018). One such group is homeowner-voters who are concerned about the value of their homes and are often successful in pushing local governments to limit housing supply (see Fischel, 2001). Environmental groups and liberal voters may also actively oppose development in some cases (Glaeser et al., 2005; Kahn, 2011). However, local governments must also navigate coalitions of economic agents who support development (the so-called "growth machines" as described by Molotch, 1967). The debate over which of these two groups prevails remains unresolved (see Anzia, 2022, ch. 4).

The real estate industry plays a crucial role in the pro-growth coalition. The companies and professions associated with this industry, such as land developers, home builders, and real estate agents, have a significant economic stake in city land use policies (Anzia, 2022, p. 12). Council decisions can directly or indirectly affect the number and types of building permits, project costs (e.g., delays, red tape, development fees), and ultimately the industry's profits. Furthermore, compared to other types of businesses, real estate candidates may not prioritize other local policies, making it common for the real estate industry to invest significant effort in influencing city council decisions on development matters. This influence can be exerted indirectly, through campaign contributions, or directly, by educating candidates on housing policy, assisting them in organizing their campaign, or even running for office with a candidate of their own (Anzia, 2022, p. 183).

In this paper, we focus on the direct channel, while acknowledging the importance of the indirect one. Specifically, we examine the impact of a developer winning a seat in the city council on housing supply, which, to our knowledge, has not been explored in previous literature. Developers may find being on the council preferable to simply contributing to the campaign of other candidates who may be less knowledgeable about housing policies, care about other issues, and are possibly less trust-worthy (Gelbach et al., 2010). Additionally, serving on the council provides developers with the opportunity to establish a direct relationship with appointed officials such as the members of the planning commission and the city's bureaucrats, which could be just as important as voting power for project success. Our database (further details provided in the following section) also supports the relevance of the direct channel, indicating that roughly 11% of all city council candidates and members have a profession related to the real estate industry.

In this study, we aim to examine whether developers exert influence in the city council by analyzing the relationship between their presence and the issuance of building permits. To test our hypothesis, we utilize data on all city council elections in California between 1995 and 2017, obtained from the California Electoral Data Archive (CEDA). Notably, California's non-partisan local elections require candidates to display their occupation on the ballot, enabling us to identify developer and non-developer candidates. To establish causality, we employ a close-elections regression discontinuity design, where we examine races in which a developer candidate competes against a non-developer opponent and where the vote shares of the two candidates are similar.

Our findings suggest that the presence of a developer in the council has a significant impact on the number of housing units permitted during their term, increasing by 0.5 log points or a 65%. Our estimates are quite precise and have conducted several specification checks to ensure their robustness. However, it is important to note that although this effect is substantial, the starting level of permitting is relatively low. On average, cities in our sample are projected to increase the number of units permitted per 1,000 residents from 14 to 23.[1] Secondly, our results suggest that these effects are persistent. We find that permitting continues to increase by a similar amount during the following term of office, indicating that policy decisions have long-term effects. Furthermore, while only a fraction of developers is re-elected for a second term, amendments to land use regulations approved during their first term can continue to impact permitting in the future, even when the developer is no longer on the council. Thirdly, we find that the effect of having a developer in the council is larger for multi-family housing than for single-family housing. In addition, a relevant share of the effect on multi-family housing is driven by the extensive margin, suggesting that having a developer in the council is particularly critical for lifting prohibitions on building multi-family housing.

Fourth, the impact of developers on housing policy does not depend on the presence of other developers within the council. This suggests that even the first developer to enter the council can exert significant influence. This could be due to their ability to negotiate effectively or their expertise in housing policy. Fifth, the influence of developers is more significant in areas with a higher concentration of minority and liberal council members. This may indicate that developers are more successful in reaching agreements with these groups. Interestingly, our analysis did not find evidence of a stronger influence by developers in councils with a higher representation of businessmen, attorneys, or finance professionals, who are often associated with the "growth coalition." Finally, we did not uncover significant evidence of the impact of institutional factors on the power of developers in shaping housing policy.

Contribution to the literature. This paper makes several contributions to the existing literature. First, it contributes to the literature that examines the political economy of local land-use regulations and housing policies. Previous research has examined the role of various agents in shaping land-use regulations. For instance, Fischel (2001) finds that homeowners resist new housing developments due to concerns about the impact on their quality of life and property values, while Brueckner and Joo (1991), Glaeser et al. (2005), and Hilber and Robert-Nicoud (2013) provide theoretical analyses of the role of undeveloped landowners and developers. Recent work has also explored the impact of electoral institutions. For example, studies by Mast (2020) and Hankinson and Magazinnik (2021) reveal that California cities with at-large (rather than by-district) elections issue more permits. In addition, some studies have examined the role of the political ideology of the party in control of local government. For example, Kahn (2011) finds a negative relationship between the

[1] Notice that the number of units permitted over a four-year period in California was about 40 at its peak level in the 1970s (Figure A2.1 in the appendix). Interestingly, a recent report suggests that this figure should be multiplied by three to address the current affordability crisis (Bughin et al., 2016).

proportion of liberal voters and the number of housing permits issued in California's cities, while Solé-Ollé and Viladecans-Marsal (2013) demonstrate that left-wing Spanish governments allow for less land development than their right-wing counterparts. More recently, De Benedictis-Kessner et al. (2022) show that electing a Democrat mayor in the US leads to more multi-family housing permits.

There is limited empirical evidence on the ability of developers to influence local housing policies. However, Solé-Ollé and Viladecans-Marsal's (2012) work provides indirect evidence of the influence of developers in Spain. They found that development is higher where there is lower electoral competition. Nonetheless, the authors do not provide direct evidence of any link between developers and the local council. This is where our study of California data can make a contribution. Anzia's recent book (2022) does document a direct association between the presence of developer interest groups in a city and permitting activity. Our work builds on this, providing causal evidence of the influence of developers on land use policies.

Second, the paper also relates to the literature on how politicians' characteristics impact policy outcomes. For instance, Lee et al. (2004) found that the roll-call votes of Democrat congressmen score higher on a liberal scale. Regarding local governments, Ferreira et al. (2010) showed that the mayor's party does not matter for fiscal policies in US cities. However, papers such as Gerber et al. (2011) and De Benediktis-Kessner and Warshaw (2016) find that the partisan identity of US mayors matters for this policy. Folke (2014) discovered that the election of an additional councilor affects immigration and environmental policies in Sweden.

Moreover, some papers have examined the effects of the politician's profession. For example, Beach and Jones (2016), Kirkland (2020), and Szarkonyi (2021) looked at the impact of businessmen candidates on local fiscal policies. While the first paper found no effect for California city councils, the other two papers found significant effects on local infrastructure policies for US and Russian mayors, respectively. Importantly, there are no works in this literature studying local housing policies or, more specifically, focusing on the effect of professions related to the real estate sector. This is a critical omission for two reasons: first, land use regulation is one of the main policy areas of local governments, and second, as we previously argued, developers have a clear interest in this policy domain, which could make them very influential.

Outline of the paper. In the next section we describe who are the developers and which is their role in local politics. This allows us to derive some theoretical expectations regarding their effect on local housing policies. In section three, we provide some institutional details, regarding how local elections work in California and also regarding the role of local governments on housing policies. In section four, we describe the data and lay out of our empirical strategy. In section five, we present the results. The last section concludes.

2.2. Theoretical expectations

In this section, we explore why companies and professionals in the real estate industry have a vested interest in local housing policies and may seek to secure positions on the city council. We also investigate the factors that contribute to this industry's ability to influence the decisions of the council and other local officials. These insights will inform the theoretical predictions for our empirical analysis.

Which are the goals of developers? Throughout this paper, we will use the term 'developer' as a general reference to any company or occupation involved in housing production. However, it is important to note that three distinct groups operate within this industry: land developers, home builders, and real estate agents. Land developers acquire large tracts of land and subdivide them, requiring them to navigate the local regulatory system to obtain necessary building permits. Home builders design and construct individual homes or apartment buildings, either by purchasing serviced lots from developers or entering into contracts with them. It is also common for developer companies to have their own construction arm. Land developers and home builders' profits are closely tied to the amount of new construction taking place. Let's assume that they have similar goals to those of owner of undeveloped land identified in previous literature (Brueckner and Jo, 1991; Hilber and Robert-Nicoud, 2013). Then, we can see that unlike owners of developed land, who experience a decrease in asset prices with increased supply, owners of undeveloped land benefit from such situations as the increase in quantity compensates for the reduction in price. However, we do not assume that these groups are interested in unlimited growth. Rather, we expect them to be only 'relatively' more inclined toward construction than homeowner-voters, who typically hold the majority within the council.[2]

There is no denying that developers and home builders have a vested interest in increasing the supply of housing. This is evident in numerous statements made by industry associations. For instance, the National Association of Home Builders (NHAB) frequently lobbies the government to "enact policies that help home builders to expand housing supply, reduce the housing deficit, and improve housing afford-

[2] One may argue that firms that have a large share of the local housing market might try to tame the path of new construction to avoid the entry of new firms and keep prices high. In fact, there is some work showing that housing construction is lower in places with more market concentration (Quintero, 2020). This suggests that in places where homeowners do not control the council and there is a high level of concentration, the entry of a developer in the council could reduce (instead of increase) the number of permits. Unfortunately, we lack data on market power at the local level to test this hypothesis. However, we will present some heterogeneity results where we split the sample regarding the number of other council members that are potentially in favor of building more (minority and liberal councilmembers). We find that the effect of the entry of the developer is positive and significant even when there are no other developers in the council. Therefore, although we cannot rule out a possible negative impact of developers on permits in specific situations, it seems that this is not the predominant case in our sample.

ability for all Americans." Similarly, the California Building Industry Association (CBIA) urges the government to "reduce barriers to home construction and ensure that California can produce the homes needed to achieve housing for all." Additionally, these associations encourage their members to "identify, educate, and support candidates for elected office who are supportive of the building industry commitment to the American dream of homeownership" (NHBA, 2020, p.10).

The third group comprises real estate agents, whose main role is to facilitate the buying and selling of houses. They can work as independent contractors or for real estate agencies, brokerages, or development companies (as in-house salespeople). Unlike land developers and home builders, their profits are dependent not only on the number of housing transactions but also on selling prices, and some level of activity may be sustained even without permits to build new houses. Consequently, it could be argued that the pro-growth preferences of this group may be weaker. For this reason, we will present results separately for both groups: land developers and home builders versus real estate agents. Nonetheless, there is also anecdotal evidence of the pro-growth preferences of realtors. The National Association of Realtors (NAR), for example, advocates for supporting their own candidates because "cities having members on key committees, such as planning and zoning, tend to have a much more development-friendly climate" (NAR, 2020, p.5), implying that one of the association's goals is to ease local development. Additionally, realtor associations in California frequently join pro-growth coalitions in support of initiatives aimed at rolling back zoning regulations.[3]

What about the type of development that developers may be interested in? When land prices are very high, developers tend to turn to multifamily construction, which is more profitable. Given the significant opposition to this type of housing, it is even more likely that developers are more willing to allow more construction of multi-family housing than the average council member. Furthermore, multifamily projects often include retail space and affordable housing, which may facilitate reaching a deal. Anecdotally, industry associations such as the NHBA and NAR tend to complain about regulatory costs associated with multifamily housing, and some of the bills they sponsored aimed at spurring high-density development (see again fn. 3).

[3]For instance, the California Association of Realtors (CAR) played a key role in co-sponsoring Senate Bill 50 alongside CA YIMBY and the Non-profit Housing Association of Northern CA. The bill aimed to preempt local government control of zoning near public transit stations and job centers. Additionally, the association supported AB 48, a California bill mandating a 55% vote on local proposals seeking to reduce density or halt development.

Why are developers powerful? Developers wield power through three main channels. First, they may buy influence over local housing policies by contributing to the campaigns of city council candidates (Leffers, 2017; Yu, 2022). However, the use of campaign finance has some limitations. For one, campaigns in smaller cities tend to be relatively inexpensive. Moreover, many California cities have strict limits on campaign donations. Second, large donations can backfire, as they may suggest capture by interest groups and harm the candidate's electoral prospects (NBHA, 2017, p.23). Finally, professional politicians are beholden to multiple donors and constituencies, which may limit their ability to bargain on behalf of developers. As a result, interest groups may view professional politicians as less reliable allies (Gehlbach et al., 2010).

Secondly, the real estate industry wields significant power in organizing its lobbying activities across different levels of government. National associations not only influence policy at the federal level, but also provide guidance and training to state and local associations. Local associations work tirelessly to identify and educate candidates, as well as support their election campaigns (NRA, 2020). Moreover, some methods employed to support candidates do not rely heavily on monetary contributions. The NHBA, for instance, advocates for "get out to vote" techniques, whereby members mobilize friends, family, to vote in low turnout elections (NHBA, 2017).

Thirdly, the real estate industry's ability to provide valuable information about land use planning is also a significant source of power. Land use planning is a complex issue that requires continuous collaboration between local governments and the real estate industry (Leffers, 2017). The industry can provide this information by participating in public meetings regarding the approval of master plans and zoning ordinances or by offering education and training to city council candidates (NBHA, 2017). Therefore, it is essential to interact directly and continuously with appointed officials such as the planning commission members and city bureaucrats. The best strategy to achieve this goal is to place an industry member on the city council or to appoint a member of the planning commission.

Anzia (2022) emphasizes the significance of these direct channels, highlighting the value of lobbying activities and the endorsement, training, and selection of candidates by the real estate industry. Moreover, there is anecdotal evidence suggesting that the industry places great importance on having their own candidates running for office. For instance, the National Association of Realtors (NAR) encourages its members to participate in local government boards and commissions. Additionally, one of the NBHA recommendations to local BHAs regarding candidate recruitment is to 'call on one of your own.' "If many or your members are politically active, try to get them to take it a step further by getting them to run for state or local government. There is nothing better than to have state and local officials that understand the issues and can sympathize with them." (NHBA 2017, p. 15).

What constraints the power of the developers? When considering the potential effectiveness of a developer's entry into the council, several factors come into play. One crucial factor is the composition of the council, as the developer's ability to exchange votes with other members will depend on their traits and priorities (Casella and Palfrey, 2019). If the majority is strongly opposed to growth and prioritizes this issue above all else, the developer may be powerless to secure approval for their preferred projects. However, if there are other council members who are more open to growth and prioritize other issues, the developer may be able to reach a deal with them. In this paper, we will explore whether the developer's influence depends on the size of groups that exhibit these characteristics and can be measured with our data, such as other developers, businessmen, minorities, and liberals.[4]

The influence of developers on local councils can be moderated by various institutional factors. One such factor is the size of the council. In smaller councils, it may be easier for developers to buy the minimum number of votes needed to get a project approved. Another factor is the electoral system, with recent research indicating that councils elected at large tend to issue more permits. However, the impact of at-large elections on the influence of developers is uncertain. In district elections, electoral outcomes may be more sensitive to voter preferences, but there may also be less competition among council members, leading to more rent-seeking behavior. The timing of elections can also affect the influence of developers. Candidates elected in off-cycle elections may face less scrutiny, making it easier for them to approve construction projects. Additionally, the existence of campaign finance limits and transparency requirements can play a role. Tight contribution limits and the obligation to publish campaign finance information online may reduce the number of deals developers can make, but these constraints may be less effective for non-professional politicians. Finally, the stringency of land use regulations can moderate the impact of developers. In places with tight regulations, developers may need to reform them rto speed up the permitting process, which could reduce or delay their influence.

[4] Zoning regulations have been found to be used as a tool to exclude certain minorities from cities, as documented by Shetzer et al. (2016). Minorities also tend to face more housing affordability challenges than other groups. Therefore, it is natural to expect that minorities are in favor of zoning reform and the expansion of housing supply. The effect of partisanship on local policy in the US is less clear, as noted by Anzia (2021). However, recent research by De Benedictis-Kessner et al. (2022) shows that the political affiliation of mayors does have an effect on housing policy.

2.3. Institutional framework

Local Elections. In California, the main structure of municipal governments is established by state law. Council members are elected for staggered terms of four years, with elections filling seats every two years. However, there are some institutional differences among cities. Most cities in California (75.6%) have five-member councils, which is the minimum required by state law. Nonetheless, some big cities, such as Los Angeles, have larger city councils (15 members). Additionally, most cities (74.3%) elect their city councils at-large, meaning that council candidates are elected by a citywide electorate. However, some cities elect their council members by district. In the vast majority of California cities (96%), the electoral rule used is plurality (first-past-the-post). This means that voters are allowed to select as many candidates as there are seats being filled, and these seats are assigned to the candidates who receive the highest number of votes. The remaining 4% of cities use a run-off rule, which means that if no candidate receives more than 50% of votes in the first round, the two candidates with the most votes go to a second round. The council-manager form of government is prevalent in California (94.5%). Under this system, the council is responsible for establishing policy, while an appointed manager is in charge of its execution.

Local housing policies. In California, cities make two types of development decisions. The first type involves the adoption of broad policies by the city council that govern development, such as general plans and zoning ordinances. These policies determine where construction is allowed, what types of housing are permitted (e.g., single-family homes or apartments), and various other details (e.g., lot sizes and building heights). The development of these legal documents is time-consuming, and their effects on housing construction may not be seen for some years after their approval.

The second process involves discretionary rezoning and building permit approvals. In California, property developers or owners who seek to change the use of their land must go through the approval process with their local council. Once their proposal is approved and a building permit is issued, development can proceed legally. However, gaining local approval for housing development is often seen as expensive, risky, and slow, deterring developers and driving up costs. While the Planning Commission is responsible for reviewing these proposals, the city council has the final authority, particularly in cases of disagreement or appeals. Additionally, council members indirectly influence the commission through appointment of its members. For instance, in some small councils, each council member appoints a commission member, leading commission members to follow the guidelines of their appointer.

2.4. Data and Research Design

Identification of developers. The primary data source for this classification is the California Election Data Archive (CEDA), which contains information on the names, number of votes received, and professions of all candidates who have participated in local government elections between 1995 and 2017. California's state election laws require candidates to disclose their profession, making available this information in the CEDA dataset.[5]

We classify a candidate as having a real estate industry affiliation if their profession falls within any of the three groups described in section two: Land developers, Home builders, and Real estate agents. We will create two distinct categories: one for Developers and Home builders, and another for Real estate agents. As explained in section two, the second category may have weaker pro-growth preferences than the first one. To identify candidates with connections to the real estate industry, we selected the names of professions that appear to be associated with these groups. For candidates whose direct connection to the real estate industry is not evident from their profession, we used additional information to determine their affiliation.[6] We use auxiliary information sources such as candidates' web pages, newspapers, and LinkedIn to identify candidates with a connection to the real estate industry. If a candidate is specialized in tasks related to real estate, such as a 'real estate lawyer', or owns or works for a real estate company, such as a 'construction firm', they will be coded as related to the industry.

Table A2.1 in the appendix presents the final list of words used to classify our candidates. The primary professions considered in the first category (Land developers and Home builders) are those that contain the words 'developer', 'investor', 'builder', and 'contractor.' Still, we also include professions related to consulting jobs used by real estate companies (such as 'architects' and 'engineers'). The second category (Real estate agents) includes professions with the words 'real estate agent,' 'broker,' or similar. It also comprises professions that participate in the commercialization of houses, such as 'appraiser' or 'mortgage broker.'

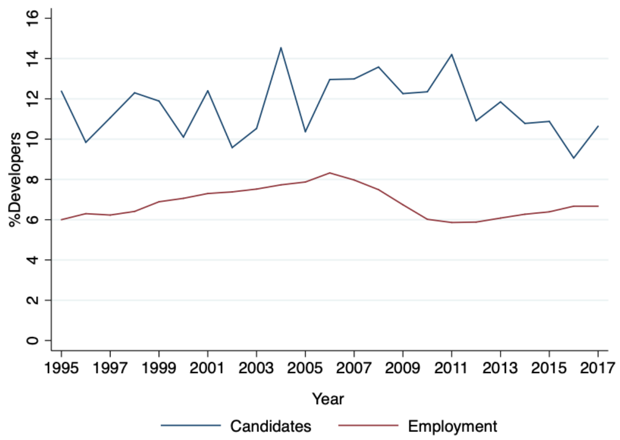
During the period under study, a total of 30,384 candidates ran for city council (refer to Table A2.2 in the appendix). By applying the classification method described above, we have identified 2,524 candidates as unambiguously related to the real estate industry. This represents 8.19% of the total number of candidates and 11.28% of those who were classified accurately. Of these, approximately half are classified

[5] Figure A2.2 in the appendix shows the image of a ballot; the profession of the candidate ('real estate agent' in this case) is displayed right below the name.

[6] The professions considered in this group are: 'lawyer', 'attorney', 'manager', 'ceo', 'executive' 'managers', 'director', 'businessman', 'engineer', 'entrepreneur', 'consultant', and 'commissioner'.

as Land developers or Home builders (5.79%) and the other half as Real estate agents (5.58%). This proportion is higher than the industry's employment share, which we estimate to be 6.68%. This suggests that the real estate industry is over-represented in local politics in California. Figure 2.1 displays the evolution of the share of developers as candidates or in employment over the period of 1995-2017. The difference between these two numbers remains relatively stable over time, with a slight increase during the housing boom and bust period.

Figure 2.1. Evolution of share of 'Developers': Candidates vs Employment



Notes: (1) California, City Council elections of the period 1995-2017; (2) The blue line is the % of Developers (Land developers + Home builders + Real estate agents) over Candidates running at these elections; Data from CEDA and own elaboration. (3) The red line is the estimated share of employment in real estate, which includes: employment in Construction (NIAC23), in Real Estate (NIAC531), and the number of state real estate licensees (less the number of employees in the sector); Data from Employment Development Department, ("Employment by Industry Data") and California Department of Real Estate.

Table 2.1 presents supplementary details that further support this idea. It shows that in 22 percent of the elections, at least one developer runs, and there are 0.15 developer candidates running for each open seat. Notably, approximately 11 percent of these open seats are won by developers, and 47 percent of the developers running actually win a seat. This evidence suggests that developers run and win in significant numbers in city council elections. The fact that they run in such high numbers implies that the decisions made in the council are crucial for their activities, as they have a high chance of being elected, and possibly because they can influence policies once inside the council. This is precisely what we aim to investigate in this paper.

Table 2.1.: Presence of 'Developers' in city council elections (1995-2017)

	Mean	SD	Min	Max	#Obs.
At least one developer running	0.219	0.441	0.00	1.000	8,769
#Developers running / #open seats	0.153	0.343	0.00	4.000	8,769
#Developers winning / #open seats	0.112	0.324	0.00	1.000	8,769
#Developers winning / #Developers running	0.473	0.483	0.00	1.000	1,921

Notes: (1) Data from the 8,769 city council elections held in California during the period 1995-2017 (#Obs.=8,769) of which 1,921 had at least one developer running. (2) Data source: see definitions and data sources in Table A2.4.

Outcome variable. We measure our main outcome variable, the number of permitted housing units, using data from the U.S. Census Building Permits Survey. This dataset covers most California cities between 1990 and 2020 and provides information on the total number of permitted units, as well as the breakdown between single-family and multifamily housing. Our dependent variable is units permitted because permitting is a political decision, whereas building completions are affected by exogenous factors such as the availability of internal financing.

Controls. We have compiled a variety of covariates to account for potential confounding factors (definitions and data sources are provided in Table A2.3 in the appendix). Our first set of controls consists of economic and demographic characteristics of each city, such as total population, median income, median gross rent, commuter share, homeownership rate, vacancy rate, and the shares of residents from different ethnic categories and educational backgrounds. We also utilize geographical data to determine whether a city is located close to the center of the urban area or to the coastline. Additionally, we incorporate information on the stringency of land use regulations from Jackson (2018). Our fourth set of controls includes city-level institutional information on council size, voting geography, and electoral rules, as well as candidate-level information such as gender, incumbent status, prior political experience, and ethnicity (coded using the wru package in R by Imai et al., 2016). Lastly, we calculate the Ideological CF score (Bonica, 2014) using campaign contribution data at the state level to classify candidates as either Liberal or Conservative.[7]

Sample. Our empirical analysis focuses on a sample of 953 mixed elections, where one of the two marginal candidates (i.e., the one winning the last seat or the runner-up) has a real estate industry background. This sample is smaller than the set of elections with developer candidates, as we only include cases where the

[7]Figure A2.3 in the appendix shows the Histogram of CF score for the candidates to the city council. The distribution is bimodal and very similar than the one reported by Bonica (2014) for other groups.

developer is one of the marginal candidates and is not competing against another candidate of the same type. Additionally, our final sample used in the estimation of the Regression Discontinuity Design will be even smaller, as it will be restricted to cases where the margin of victory is thin.

One concern with this approach is that the cities used for the identification of the developer's effect might be different from the average California city. This is because cities without mixed elections might not have any developer running, and developers might be more likely to run in places with certain characteristics, such as minority and liberal councils, less restrictive campaign finance and land use reform procedures, etc. Moreover, close elections might also differ from other mixed elections, as elections might be more competitive in at-large elections or in places with campaign finance limits. To assess whether this is the case, in Table A2.4 in the appendix we compare the means of a large number of variables in three samples: All cities, Cities with mixed elections, and cities with Close elections, which are a subset of those with mixed elections. The table shows that the values of all the variables are very similar in the three samples. All the socioeconomic variables are virtually identical. If anything, close elections seem to happen more often in cities with slighter smaller council sizes, with at large elections, using plurality rule (instead of runoff elections), or requiring special rules for the approval of regulatory reforms. The differences are, however, rather small. In any case, we will explore the effects of these institutional differences in heterogeneity analyses.

Regression Discontinuity Design. We employ a regression discontinuity design (RDD) to isolate the impact of electing a developer to the council on local housing policy. Specifically, we compare cities where a developer won a seat in the council by a narrow margin to cities where a developer lost a seat also by a slim margin. In such close elections, the winner is essentially determined by chance, which allows us to identify the causal effect of real estate representation on policy outcomes. We focus on the marginal candidates in elections where more than one seat is up for grabs, namely the candidate who won the last seat and the runner-up, where one of the two is related to the real estate industry and the other is not. This RDD approach can be thought of as a hypothetical scenario where we randomly replace one non-developer council member with a developer. Of course, this assumes that adding a developer to the council does not affect the chances of other developers getting elected, an assumption that we will verify later on.

The estimated equation is the following:

$$\log u_{it}^{t+k} = \alpha_1[Developerwins_{it} = 1] + \alpha_2[VoteMargin_{it}] + \alpha_3[Dev.wins_{it} = 1] \times [VoteMargin_{it}] + \rho \log u_{it}^{t-k} + \lambda_t + \eta X_{it} + \varepsilon_{it} \tag{2.1}$$

In our analysis, we use as our outcome variable the natural logarithm of the cumulative number of units permitted (per capita) since the election and until k years after, denoted by $\log u_{it}^{t+k}$. Here, the subscript i indicates the city, subscript t indicates the year, and k represents the number of years after the election. We primarily focus on k=4, which represents the full term-of-office of the politician. However, we also present results for shorter and longer time horizons. We choose to log the variable because the original data is heavily skewed, as shown in Figure A2.2 in the appendix. Logging the variable allows us to mitigate the potential influence of observations with large outlier values. The variable $1[Developer\ winds_{it} = 1]$ is a binary variable that takes on a value of one if the new council member is a developer who won the election, and zero otherwise. We also include a variable Vote margin, which represents the difference between the vote share of the developer and that of the competitor. This variable enters our analysis through a flexible function *f(.)*, estimated using a local polynomial on the optimal bandwidth.

Our main specification includes controls for the lag of permitted units, which represents the cumulative number of units permitted previously to the election, and time fixed effects. It is worth noting that controlling for the lag of the outcome is similar to computing the dependent variable in differences (that is, subtracting the prior value of the outcome). This approach was proposed by Lee and Lemieux (2010) and has been used in recent RD papers, including de Benedictis-Kessner et al. (2022).[8] If the lagged outcome is balanced at the cutoff, introducing it as a control will not bias the results and can increase the efficiency of the estimates, as Calonico et al. (2019) have shown. In addition to this control, our equation includes city-level variables (X_{it}), such as total population, income, the percentage of residents in each ethnic and education category, and dummies indicating whether the city is the center of an urban area or near the coast. We use these variables as a further validity test because if they are balanced at the cut-off, their inclusion should not affect the results.

The parameter, α_1, represents the causal impact of electing a candidate from the real estate sector to the city council. It reflects the local average treatment effect (LATE) of the entry of a developer in close elections, rather than the effect of having an additional developer on the council overall. However, we have already

[8]For transparency, we will present first the results without controlling for the lagged outcome and, as a robustness check, the results using the differenced specification.

demonstrated that cities with close elections in our sample are highly similar to all California cities in terms of a broad range of economic and political characteristics. Additionally, close elections make up a significant proportion of our sample, with around 60% of the observations falling within the optimal bandwidth and 40% within half the optimal bandwidth.

Estimation. For our estimation, we utilize a local polynomial of order one, estimated using a triangular kernel and the optimal mean squared error (MSE) bandwidth.[9] In the robustness checks section, we will present estimates based on different polynomial orders and kernel types, as well as a wide range of bandwidths. The tables will include both the conventional estimator and the bias-corrected standard error and confidence interval, as recommended by Cattaneo et al. (2019). We also cluster the standard errors at the city level.

We may encounter a potential complication when estimating our equation, as some cities may have zero permitted units in certain periods. While this is not problematic for total units and single-family units, the percentage of zeroes is higher for multi-family housing. To address this issue, we look both at the intensive and extensive margins, a has been recently recommended (Dong, 2019; Chen and Roth, 2023). For the intensive margin, we use the same RDD approach as before but define $\log u_{it}^{t+k}$ only for positive observations. For the extensive margin, we estimate a linear probability model that examines the decision to permit at least one multi-family unit.[10]

Validity. First, we assess whether the treatment and control groups have similar observable characteristics near the threshold. The premise is that if units are unable to precisely manipulate the forcing variable, there should be no systematic differences between units with similar values of the forcing variable (Cattaneo et al., 2019). We expect units just above and just below the threshold to be alike in all variables that could not have been affected by the treatment. According to Eggers et al. (2015), a falsification test like this can be conducted using lagged values of the outcome variable. In Table 2.2, we present the results of this test for our main outcome variable, the number of housing units.

[9] We also show results for the bandwidth minimizing the coverage error of the confidence interval (CER), which is preferred for inference (Cattaneo et al, 2020).

[10] For a full understanding of the impact on multi-family housing, we also examine the results using log(x+1). These results are affected by both the extensive and intensive margins, making it difficult to determine which one holds more weight. Chen and Roth (2022) caution against using this approach for this reason. The log(x+1) transformation has been employed in studies that use similar, such as Mast (2020) and De Benedictis-Kessner et al. (2023).

Table 2.2.: Effect of Developers on *lagged* Housing units permitted.

Dep. Variable: lagged log Housing units p.c					
	years 0 and -1			years 0 to -3	
	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.002	0.005	-0.041	0.075	0.039
Pr > $\|z\|$	[0.964]	[0.939]	[0.644]	[0.628]	[0.784]
Robust c.i.	(-0.460, 0.439)	(-0.417, 0.404)	(-0.391, 0.269)	(-0.291, 0.481)	(-0.323, 0.428)
Bandwidth	0.122	0.129	0.118	0.117	0.120
#Observations	969	969	969	980	980
#Effective Obs.	663	675	656	662	668
lag log Housing units	No	No	Yes	No	No
Year f.e.	No	Yes	Yes	No	Yes

Notes: (1) RD Estimates: triangular kernel with first-order polynomial fitted on the MSE optimal bandwidth. (2) We report the Bias-Corrected RD estimates, the robust p-value (Pr > $|z|$), and the Robust 95% c.i. computed as per Calonico et al. (2014). Standard errors clustered at the city level. (3) Dependent variable: lagged log Housing units per capita during the previous term of office, years 0 and -1 or years 0 to -3. (4) Year f.e. and lag Housing units per capita included as controls in some specifications; lag computed with years -2 to -3.

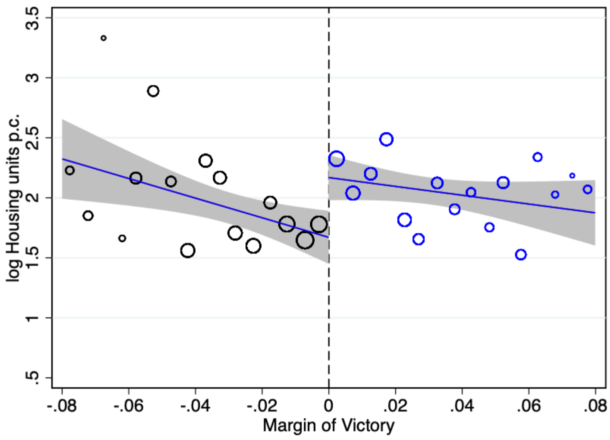
None of the specifications show an effect of the number of housing units permitted in the term prior to the developer's entry into the council (years 0 to -1 or years 0 to -3). The coefficients are negligible, and the p-value is high. Therefore, our findings are not attributable to the fact that developers are more likely to win in locations with more construction in the years preceding the election. Another falsification test that could be performed is to test for a discontinuity in other pre-determined covariates. The results are reported in Panel a of Table A2.5 in the Appendix. All point estimates are small with large p-values. There is no evidence that these predetermined covariates are discontinuous at the cutoff.

A second type of validity test examines whether the number of observations above and below the threshold is substantially different close to the threshold. The underlying assumption is that, if candidates lack the ability to precisely manipulate their vote share, the number of treated observations above the threshold should be similar to the number of control observations below it. To test for manipulation, we analyze the histogram and formally test for the continuity of this variable at the cutoff using both the conventional Mcrary test and the robust test proposed by Cattaneo et al. (2018) (refer to Figure A2.5 in the appendix). Neither of the two tests provide any evidence of manipulation.

2.5. Results

Regression Discontinuity Estimates: Main results. This section examines the impact of a developer's election to the city council on the number of permitted units. We primarily focus on a four-year period, which aligns with the council member's full term of office. However, we also present results at different time horizons later on.

Figure 2.2. Regression Discontinuity Estimates: Main results.



Notes: (1) Each point represents the sample average of the dependent variable for 0.5% bins of the Margin of Victory. (2) Dependent variable: logged number of total units permitted per capita during the first term of office of the council member (years 1 to 4). (3) The straight line is a first order polynomial in the Developer's Margin of Victory. (4) The grey areas show the 95% c.i. and the box includes the RD point estimate and the robust p-value.

The discontinuity in housing units permitted around the cutoff is illustrated in Figure 2.2, which shows the plot between housing units issued and the forcing variable. The graph provides evidence of a clear and sizeable discontinuity: cities marginally to the right of the cutoff (i.e., those with a developer elected to the city council) permit the construction of more units than those marginally to the left (i.e., cities where a developer is not elected). This finding reveals that during the four-year term following a developer's win, the number of permitted housing units increases by approximately 0.5 log points, which is equivalent to a 65% increase. [11]. This translates to a rise of 10 permitted units per 1,000 residents, from 14 to 24. However, it is worth noting that the number of permitted units in our sample is an historical low (for example, it was close to 40 before 1990). Thus, to put our results in perspective, addressing the current housing affordability problem by returning to that level of housing production would necessitate nearly doubling the number of permits. In fact, a recent study by a consultancy firm (Bughin et al., 2016) suggested that housing production should triple by 2025.

The full regression discontinuity results are presented in Table 2.3. All the specifications in this table use a local linear regression with a triangular kernel and the optimal bandwidth. Columns 1-3 use the bandwidth that minimized the Mean Squared Error (Calonico et al., 2014). The first column presents the raw estimates without any type of control. The second column controls for the lag of housing

[11]Figure A2.6 in the appendix replicates this graph for two distinct dependent variables. We plot the residuals of a regression between log Housing units p.c. and its lagged value and Year f.e.. Then, we plot the variable computed in first differences (log Housing units p.c. minus its lag). The outcomes appear comparable, albeit the graph indicates that the estimation is more accurate.

units and year fixed effects. The third column includes city-level controls. The point estimates are very similar in all these specifications. The estimates become clearly more efficient when we include controls for lagged units and year fixed effects, while the inclusion of city-level controls leads to a relatively smaller improvement in efficiency. In column 4 we present the results when using the coverage error-rate optimal bandwidth (CER). In column 5 we report the results using a polynomial of order two. The results remain very similar in both cases.

Table 2.3.: Effect of developers on Housing units permitted.

Dep. Variable: log Housing units p.c.					
	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.522*	0.556***	0.527***	0.587***	0.622***
Pr $>|z|$	[0.011]	[0.000]	[0.000]	[0.000]	[0.000]
Robust c.i.	(0.133, 1.013)	(0.292, 0.969)	(0.263, 0.879)	(0.287, 0.940)	(0.295, 0.914)
Effect in %A	0.685	0.744	0.694	0.799	0.863
Mean dep. Var.	14.123	14.123	14.123	14.123	14.123
Bandwidth selector	MSE	MSE	MSE	CER	MSE
Polynomial order	1	1	1	1	2
Bandwidth	0.093	0.085	0.082	0.061	0.135
#Observations	945	945	945	945	945
#Effective Obs.	588	561	549	495	667
lag log Housing units	No	No	Yes	No	Yes
Year f.e.	No	Yes	Yes	Yes	Yes
City controls	No	No	Yes	No	No

Notes: (1) RD Estimates: triangular kernel with a polynomial of order one fitted on the optimal bandwidth; optimal bandwidth computed using the MSE selector in columns 1 to 4, and the CER selector in column 5. (2) We report the RD Estimate, the robust p-value (Pr $>|z|$), and the Robust 95% c.i. computed as per Calonico et al. (2014). Standard errors clustered at the city level. (3) Dependent variable: log Housing units per capita during the full term of office of the council member (year 1 to 4). Lagged units refer to the years 0 and -1. (4) Municipal controls: lagged population, logged land area, count by tenure, urban dummy, CBD dummy, share of white, share with college education, and share of democratic voters. (5) Sample: elections of the period 1995-2017.

Additional validity checks. In this section, we examine additional factors that may impact the identification of our results. Firstly, there is the possibility that other candidate characteristics that are correlated with the preference for/against development could influence development outcomes. For example, assume that there is a higher fraction of developers that are white or conservative, and that white and conservative candidates prefer less development. If this was the case, then we would be underestimating the effect of a developer. To investigate the importance of these correlated effects, we conduct tests for discontinuity in candidate-level covariates such as ethnicity, gender, incumbency status, experience, and ideology. The results, presented in Panel b of Table A2.5, reveal that none of the estimates are statistically significant at conventional levels. However, the RDD coefficients on incumbency status and conservative binary variables might not be considered small (-0.167 and 0.199, with p-values of 0.211 and 0.324, respectively). Therefore, to further validate our findings, we re-estimate our model for samples in which the developer and non-developer share another characteristic (for example, none is an incumbent or both have the same ideology). Table A2.6 in the appendix presents this analysis for all candidate-level characteristics. The estimated coefficient is highly similar across all cases and almost identical to that estimated with the full sample. Consequently, we can confidently conclude that our estimated coefficient genuinely reflects the effect of electing a developer and not that of other candidate characteristics.

Secondly, to ensure that the RDD coefficient accurately represents the effect of electing a marginal developer to the council, it is necessary to examine whether the election of such a developer would displace other developers or related professions that were not included in our primary definition of developer. We address this possibility by presenting the RDD coefficients with the total number of elected developers or candidates with related professions as the dependent variable in Table A2.7 of the appendix. Similarly, we also look at the effect on the number of developers and candidates from other professions that are running at these elections. Our findings suggest that the increase in the total number of developers is not statistically different from one, which is what we expected. The number of council members with related professions and the number of developers running in the same election are not affected by the treatment.

Robustness checks. The results presented above are statistically significant and quantitatively meaningful. Moreover, they are robust to several robustness checks. First, we show that the estimated coefficient is fairly stable when we employ a broad range of bandwidths (Figure A2.7 in the appendix). The coefficient is less precisely estimated as we approach the threshold and is a bit smaller for large bandwidths. The figure also presents the bias-corrected coefficient which, as expected, is more stable across bandwidths. The two coefficients are virtually the same for bandwidths equal or smaller than the optimal one.

Second, we show that the results do not change much when we employ other kernels and/or a polynomial of order two (Figure A2.8 in the appendix). The results do not depend at all on the type of kernel; the coefficients are a bit larger and less precisely estimated when using a polynomial of order two. In accordance with the suggestion by Pei *et al.* (2020), we compute the MSE for each of these cases. We find that the local linear specification performs better in terms of MSE than the second order polynomial.

Third, we also show that our results do not depend much on the way we have selected our sample. We repeat the estimation with the following changes: (a) we add to the treated group other professions for which we have some doubts regarding whether there are related or not to the real estate industry; (b) we exclude these professions entirely from the sample to account for the possibility that the control group is contaminated; (c) we exclude the largest cities (Los Angeles, San

Diego, San José and San Francisco) since their housing market and local political dynamics might be quite different from the rest of the state; (d) we exclude short-term elections, that is elections where the elected candidate has to run again in two years (and therefore it is less clear that can impact policy); and (e) we exclude run-off elections, that do not use the plurality system and, for which, the close-elections RDD might be less appropriate. The results are reported in Table A2.8 and are very similar to the previous ones in all cases.

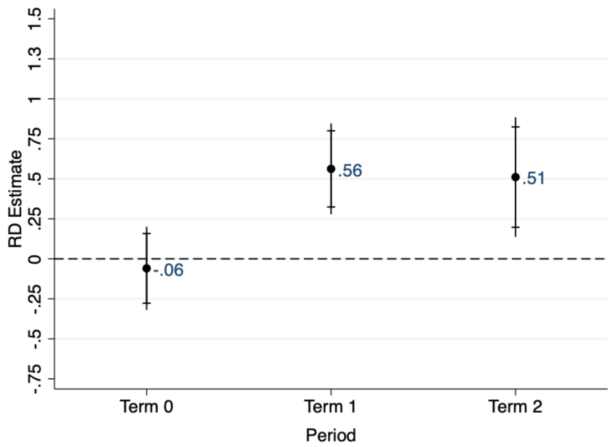
Finally, in Tables A2.9 and A2.10, we report the results when we use the dependent variable computed as a first difference (that is, log of housing construction in the current term vs in the previous term) and when using different clustering options (none, place, place x year, or county). The results are very similar in all cases.

Additional results: Developers vs Real estate agents. In Table A2.11 we look separately at the two main categories of real estate candidates: developers (which include home builders and real estate developers) and realtors (which include real estate agents, brokers, and the like). Recall that we anticipated a possible smaller interest (and so impact) of real estate agents on expanding housing construction due to an interest in keeping prices high. Instead, the results suggest that the effect are mostly similar for the two groups; the coefficient is slightly smaller for realtor than for developers (0.46 vs 0.57) but both are quite precisely estimated and so we cannot conclude that they are really different.

Additional results: Developers vs Businessmen. We also examine what happens when we include in our analysis other professions potentially related to local development: businessmen, attorneys and finance-related professions. Recall that we already included some of these candidates in our definition of developers. However, we were not able to find evidence on this aspect for many of them, which suggest that our control group might be contaminated. Furthermore, some authors have suggested that these are precisely the professions that should be included in the ‘growth machine’ concept and, therefore, that we could expect they have also interest in fostering development.

Therefore, we repeat the estimation with the following changes: (a) we exclude businessmen from the control group, (b) we exclude all three groups from the control group, (c) we use businessmen as the treated group but exclude the real estate developers from the control group, and (d) we do the same for the three groups (businessmen, attorneys and finance). The results are reported in Table A2.12 and show that the estimated effect is even larger when we exclude these groups from the control group. This suggests either that there are some remaining unclassified real estate candidates in these groups, or that these professions also have some impact on development when they are in the council. Nevertheless, the effect is not substantial and suggest that the real effects of might be even larger than what those presented in Table 2.3. Moreover, when we use our RDD to estimate the effect of the entry of a businessman in the council, we find a positive effect, but the coefficient is rather small (compared to that of real estate candidates) and not statistically significant: 0.12 with a p-value of 0.5. These results suggest that the effect on housing construction is restricted to real estate candidates and does not extend to other professions traditionally believed to have also a high interest in local development. We can not proof that these other professions also have interests and influence on local development in general (e.g., commercial and industrial development) but certainly, in our sample, they do not have a meaningful impact on the housing supply.

Figure 2.3. Dynamic effects. Housing units permitted by Term



(1) Dependent variable: logged number of total units permitted in each term per capita. Term 0: years 0 and -1, Term 1: years 1 to 4, Term 2: years 5 to 8. (2) RD Estimates: triangular kernel with first-order polynomial, including lagged Housing units p.c. and Year f.e. as controls. (2) 95% and 90% c.i. displayed; standard errors clustered at the city level.

Additional results: Dynamic effects. Figure 2.3 below shows that the effect of the entry of a developer persist beyond the years where we know for sure he is in the council. The coefficient for the second term is a bit smaller (0.51 vs 0.56 for the first term) but it is still statistically significant. The persistence effect can be due either to two different factors. First, the policy changes adopted in the first term can have long-term effects if they are difficult to revert afterwards. This may involve projects initiated during the first term with the support of the developer that got the final permit in the following term or the effect of the modification of land use regulations that ease the permitting process in the future. Second, the effect might also be due to the fact that some of the developers are reelected for a second term. Notice, however, that this percentage is not very large: only around one half of all developer candidates and one third of developers elected in close elections are re-elected (refer to results presented in Table A2.13 in the appendix). This suggest that the most important part of the effect is due to policy persistence.

For a more detailed analysis of the dynamic effects of a developer's entry into the council, refer to Figure A2.9 in the appendix. The figure presents results for two-year periods, indicating that the developer's impact is noticeable during the first half of their term, with a coefficient of 0.45 implying a 56% increase in units at the threshold. However, the effect intensifies during the second part of the term, with a coefficient of 0.87 indicating a 140% increase in the number of units. These findings suggest that decisions regarding delayed permits can be made swiftly after the developer wins the election. Nonetheless, the developer's entry may stimulate the application for new permits, which may take time to prepare and move through the bureaucratic process. As the effect drops significantly between the second half of the first term and the first half of the second term, it is crucial to expedite the process during the developer's tenure. This is because there is a high chance that the developer may not be re-elected for the next term.

Additional results: Housing typology. Table 2.4 presents the regression discontinuity results for different types of units, namely single-family and multi-family housing. Multi-family housing comprises units divided into two or more independent units, such as apartment buildings, condominiums, and duplexes, triplexes or fourplexes. We distinguish between these two types of units because zoning in most California cities restrict the areas where multi-family housing can be built, making it harder to increase production than in the case of single-family units. Additionally, multi-family housing has been a focal point of discussions around housing affordability, with opponents of new construction often targeting this type of housing, while supporters argue that it is crucial for easing the housing crisis. Therefore, to provide a comprehensive understanding of the impact of developers on housing supply, it is important to examine their ability to increase production of both single- and multi-family units.

Table 2.4 presents compelling evidence that the entry of a developer into the council has a significant impact on the permitting of both single-family and multi-family units. For single-family units, the estimated coefficient is around 0.4, indicating an increase of 50% in the number of units. The results for multi-family units, presented in columns 2-4, reveal that the entry of a developer affects both the intensive and extensive margins. The number of units permitted (provided that there is at least one permit) rises by a 56% (column 2), and the probability of permitting at least one unit increases by 11 percentage points, from 72% to 83% (column 3).[12] Both of these effects are statistically significant at the 10% level. To capture the joint impact of these

[12]Figure A2.10 in the appendix shows the dynamic results separately for single-family and multi-family units, and both for the intensive and extensive margins. The results mirror the dynamic effects found for total units.

two effects, we re-run our RDD analysis, this time using the log of housing units plus one as the dependent variable. Our results show that the entry of a developer increases the production of multi-family units by an 85%. This corresponds to a shift from a rate of about 3 units per 1,000 residents to nearly 6 units per 1,000 residents.

Table 2.4.: Typology of construction: Single-family vs. Multi-family housing

Type of Housing	(a) Single Family	(b) Multi family		
Depedent variable	log Housing units p.c.	log Housing units p.c.	l(Housing nits>0)	log (Housing units p.c. + 1)
	(1)	(2)	(3)	(4)
RD Estimate	0.399***	0.447	0.110	0.620*
Pr > $\|z\|$	[0.003]	[0.065]	[0.079]	[0.039]
Robust c.i.	(0.137, 0.731)	(-0.031, 1.023)	(-0.013, 0.241)	(0.218, 1.253)
Effect in %Δ	0.499	0.563	0.153	0.851
Mean dep. Var.	11.673	3.190	0.719	3.190
Bandwidth	0.097	0.093	0.110	0.096
#Observations	952	713	956	956
Effective #Obs.	602	445	630	576
Year f.e.	Yes	Yes	Yes	Yes
Lag log Units p.c.	Yes	Yes	Yes	Yes

Notes: (1) In columns 1 and 2 we present the same RD Estimates as in Table 2.3 for Single family and for Multifamily housing: using only the observations with positive values of the dependent variable. (2) In column 3 we repeat the same analysis for multi-family housing using as the dependent variable l(Housing units>0), which is a binary variable that indicates if there has been at least one multi-family unit permitted during the period. We use the whole sample to estimate a Linear Probability model. (3) In column 4 we show the results for multifamily housing using the log(x+1) transformation. (4) Sample: elections of the period 1995-2017.

The findings demonstrate that developers not only affect the production of single-family housing but also impact the politically complex process of developing multi-family housing. Moreover, it's worth noting that the presence of a developer on the council raises the likelihood of the council granting permission for the construction of multi-family housing, which was previously unapproved. This reinforce the idea that part of the effect of developers happens through the modification of existing zoning regulations.

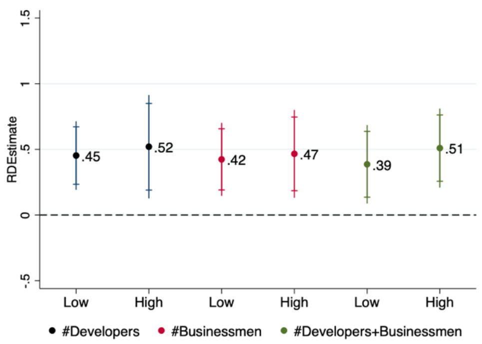
Moderators. In this section we explore several factors that might contribute to enhance or mitigate the influence of developers inside the city council –the moderators. With this aim we perform several heterogeneity analyses by interacting a binary variable for high and low values of the moderator with our treatment. This model is estimated parametrically using a polynomial of order one and the optimal bandwidth of the main analysis.

One limitation of this type of analysis is the challenge of ascribing the difference in coefficients between the two categories solely to the moderator under study. To mitigate this issue, we present additional results that simultaneously control for interactions between our treatment and all other moderator variables, as well as the city-level controls employed in our main analysis (see Table A2.14 in the appendix for these findings). By adopting this approach, we can attribute the results associated with a specific moderator to the variable used to divide the sample, rather than to other correlated variables that may influence both the moderator and the impact of developers on housing policy. However, it's worth noting that this analysis is still not perfect, as there could be other unobservable factors that are correlated with our moderators and may affect the results.

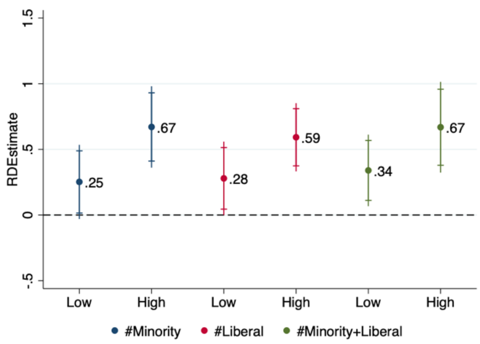
Figure 2.4 illustrates the first set of moderators, which examine the composition of the council. In Panel a, we investigate the impact of having other developers and council members with professions that are potentially interested in development (e.g., businessmen). The first graph in this panel examines the effect of having other developers on the council. The "low" category represents situations where no other developers are on the council, while the "high" category refers to cases where at least one other developer is present, in addition to the developer who won the election. Our results suggest that the influence of a developer on housing policy is not contingent on the presence of other developers on the council. The first developer raises the number of housing units by 56%, while the second developer increases it by 72%. However, the difference between these two coefficients is not statistically significant. The significant impact of the first developer could be due to her ability to negotiate effectively with other council members or her expertise in housing policy matters.

Figure 2.4: Moderators: Council composition

(a) Other developers and Businessmen



(b) Minority and Liberals



Notes: (1) Panel (a) reports the effect of electing a Developer depending on the number of other council members that are Developers and Businessmen (Low=there are no other councilmembers in this category, High = there is at least another councilmember in this category). Panel (b) reports the effect of electing a Developer depending on the number of Other council members that belong to a minority (i.e., are not Non-Hispanic white), which could be Low (zero or one Minority members in the council, besides the Developer) or High (two or more); this panel also reports the effect depending on the number of other council members that are Liberal less the number that are Conservative), which could be Low(#Liberals <#Conservatives) or High (#Liberals >= Conservatives). (2) RD Estimates: triangular kernel with first-order polynomial. (3) 95% and 90% c.i. displayed; standard errors clustered at the city level. Above each graph we report a t-test of the hypothesis that the difference between the two coefficients is zero (4) Dependent variable: logged number of total units permitted per capita during the full term of office (years 1 to 4); the variable is the residual of a regression against T and lagged logged Housing units p.c. (5) Sample: elections of the period 1995-2017.

The two other graphs in Panel a examine the impact of a developer in locations where there is already a businessman on the council, as well as in places where there is either another developer or a businessman present. The coefficients for the "high" category are slightly larger in both cases, but the difference is not statistically signifycant. It appears that the influence of developers is not contingent on the presence of businessmen on the council. This finding may be due to the fact that the first developper is the one who has the most significant impact, but it also aligns with the evidence presented in Table A2.12, which indicates that the entry of a businessman into the council does not have a substantial effect on housing policy outcomes.

Panel b in Figure 2.4 examines the council's composition from various perspectives. To begin with, we split the sample into two categories based on the ethnic make-up of the council. We investigate the effect of electing a developer on the number of housing units based on the number of minority council members (i.e., those who are not Non-Hispanic white) apart from the developer. The "low" category includes councils with either no minority members or just one, while the "high" category comprises councils with two or more minority members. The results in Figure 2.4 demonstrate that the RD coefficient is more substantial when there is a significant number of minority members on the council. Furthermore, the difference between the coefficients of the high and low categories is statistically significant at the 5% level. When there are already several minority members on the council, the entry of a developer results in a 95% increase in the number of housing units. In contrast, when there are not as many minority members, the effect is smaller, at 28%. This outcome is somewhat predictable since minorities are more adversely affected by the housing affordability crisis than the general population. The second graph in Panel b carries out a comparable examination for the case of ideology. We form two groups: councils with fewer liberal than conservative council members and councils with more conservatives.[13]

The results presented in the figure demonstrate that the effect of a developer's entry is much more substantial in councils with a higher number of liberal members. This difference between the two groups is large and statistically significant at the 1% level. Additionally, we are controlling for interactions with other moderators, so the effect is not confounded by other factors, such as the presence of more minority council members in liberal councils (see Table A2.14 in the appendix). This finding may be surprising since previous research has not shown a significant relationship between ideology or partisanship and local housing policies. While conservatives may oppose construction because they tend to be wealthier, white, and reside in sub-urban areas, many Democrats oppose development for different reasons. A study

[13]Notice that this count does not take into account the councilmembers for which we have not been able to compute the Ideological CF score.

by Kahn (2011) found that liberal cities tend to oppose development, but a more recent study by De Benefictis-Kessner et al. (2023) found that Democratic mayors allow more multi-family housing. Lastly, the final graph in Figure 2.4 confirms similar results by combining the previous two characteristics of council composition, minority status, and ideology.

In Figure 2.5, we examine the moderating effect of institutions. Panel a explores the effect of council size, election system, and election timing. The first graph shows that the effect is more significant for small councils (with a size below five) than for large councils. The size of the difference remains significant even when we control for interactions with other variables (see Table A2.14), suggesting that vote-exchange might be more manageable in small councils, even after fixing its composition (number of minority and liberal councilmembers). However, note that the difference between the coefficients of the two categories is not statistically significant. This is somewhat expected because there are likely too few cities with large districts to provide a reliable answer to this question.

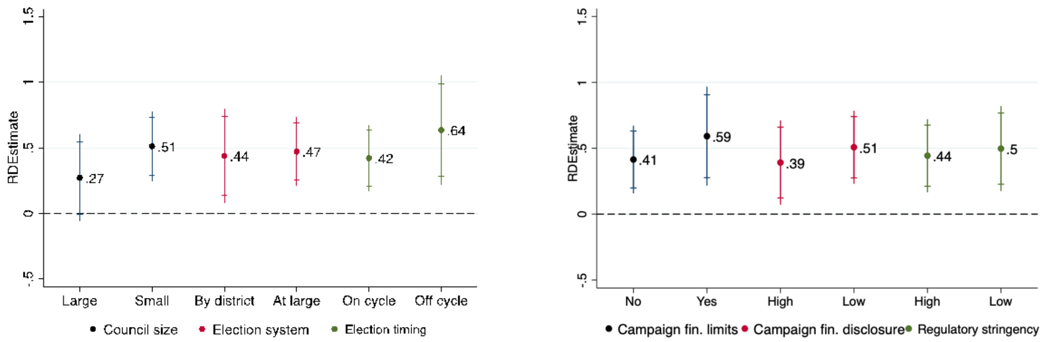
The second graph in Panel a reports very similar coefficients for councils with at-large vs. district elections. This institutional feature is a significant driver of local housing policies; as Hankinson and Magazinnik (2023) demonstrate, at-large districts tend to produce more housing because they prioritize the production of citywide public goods. However, we are exploring a different question – the impact of an additional developer – where the ex-ante theoretical prediction is less clear. On the one hand, in district elections, developers may find it more difficult to deviate from voter preferences. On the other hand, there are fewer candidates in these elections, and elections are less competitive, leading to more rent-seeking behavior.

The last graph in Panel a of Figure 2.5 examines the influence of institutional factors on the effect of the entry of a developer. Specifically, we investigate the moderating effects of council size, election system, and election timing. The first graph indicates that the effect of a developer is greater in small councils (with size below five) than in large ones, although the difference between the two categories is not statistically significant. This suggests that in small councils, vote-exchange might be more prevalent, even after accounting for other factors such as the number of minority and liberal councilmembers (as shown in Table A2.14, when we controls for all the interactions). However, due to the limited number of cities with large districts, we cannot draw definitive conclusions about the impact of council size.

In the second graph of Panel a, we show that the coefficients are similar for councils with at-large and by-district elections, despite the fact that election system has been reported to be an important driver of local housing policies in previous research (Hankinson and Magazinnik, 2022). However, our focus is on the question of the

Figure 2.5: Moderators: Institutions

(a) Electoral institutions (b) Other institutions



Notes: (1) Panel (a) reports the effect of electing a Developer depending on Electoral institutions: Council size (large= more than 5 members, small=5 members), Election system (By district or At large), and Election timing (On cycle vs Off cycle). Panel (b) reports the effect of electing a Developer depending on other institutions: whether there are Campaign finance limits (No=the limits passed by the State is applied, Yes=a tighter limit is applied), whether there are special Campaign finance disclosure requirements (High=information should be published online, Low=information need not be published online), and Regulatory stringency (High/Low=number of residential regulatory rules higher/lower than the median). (2) See Figure 2.4.

influence of an additional developer, which has a less clear theoretical prediction in the context of election system. On the one hand, in district elections, developers may face more constraints in deviating from voter preferences, but on the other hand, the fewer candidates and less competition may lead to more rent-seeking.

Finally, the last graph in Panel a examines the effect of election timing on the influence of a developer. The results indicate that the effect is larger in both off-cycle and on-cycle elections, with a larger difference between the two categories when controlling for other covariates (as shown in Table A2.14). This suggests that developers are more able to shape local housing policy in off-cycle elections, which typically have lower turnout rates. In such elections, interest group members are more mobilized and their activities may have a greater impact on voter behavior. This result is consistent with previous research, such as Anzia (2011) who found that teachers' salaries in California tend to increase more after off-cycle elections.

Panel b in Figure 2.5 examines other institutions that may similarly influence the power of developers, including campaign finance limits, campaign finance disclosure obligations, and the strictness of local land use regulations. When campaign finance limits are tighter than those enforced by the State of California, the effect of developers is significantly greater, making it more crucial to have a developer on the council, as it is more challenging to channel financial resources to professional politicians. However, there is no statistically significant difference between the two groups.

The second graph in Panel b explores the impact of campaign finance disclosure rules. We found that the effect of developers is less pronounced when there is a requirement to publish all information online, which is consistent with the results observed for off-cycle elections. Nonetheless, there is also no significant difference between the two groups in this case. It's worth noting that these last two tests may lack sufficient statistical power due to the small number of cities implementing stricter limits or transparency requirements.

Finally, the last graph in Panel b examines the impact of land use stringency. To do so, we utilized information from Jackson (2018) to calculate the number of different residential land use regulations implemented by each city. We found that the effect of developers remains relatively consistent, regardless of whether land use is heavily regulated or not. It is worth noting that the predictions in this case are not entirely clear. On one hand, if land use is unregulated, developers may have little incentive to influence policy decisions. On the other hand, if land use is tightly regulated, developers may find it challenging to modify regulations.

In conclusion, this section's main finding is that the composition of the council is the primary driver of developer influence. A council that is not entirely against development, preferably a smaller one, is more conducive to reaching agreements with other members. The presence of minority and liberal council members appears to be crucial for developers. The presence of businessmen and other professionals interested in local development appears to be less significant. While low turnout and lack of transparency, common in off-cycle elections, may be relevant factors, the evidence is weaker in this regard. Finally, other institutional features appear to have less of an impact.

2.6. Conclusions

This paper examines the influence of politicians with ties to the real estate industry, whom we colloquially refer to as "developers," on local housing policies. While much of the literature focuses on the role of homeowner-voters, there is limited evidence to suggest their preferences differ significantly from those of renters (Hankinson, 2018; Marble and Nall, 2021). We demonstrate that real estate interests are well-represented in local elections in California, with developers accounting for 11% of local council candidates. Additionally, 22% of elections feature at least one developer candidate, and 11% of open seats are ultimately won by developers. These are substantial figures, indicating that this industry is heavily involved in city politics in California. Our paper explores the reasons why developers are interested in running for local office, which are linked both to the impact of local policy on their activity and their ability to exert influence if elected.

While the paper addresses various issues related to the involvement of developers in local politics, its main objective is to provide concrete evidence regarding the extent to which developers can influence local housing policy if elected to the city council. The study's findings suggest that developers wield significant influence. For instance, the entry of a developer into the council results in a 65% increase in the number of housing units permitted in the city during their entire term of office. Although this is a substantial figure, it still falls short of addressing California's current housing shortage. Furthermore, our research indicates that the effects of developer influence are long-lasting. We observed that the number of permitted units continues to rise by a similar margin during the following term of office. Given that only a fraction of developers secure re-election for a second term, we interpret these findings as suggestive of the long-term impact of policy decisions. Finally, our results show that having a developer in the council is particularly influential in matters pertaining to multi-family housing rather than single-family housing. Additionally, a significant portion of the effect on multi-family housing appears to be driven by the extensive margin, implying that having a developer in the council is crucial for lifting bans on the construction of multi-family housing.

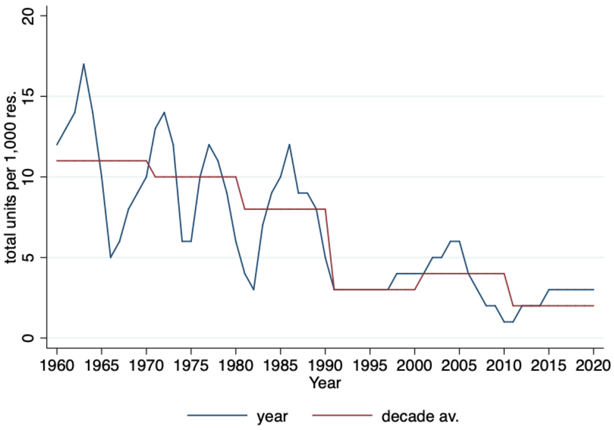
Furthermore, our study reveals several additional findings regarding the influence of developers on local housing policy. Firstly, we observe that having a developer on the council leads to a 65% increase in permitted units over the course of their full term, indicating that developers have significant sway over housing policy. However, this increase is not enough to meet California's housing shortage. Secondly, we find that the effects of developers on housing policy persist even after they leave office, as permitting continues to increase at a similar rate during the following term. Thirdly, we find that developers have a greater impact on multi-family housing than on single-family housing, and that a significant portion of the effect is due to lifting prohibitions on building multi-family housing.

Interestingly, we also find that a developer's impact on housing policy is not contingent on the presence of other developers on the council, suggesting that even a single developer can exert significant influence through effective negotiation or expertise. Additionally, we find that the influence of developers is greater in areas with a higher concentration of minority and liberal council members, potentially indicating greater success in reaching agreements with these groups. Surprisingly, we do not find evidence of a stronger influence by developers in councils with a higher representation of businessmen, attorneys, or finance professionals, who are typically associated with the "growth coalition." Finally, we did not observe significant evidence of the impact of institutional factors on the power of developers in shaping housing policy, aside from the effect of off-cycle elections.

Overall, our findings present a mixed picture. While the impact of developers on housing supply is large in absolute terms, it is less substantial when considered relative to California's overall housing shortage. This is particularly true for multi-family housing, which starts from very low permitting levels. Developers do have the ability to influence local housing policies and increase housing supply, albeit with limited impact on housing affordability. However, this effect is more pronounced in areas where the barriers to expanding housing supply are relatively low.

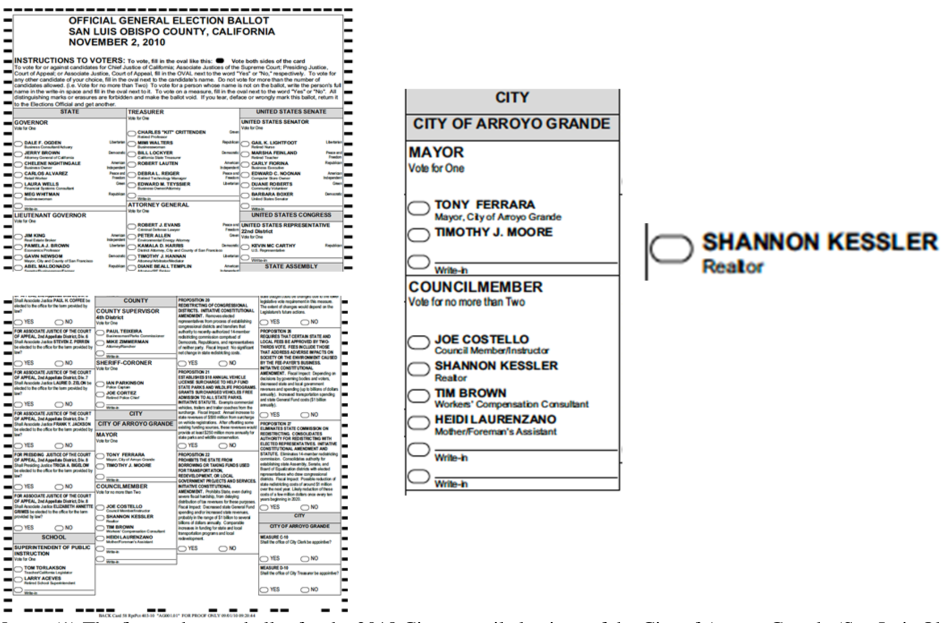
2.7. Appendix

Figure A2.1: Evolution of the number of units permitted



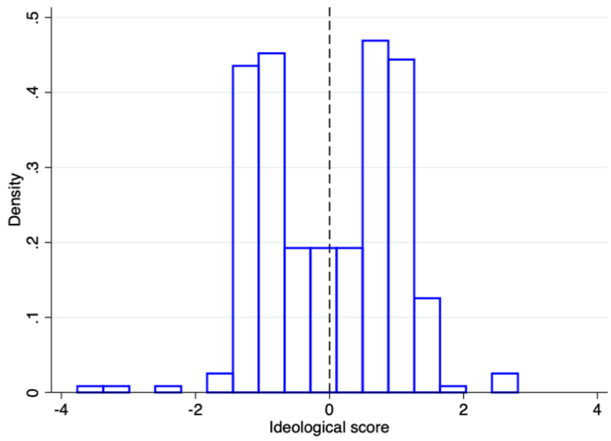
Notes: (1) California 1960-2020; (2) Variables expressed per capita x 1000. The blue line is the year value and the red line is the decadal average. (3) Source: US Census, Building Permits Survey.

Figure A2.2: Example of profession information in a ballot



Notes: (1) The figure shows a ballot for the 2018 City council elections of the City of Arroyo Grande (San Luis Obispo county). One of the candidates (Shannon Kessler) reports the profession of Realtor.

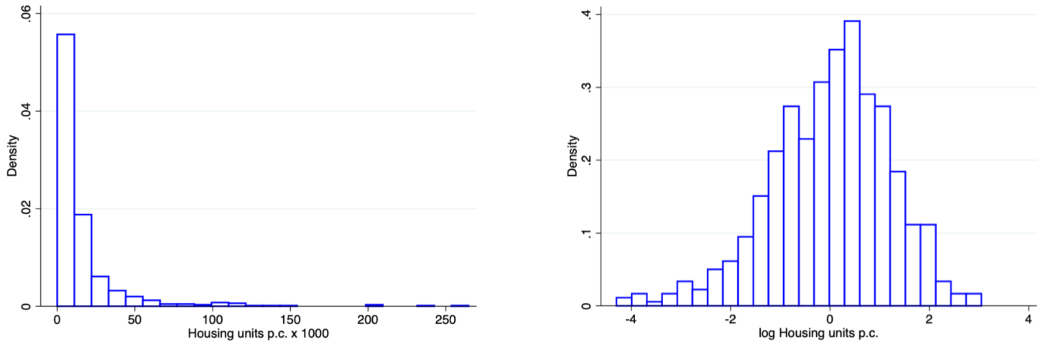
Figure A2.3: Histogram of the Ideological score



Notes: Ideological Score (Bonica, 2015). -/+: Lib./Cons.

Figure A2.4: Histogram of outcome variable

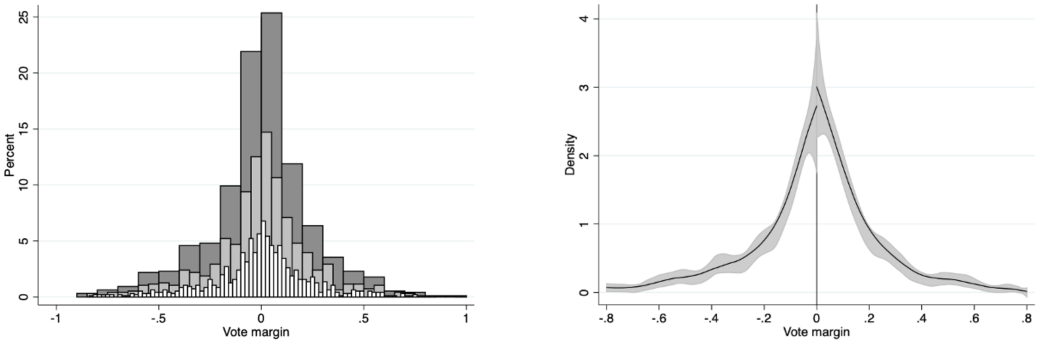
(a) Housing units p.c.   (b) Log Housing units p.c.



Notes: Panel (a) show the histogram of the number of Housing units permitted per 1000 residents; Panel (b) shows the logged value of this variable. (2) Sample: elections of the period 1995-2017.

Figure A2.5 Manipulation test

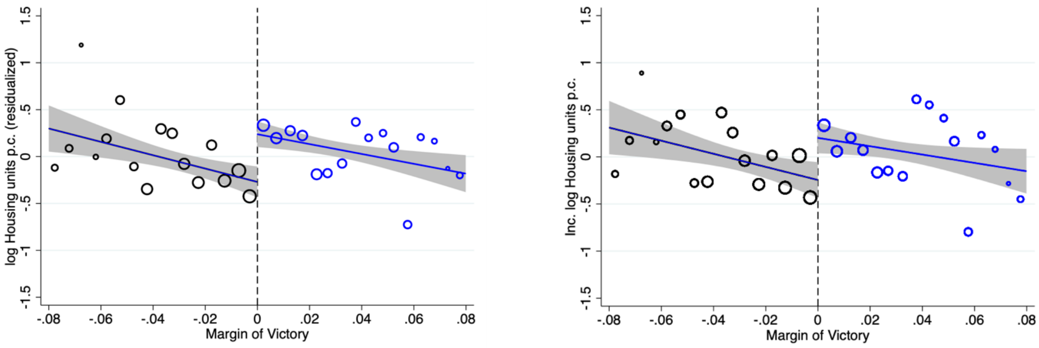
(a) Histogram (b) Cattaneo et al. (2017)



Notes: (1) Panel (a) shows the histogram of the forcing variable for bins of size 10%, 5% and 1%. (2) Panel (b) shows the result of the Cattaneo et al. (2017) test for the default options; we report in the box the value of the conventional and robust tests and the p-values.

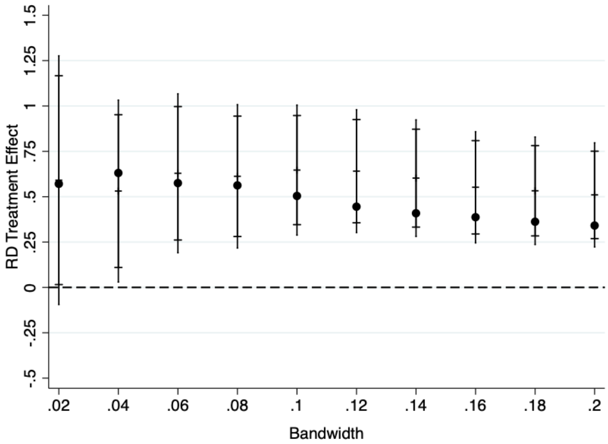
Figure A2.6: Effect of Developers on Housing units. Different specifications

(a) Residualized (b) Differenced



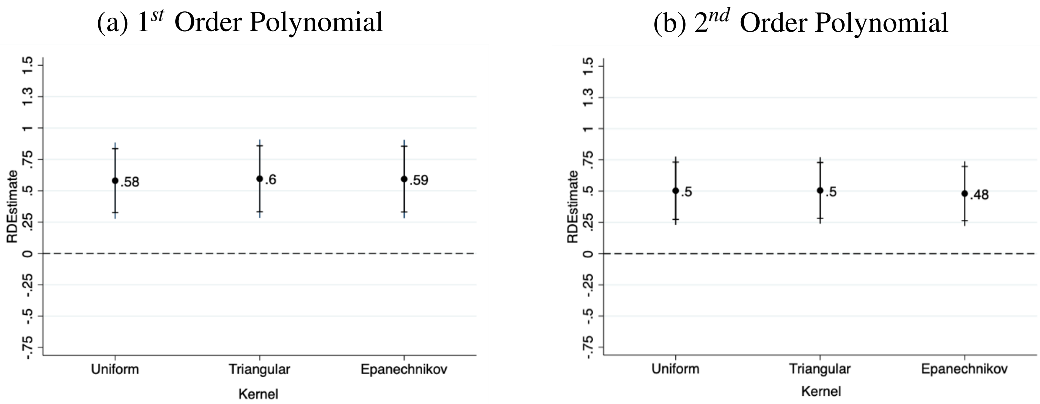
Notes: (1) Each point represents the sample average of the dependent variable for 0.5% bins of the Margin of Victory. (2) Dependent variable: logged number of total units permitted per capita during the first term of office of the council member (years 1 to 4); in Panel (a) the variable is the residual of a regression against lagged Housing permits p.c. and Year f.e.(3); in Panel (b) the variable is the increases wrt the previous term of office (year 0 and -1). (3) Sample: observa-tions with positive values of Housing permits. (4) The straight line is a first order polynomial in the Developer's Margin of Victory. (5) The grey areas show the 95% c.i. and the box includes the RD point estimate and the robust p-value.

Figure A2.7: Robustness: Bandwidths



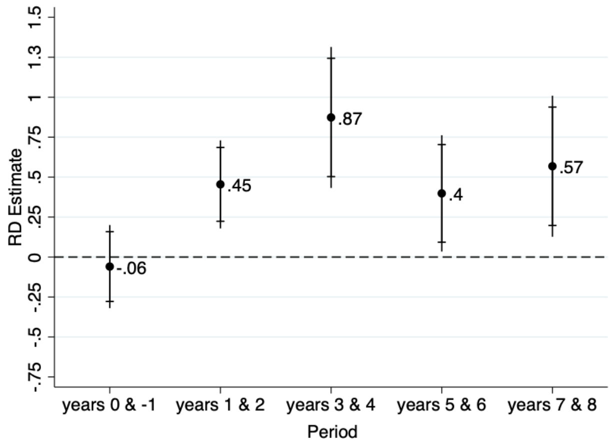
Notes: (1) The figure displays the results of the main RDD specification (triangular kernel with first order polynomial) for different bandwidths, including lagged Housing units p.c. and Year f.e. as controls. (2) We report the Conventional RD estimates (indicated by a black dot), and the Robust 95 and 90% c.i., computed as per Calonico et al. (2014). Standard errors clustered at the city level. (4) Dependent variable: log of Housing units p.c. permitted during the full term of office (years 1 to 4.).

Figure A2.8: Robustness: RD specification



Notes: Panel (a) reports the results for different types of kernel of fitting a local polynomial of order one on the optimal bandwidth. Panel (b) reports the same results for a polynomial of order two. (2) RD Estimates: triangular kernel including lagged Housing units p.c. and Year f.e. as controls. (3) 95% and 90% c.i. displayed; standard errors clustered at the city level. (4) Dependent variable: log of number of total units permitted per capita during the full term of office (years 1 to 4).

Figure A2.9: Dynamic effects. Total Units by 2-year periods.



Notes: (1) Dependent variable: logged number of total units permitted per capita. The figure displays the results for the number of permitted units awarded during each two-year period (i.e., years 1 and 2, years 3 and 4, etc.). (2) RD Estimates: trian-gular kernel with first-order polynomial, including lagged Housing units p.c. and Year f.e. as controls. (3) 95% and 90% c.i. displayed; standard errors clustered at the city-level.

Figure A2.10: Dynamic effects by Housing typology. Terms

Figure 2.1.: (a) Single family units (intensive margin)

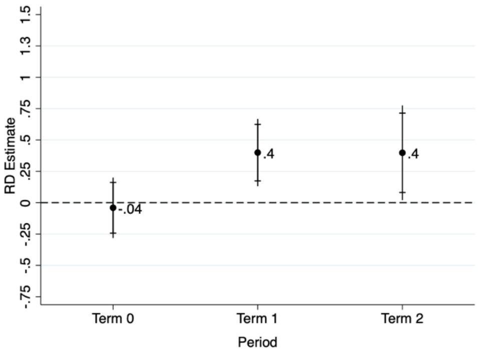


Figure 2.2.: (b) Single family units (extensive margin)

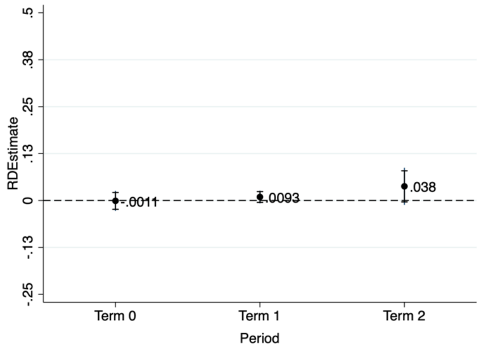


Figure 2.3.: (c) Multifamily units (intensive margin)

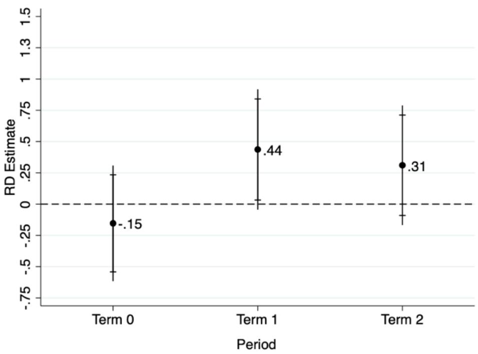
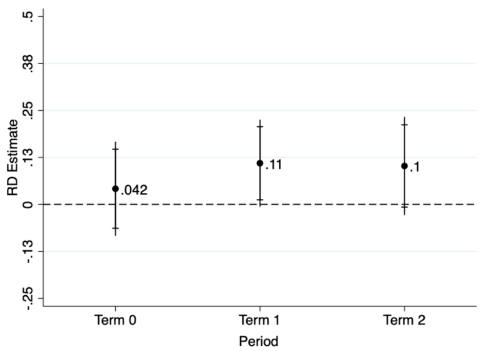


Figure 2.4.: (d) Multifamily units (extensive margin)



Notes: (1) Panel (a) reports the Conventional RDD coefficients for Single family units and for Multi-family units (intensive margin); the coefficient is expressed in log points, so in order to recover the effect in % Δ it has to be transformed using the expression exp(coef.)-1. Panel (c) reports the Conventional RDD coefficient for the RDD-LPM model; the coefficient can be interpreted as a in % Δ . Panel (c) reports the estimated% Δ of the two-part model estimated with the transformed variable; the coefficient can also be interpreted as a % Δ. (2) See notes to Table 5 for more details on the estimation of the Two-part model.

Table A2.1: Words used to identify professions related to the real estate industry

Real estate developers and *Home builders*	*Real estate agents*
'Developer' 'Property developer' 'Real estate developer' 'Affordable housing developer' 'Builder' 'Homebuilder' 'Contractor' 'Civil engineer' 'Architect' 'Construction firm'	'Real estate agent' 'Realtor' 'Real estate broker' 'Commercial property broker' 'Property investor' 'Property manager' 'Real estate investor' 'Real estate appraiser' 'Mortgage broker' 'Mortgage banker'

Table A2.2: Share of developers in the candidates to city councils, 1995-2017

Candidate category	# Candidates	% Over total	% Over class.
Non-developers	19,855	64.39	88.72
Developers	2,524	8.19	11.28
Land dev. and builders	1,276	4.14	5.70
Real estate agents	1,248	4.05	5.58
Total classified	22,379	72.58	100.00
Unclassified	8,455	27.42	
Unclear relationship	504	1.63	
Unclassifiable profession	1,321	4.28	
Unknown profession	6,630	21.50	
Total # of candidates	30,834	100.00	

Notes: The table reports the number of candidates running at city council elections in California for the period 1995-2017 classified in different categories according the relationship between her profession (or the activity of her company). Source: Own elaboration using data from the 'California Elections Data Archive' (CEDA) and several auxiliary information sources (web pages, newspapers, LinkedIn).

Table A2.3: Variable definitions and data sources

Variables	*Definition*	*Source*
		(a) City-level variables
Housing Units	Number of units permitted, total, single family, and multifamily	U.S. Census Bureau Building Permits Survey (1990-2019)
Population	Resident population	U.S. Census population estimates.
Land area	Total land area of the city	Own calculations using GIS
l(Coast=1)	City has access to the sea	Own calculations using GIS
l(Urban=1)	City belongs to a MSA	U.S. Census Bureau
l(CBD=1)	City is largest city of the MSA	
Income p.c.	Household Median Income	U.S. Census Bureau
%Homeowners	% living in owned properties	American Community Survey (ACS)
Housing price	Median home value	(1990, 2000, 2010-2017)
Rent	Median gross rent	
% Vacant houses	% of vacant houses	U.S. Census Housing Vacancies and Homeownership (CPS/HVS) (1990, 2000, 2010-2017)
% White	% of white population	U.S. Census Current Population Survey (1990, 2000, 2010-2017)
%College	% with college education	
% Democrat	% of individuals registered as Democratic political party	Voter Registration Statistics: California Secretary of State (1999-2017)
%Turnout	Voter participation in local elections	Own calculations using California Election Data Archive (1995-2017)
% Margin	Difference between the vote share of the developer and that of the competitor	
Council size	Number of seats in the council	California Election
l(At large=1)	One if electoral geography is at large, zero if it is by-district	Data Archive (1995-2017)
l(Plurality=1)	One if plurality rule, zero if runoff	
l(Mayor-Council=1)	One if the city has the mayor-council form of government (instead of Council-Manager or Commission)	
l(Off cycle elections=1)	Election held in even years	
Land use regulation index	Number of residential land use regulations	Own calculations using the data in Jackson (2016a and 2016b)
l(Land use approval =1)	Land use changes require either a super-majority in the council or a referendum	
l(Contribution limit=1)	City has own campaign contribution limit (contribution < 4,200$)	California Fair Political Practices Commission (www.fppc.ca.gov)
l(Contribution disclosure=1)	City campaign finance data published on the web	
		(b) City-level variables
l(White=1)	One if candidate is non-Hispanic white, zero if not	Coded using the wru package in R (Imai. K.. and Khanna. K. 2016).
l(Woman=1)	One if woman, zero if male	Coded based on list of common male and female names
l(Incumbency=1)	One if the candidate is the incumbent, zero if not	California Election Data Archive (1995-2019)
Ideology Score	Ideology CF score; negative values indicate Liberal, positive Conservative	
l(Conservative=1)	One if the candidate is Conservative, zero if Liberal (CF score >0)	Bonica (2014) CF score, own calculation for candidates to city councils
l(Ideology Score $\neq$.)	One if there is information about the CF score	

Table A2.4: Descriptive statistics

	(a) City level variables					
	All Cities		Cities with mixed elections		Cities with close elections	
Variables	Mean	(Std.)	Mean	(Std.)	Mean	(Std.)
Population	81379	114611	87638	126887	79106	140587
Land area	1335	3578	1387	3132	1524	3450
[Coast]=1	0.202	0.401	0.179	0.383	0.149	0.356
[Urban]=1	0.309	0.462	0.330	0.470	0.285	0.452
[CBD]=1	0.123	0.114	0.114	0.318	0.091	0.288
Income p.c.	61207	32326	63793	48851	63933	48878
Housing price	250	140	261	159	261	169
Rent	1151	333	1162	383	1160	390
% Vacant houses	0.085	0.076	0.086	0.075	0.082	0.071
% Homeowners	0.583	0.133	0.578	0.129	0.591	0.128
% White	0.661	0.191	0.672	0.196	0.692	0.191
% College education	0.272	0.169	0.275	0.174	0.275	0.173
% Democrat[1]	0.358	0.116	0.358	0.119	0.349	0.121
% Turnout	0.310	0.179	0.313	0.178	0.342	0.333
Council size	5.744	1.500	5.634	1.434	5.451	1.831
[At large election]=1	0.723	0.372	0.741	0.438	0.820	0.392
[Plurality rule]=1	0.834	0.477	0.854	0.353	0.903	0.296
[Mayor-Council]=1	0.297	0.289	0.924	0.265	0.934	0.248
Land use regulation index	0.233	0.198	0.219	0.199	0.222	0.208
[Land use approval]=1	0.378	0.151	0.444	0.151	0.428	0.450
[Campaign fin. Limit]=1	0.314	0.269	0.262	0.444	0.231	0.258
[Camp. fin. disclosure]=1	0.325	0.468	0.349	0.475	0.244	0.471
# Obs.	476		330		281	
	(b) Election level variables					
	Mean	(Std.)	Mean	(Std.)	Mean	(Std.)
#Seats	2.069	(0.868)	1.899	(0.829)	2.178	(0.805)
#Open seats	1.858	(0.976)	1.578	(1.112)	1.852	(1.028)
#Candidates	5.809	(3.282)	4.792	(2.634)	5.511	(2.691)
#Candidates per seat	3.107	(2.339)	2.673	(1.432)	2.468	(1.211)
1(Off cycle election=1)	0.292	(0.455)	0.272	(0.445)	0.271	(0.444)
%Margin of victory	0.098	(0.126)	0.118	(0.749)	0.035	(0.026)
# Obs.	5728		953		593	

Notes: See Table A.4 in the appendix for variable definitions and data sources.

Table A2.5: Covariate balance: City characteristics

			(a) City level variables			
Variables	Coeff.	P>|z|	Robust c.i.	Bandw	#Obs.	#Eff. ob.
log Population	0.174	0.295	(-0.179, 0.592)	0.128	953	693
log Land area	-0.088	0.512	(-0.372, 0.185)	0.099	953	635
[Coast]=1	-0.151	0.278	(-0.521, 0.189)	0.105	953	645
[Urban]=1	-0.009	0.983	(-0.331, 0.323)	0.125	953	692
[Suburb]=1	-0.017	0.913	(-0.312, 0.279)	0.136	953	692
[CBD]=1	0.010	0.943	(-0.338, 0.363)	0.132	953	702
log Income p.c.	-0.046	0.524	(-0.195, 0.099)	0.123	809	541
log Housing price	-0.092	0.356	(-0.344, 0.144)	0.099	878	545
log Rent	-0.050	0.314	(-0.187, 0.060)	0.091	858	552
%Vacant houses	-0.170	0.367	(-0.632, 0.233)	0.143	772	498
% Homeowners	-0.138	0.334	(-0.527, 0.179)	0.098	953	616
% White	0.073	0.528	(-0.245, 0.477)	0.099	953	636
%College education	0.051	0.687	(-0.241, 0.365)	0.106	953	643
%Democrat[1]	-0.077	0.528	(-0.433, 0.222)	0.110	953	649
%Turnout	-0.057	0.768	(-0.347, 0.757)	0.118	953	676
Council size	0.100	0.394	(-0.154, 0.390)	0.106	953	649
[At large election]=1	0.014	0.960	(-0.315, 0.332)	0.135	953	699
[Plurality rule]=1	0.024	0.865	(-0.276, 0.328)	0.160	953	754
[Mayor elected]=1	0.014	0.932	(-0.326, 0.298)	0.135	953	707
Land use reg. index	0.197	0.286	(-0.191, 0.647)	0.137	953	560
[Land use approval]=1	0.060	0.642	(-0.236, 0.450)	0.134	788	507
[Campaign fin. Limit]=1	0.158	0.338	(-0.175, 0.510)	0.114	995	560
[Camp. fin. disclosure]=1	0.189	0.242	(-0.267, 0.435)	0.137	995	674
			(b) Candidate-level variables			
	Coeff.	P>|z|	Robust c.i.	Bandw	#Obs.	#Eff. ob.
[White]=1	0.035	0.655	(-0.211, 0.355)	0.134	953	704
[Woman]=1	-0.062	0.804	(-0.322, 0.249)	0.149	953	770
[Incumbent]=1	-0.167	0.211	(-0.503, 0.111)	0.098	953	740
[Experience]=1	0.043	0.989	(-0.297, 0.291)	0.107	953	3652
Ideology Score	0.076	0.754	(-0.471, 0.651)	0.131	535	370
[Conservative]=1	0.199	0.324	(-0.231, 0.731)	0.118	535	359
[Ideolo. Score-missing]	0.101	0.432	(-0.288, 0.486)	0.088	953	614

Notes: (1) RD Estimates: triangular kernel with first-order polynomial fitted on the optimal bandwidth. (2) We report the Conventional RD estimates, the Robust 95% c.i., computed as per Calonico et al. (2014), and the robust p-value (P>|z|). Standard errors clustered at the city level. (3) All variables are standardized except those in logs.

Table A2.6: Validity checks: Correlated effects

	Same ethnicity (1)	Same gender (2)	No incumbents (3)	Same experience (4)	Same ideology (5)
RD Estimate	0.530	0.531	0.511	0.549	0.525
Pr > |z|	[0.001]	[0.006]	[0.001]	[0.004]	[0.000]
Robust c.i.	(0.221, 0.935)	(0.169, 0.989)	(0.222, 0.891)	(0.189, 1.008)	(0.255, 0.883)
Bandwidth	0.093	0.076	0.111	0.102	0.088
#Observations	715	633	637	513	929
Effective #Obs.	431	355	458	338	567
Lag log Units p.c.	Yes	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes	Yes

Notes: In column 1 we look at mixed elections between a Developer and a Non-Developer that have the same ethnicity; (2) We report the Conventional RD estimates, the Robust 95% c.i., computed as per Calonico et al. (2014), and the robust p-value (Pr > |z|). Standard errors clustered at the city level. (4) Sample: elections of the period 1995-2017.

Table A2.7: Validity checks: Effect on other developers

Dependent variable:	#Developers elected (1)	#Related oc. elected (2)	#Developers running (3)	#Related oc. running (4)
RD Estimate	0.871	0.004	0.092	-0.045
Pr > |z|	[0.001]	[0.978]	[0.789]	[0.594]
Robust c.i.	(0.674, 1.065)	(-0.143, 0.147)	(0.438, 0.576)	(-0.247, 0.141)
Bandwidth	0.127	0.120	0.096	0.117
#Observations	953	953	953	953
Effective #Obs.	661	650	597	644
Lag log Units p.c.	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes

Notes: (1) In column 1 we look at the effect on the number of developers elected (including the winning candidate). In column 2 we look at the effect on the number of professions related to development but that have not been included in the main variable used through the analysis. In columns 3 and 4 we look at the number of candidates running for each of the aforementioned groups. (2) RD Estimates: triangular kernel with first-order polynomial fitted on the optimal bandwidth. (3) We report the Conventional RD estimates, the Robust 95% c.i., computed as per Calonico et al. (2014), and the robust p-value (Pr > |z|). Standard errors clustered at the city level. (4) Sample: elections of the period 1995-2017. (2) See Table A2.5.

Table A2.8: Robustness checks: Sample selection

	Add related professions to treated	Exclude unclear cases from controls	Exclude largest cities	Exclude short term elections	Exclude Runoff elections
	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.404***	0.524***	0.533***	0.560***	0.601***
Pr > |z|	[0.002]	[0.001]	[0.000]	[0.000]	[0.000]
Robust c.i.	(0.161, 0.713)	(0.228, 0.919)	(0.251, 0.897)	(0.042, 0.965)	(0.017, 1.077)
Bandwidth	0.086	0.086	0.076	0.077	0.074
#Observations	1290	796	937	915	808
Effective #Obs.	767	466	531	527	483
Lag log Units p.c.	Yes	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes	Yes

Notes: (1) In column 1 we add to the treated group all those cases that could be defined as Developers under a laxer definition. In column 2 we exclude from the control group all cases that might be potentially contaminated (there is some chance they are unidentified developers). In column 3 we exclude the largest cities of California from the sample [Los Angeles, San Diego, San Jose and San Francisco]. In column 4 we exclude short term elections (where the candidate has to run again in two years). In column 5 we exclude runoff elections (that is elections that do not use the plurality system). (2) RD Estimates: triangular kernel with first-order polynomial fitted on the optimal bandwidth. (2) We report the Conventional RD estimates, the Robust 95% c.i., computed as per Calonico et al. (2014), and the robust p-value (Pr > |z|). Standard errors clustered at the city level. (4) Sample: elections of the period 1995-2017.

Table A2.9: Robustness checks: Differenced dependent variable

	Dependent variable: Δlog Housing units p.c.				
	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.433**	0.481**	0.514**	0.540**	0.566**
Pr > |z|	[0.005]	[0.001]	[0.002]	[0.002]	[0.003]
Robust c.i.	(0.147, 0.816)	(0.206, 0.853)	(0.207, 0.917)	(0.217, 0.918)	(0.195, 0.978)
Bandwidth selector	MSE	MSE	MSE	CER	MSE
Polynomial order	1	1	1	1	2
Bandwidth	0.108	0.104	0.081	0.074	0.145
#Observations	945	945	945	945	945
Effective #Obs.	618	608	546	556	694
Year f.e.	No	Yes	Yes	Yes	Yes
City controls	No	No	Yes	No	No

Notes: (1) Δlog Housing units p.c. = log Housing units p.c. build during the term of office (years 1 to 4) - log Housing units p.c. build during the previous two years (years 0 and -1). See Table 3.

Table A2.10: Robustness checks: Clustering options

	Dependent variable: Δlog Housing units p.c.			
	(1)	(2)	(3)	(4)
RD Estimate	0.558***	0.556***	0.556***	0.562***
Pr > |z|	[0.000]	[0.000]	[0.000]	[0.000]
Robust c.i.	(0.295, 0.908)	(0.289, 0.911)	(0.295, 0.905)	(0.316, 0.891)
Bandwidth selector	MSE	MSE	MSE	MSE
Bandwidth	0.084	0.085	0.085	0.083
#Observations	945	945	945	945
Effective #Obs.	556	561	561	549
Year f.e.	Yes	Yes	Yes	Yes
City controls	Yes	Yes	Yes	Yes
Cluster	None	Place and Year	Place	County

Notes: (1) See Table 3.

Table A2.11: Additional results: Developers vs Realtors

	Dep. Variable: 1(Housing units>0)		
	All candidates	Developers	Realtors
	(1)	(2)	(3)
RD Estimate	0.565***	0.576**	0.462
Pr > |z|	[0.000]	[0.008]	[0.012]
Robust c.i.	(0.289, 0.925)	(0.163, 1.096)	(0.110, 0.910)
Bandwidth	0.078	0.090	0.095
#Observations	953	414	539
Effective #Obs.	539	254	337
Lag log Units p.c.	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes

Notes: (1) The table reports the effect on Housing Units permitted by All candidates and also separately by Developers and Realtors. (2) RD Estimates: triangular kernel with first-order polynomial fitted on the optimal bandwidth. (3) 95% and 90% c.i. displayed; standard errors clustered at the city level. (4) Sample: elections of the period 1995-2017.

Table A2.12: Additional results: Real estate vs Businessmen

Included in the treated group	Real estate	Real estate	Businessmen	Businessmen, Attorneys and Finance
Excluded from control group	Businessmen	Attorneys and Finance	Real estate	Real estate
RD Estimate	0.687***	0.685***	0.120	0.042
Pr > |z|	[0.000]	[0.000]	[0.503]	[0.816]
Robust c.i.	(0.391, 1.106)	(0.371, 1.126)	(-0.296, 0.604)	(-0.300, 0.380)
Bandwidth	0.068	0.074	0.075	0.114
#Observations	768	670	758	986
Effective #Obs.	407	364	390	635
Lag log Units p.c.	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes

Notes: (1) In column 1 we study mixed elections between Real estate candidates (Developers or Realtors) and candidates from all other occupations except Businessmen. In column 2 we add to this latter group candidates with occupations as Attorneys or Finance. In column 3 we study mixed elections between a Businessmen and a candidate from all other occupations except Real estate. In column 4 we include in the same group candidates with occupations as Businessmen, Attorneys or Finance, and compare them to candidates from all other occupations except Real estate. (2) RD Estimates: triangular kernel with first-order polynomial fitted on the optimal bandwidth. (2) We report the Conventional RD estimates, the Robust 95% c.i., computed as per Calonico et al. (2014), and the robust p-value (Pr > |z|). Standard errors clustered at the city level. (3) Dependent variable: log Housing of units permitted per capita during the term (1 to 4 years).

Table A2.13: Reelection rates of developers and non-developers (1995-2017)

	% Run again	% Run and win	% Win	ΔVotes (in %)
		(a) All elections		
Developers	59	53	89	17
Non-developers	56	51	91	33
		(b) Close elections		
Developers	46	33	72	61
Non-developers	36	29	80	69

Notes: (1) % Run again: percent of incumbents that decide to run again after four years; %Run and win: percent of incumbents that both run again and are re-elected; %Win: percent of incumbents that run again and that are re-elected; ΔVotes (in %): percent increase in votes relative to the previous elections, for the incumbents that run again. (2) We report the results for all the incumbents and for those that were elected in close elections, which are races between a developer and a non-developer with a margin of victory lower than 8%. (3) Data: See Table 2.2 and Table A2.4.

Table A2.14: Moderators: full results

	Dep. Variable: log Housing units p.c.				
	(1)	(2)	(3)	(4)	(5)
I[Dev. wins = 1]	0.828	0.820	0.745	1.158	1.332
× [Low # Minority=1]	-0.457	-0.416	-0.384	-0.499	-0.422
×[Low # Liberals=1]	-0.337	-0.348	-0.308	-0.367	-0.308
×[No Other Developers=1]	–.–	–.–	–.–	–.–	–.–
×[No Businessmen=1]	–.–	–.–	–.–	–.–	–.–
×[Large council=1]	–.–	–.–	–.–	–.–	–.–
×[At large=1]	–.–	–.–	–.–	–.–	–.–
×[On cycle=1]	–.–	–.–	–.–	–.–	–.–
×[No Campaign fin. limit=1]	–.–	–.–	–.–	–.–	–.–
×[Low Campaign fin. disclosure=1]	–.–	–.–	–.–	–.–	–.–
×[High Reg. Stringency=1]	–.–	–.–	–.–	–.–	–.–
Polynomial order	1	1	1	1	2
Bandwidth	0.084	0.084	0.084	0.084	0.084
#Observations	953	953	953	953	953
Effective #Obs.	590	590	590	590	590
Lag log Units p.c.	Yes	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes	Yes
Interactions with City controls	No	Yes	Yes	Yes	Yes
Flexible polynomial	No	Yes	No	No	No

Notes: (1) RD Estimates: OLS regression with interactions with triangular weights, estimated on the optimal band-width fro the main analysis (Table 3). (2) The first row of the table shows the coefficient for the base category (e.g., in the first column are the observations with a High number of Minority and Liberal Councilmembers) and the other rows show the coefficient on the interaction of the treatment and the moderators. (3) In columns 1 and 4 to 6 the local polynomial is not interacted and is restricted to have the same slope in both sides of the cut-off; in column 2 we employ a fully flexible polynomial. (4) In columns 4 and 6 we add to the equation City controls and their interac-tion with the treatment.

3. A Country of Waiters: The Economic Consequences of Tourism Specialization

3.1. Introduction

The tourism industry plays a crucial role in generating income for many countries and is often considered a significant driver of economic growth. Before the COVID-19 pandemic, the World Travel and Tourism Council (WTTC) reported that tourism contributed to one in every four new jobs globally, accounting for 10.3% of both total employment and global GDP.[1] As a result, both developed and developing nations heavily invest in tourism infrastructure and promotional efforts, increasingly relying on the revenues it generates.

The impact of tourism on economic development is diverse, with both positive and negative effects on either local economies or countries as a whole. While tourism can lead to positive changes by redistributing economic activities across sectors, creating jobs, and attracting foreign exchange and investments (Song et al., 2012), some studies point out its drawbacks, including reduced productivity, potential displacement of industries undergoing technological progress, and a prevalence of temporary and low-skilled employment and, therefore, lower wages (Parrilla et al., 2007; Sequeira and Maças Nunes, 2008; Arezki, 2009; Brida et al., 2016).

This paper brings new evidence to the ongoing discussion about the lasting effects of tourism on local economies. It does so by presenting new empirical findings regarding the localized impact of tourism specialization on per capita income levels in Spanish municipalities. The study also seeks to delve into the mechanisms behind this impact, including its effects on job stability, sector displacement, and education levels. Spain provides an ideal backdrop for this research due to its remarkable surge in tourism since the 1960s, with the industry constituting a significant portion of employment (13% in 2019). In that year alone, Spain welcomed over 83 million international visitors, generating 71.2 billion euros in tourism revenue, making it one

[1] Source: https://wttc.org/Research/Economic-Impact (consulted on June 3, 2022).

of the world's top tourist destinations. The country's favorable climate and scenic beaches are major draws for tourists, leading to a concentration of tourism in warmer and coastal regions. In 2019, coastal municipalities hosted more than 70% of all tourism activity, as reported by the National Statistics Institute.

We use two distinct approaches in our study. The first one examines tourism development since the initial tourism boom in the 1960s, measuring per capita tourism growth from 1960 to 2019 using data from the local business tax base of tourism-related activities. The second method looks at the later wave of tourism development in the 1990s, triggered by external changes such as improved air accessibility and the introduction of new accommodation options. In this approach, per capita tourism is approximated by the number of overnight stays in hotels per capita. By combining these two approaches we are able to provide evidence from two different periods and using different identification strategies.

A key challenge in examining the causal impact of tourism on economic outcomes is the difficulty of separating the effects of tourism from other factors that could simultaneously influence overall economic growth in a tourist destination. For example, local investments and infrastructure development may independently contribute to economic growth in a municipality focused on tourism, apart from the influence of the tourism industry. This complexity makes it challenging to pinpoint the specific effects of tourism on economic outcomes, as tourism effects may be entangled with other contributing factors.

To address the potential endogeneity issue between tourism growth and local economic development, I adopt two identification strategies. First, when studying tourism development for the whole period 1960-2019, I employ a methodology that relies on cross-sectional variation in tourism exposure, using the presence of a beach as an instrument for tourism growth. This is inspired by Fabert and Gaubert (2019) and draws insights from the tourism literature which highlights the significance of local natural amenities in influencing tourism activities. Because there aren't so many coastal municipalities in Spain without access to a beach, what we do is to compare municipalities with sandy beaches (treated municipalities) to those located inland (control municipalities) within coastal provinces.

Additionally, I take leverage of the variation in the intensity of beach amenities along the Spanish coastline. More precisely, I use aerial photos from the PNOA hist´orico dataset (1956-57, Instituto Geográfico Nacional) to measure the surface area of sandy beaches across Spanish municipalities—specifically, the fraction of the onshore coastline covered by beach. I also consider weather features to further measure tourism attractiveness. This allows me to account for differences between the treated and control municipalities in exposure to tourism that are larger in places where beach amenities are stronger.

Second, when examining the second wave of tourism development in the 1990s, we adopt an identification strategy using a shift-share analysis. In this period, changes in the tourism industry were influenced by European policy regulations that improved air accessibility and facilitated the entry of low-cost airlines, as well as by the introduction of platforms like Airbnb. These changes had a significant impact on tourism inflows to Spain and were arguably not driven by the tourist sector's situation in Spain, making them exogenous. To leverage this context, I observe that municipalities with a higher number of residents from tourist-source countries in 1996 experienced a more substantial increase in the number of overnight stays in hotels over the following years. The strategy used involves distributing the positive surge in tourism inflows by source country across municipalities based on their pre-established composition of residents by nationality. This method is similar to the shift-share analysis proposed by Bartik (1991). Our shift-share instrument combines the share of residents from tourist-source countries in 1996 in each municipality with the growth rate of overnight stays by tourists from these countries across Spain for the period 1997-2019. Notice that differently than in the first approach, the sample used here includes only beach municipalities and exploits between-municipality differences in the exposure to tourism development drivers.

Using both research designs, the results indicate that municipalities with the highest growth in tourism specialization now exhibit lower levels of per capita income. Specifically, as of 2019, a municipality experiencing a median increase in tourism per capita over the period has a per capita income a 22,2% lower–when using the first approach, based on accessibility to beach amenities- and a 21,5% lower –when using the second approach, that relies on a shift-share analysis. I also investigate the potential channels through which tourism negatively affects income per capita. I find that tourism led to a shift in the local economy, resulting in a greater specialization in the service sector, as well as an increase in demand-related industries such as construction, real estate, hotels, and foodservice. Tourism also resulted in a decrease in manufacturing employment, lower educational attainment, and ultimately, higher job instability, since the tourism sector is more reliant of part-time and temporary contracts.

This paper contributes to several strands of literature. First, this analysis adds to the list of papers studying the causal impact of tourism on economic development. The paper complements the work conducted by Faber and Gaubert (2019). Their study focuses on the long-term effects of tourism on the development in Mexico. However, it is essential to acknowledge the distinctions between their research and mine. The economic context differs significantly; while Faber and Gaubert's analysis pertains to a developing country, my investigation centers on a developed country. Also, our examination explores other outcomes related to tourism's influence besides economic activity, namely job instability and education levels. Finally, their analysis pays specific attention to general equilibrium effects and show that coastal tourism impacts manufacturing activity inland. My analysis focus on the localized effects of tourism specialization and so the aggregate effects are out of the scope of the paper. I do take care of the possible spillovers from coastal tourism towards immediate inland municipalities. This is done in order to make sure that the estimates under the first approach (which compares beach and inland municipalities) are not biased due to the contamination of the control group, but it also brings some light on the effects of tourism beyond the municipalities hosting the amenity. However, I do not find any impact for inland municipalities beyond those adjacent to the beach municipality, which are also affected negatively, but to a much lesser extent. Of course, I cannot rule out that there are spillovers over more distant municipalities and regions (e.g., hosting industries that provide inputs to the accommodation and construction sectors) or on the country on the aggregate (e.g., due to the inflow of exchange during the early period).

There are other relevant papers in this strand of literature that also focus on the localized effects of tourism on development. Noticeable examples are the works by Nocito et al. (2021) and González and Surovtseva (2020), for Italy and Spain, respectively. The last paper focus on the short-time impact of tourism shocks on employment at the provincial level in Spain. The difference between the second paper and mine is that I focus on the long-run effects of tourism and I look at smaller geographical units, the municipalities. There is evidence that tourism is a highly localized phenomenon, as tourists are attracted by immobile resources, and consu-me mainly within the place they visit (Bronzini et al. 2019). Moreover, in Spain, there is a very high concentration of tourists in specific municipalities within a region.[2] All of this suggests that the municipality should capture the effect of tourism on local outcomes more accurately than the region or the province. Interestingly, and as already discussed, the results of the paper hold only for municipalities located very close to shore.

Secondly, this work contributes to a broad literature exploring the impact of structural transformation. Unlike most studies focusing on labor reallocation from agriculture to non-agriculture in developing countries (Makarski et al., 2022; Hjort and Poulsen, 2019; Bustos et al., 2016; Gollin et al., 2016; Herrendorf et al., 2014; McMillan and Rodrik, 2011; Kuznets, 1973), this paper centers on a developed country. Utilizing labor income data, it investigates the within-country impact of structural transformation, a departure from the focus on productivity gains in poorer

[2]For example, according to the National Statistic Institute, in 2019 the municipality of San Bartolomé absorbs the 67% of the total tourist presences of Gran Canaria (and the 12,7% of Spain) and Benidorm the 64% of Alicante (11,6% of the national total).

nations where income data from population censuses is often unavailable. Additionally, this study connects with literature exploring “Dutch disease" effects associated with natural resource booms, exploiting both between and within-country variation (e.g., Allcott and Keniston, 2018; Caselli and Michaels, 2013; Corden, 1984; Corden and Neary, 1982). While focusing on tourism as a unique form of natural resource boom, the economic questions addressed align closely with existing research.

The main contributions of this paper are as follows: firstly, an examination of the long-run impact of tourism on a developed country heavily reliant on tourism, where many municipalities depend on tourism income. Secondly, the use of unique municipality-level datasets to track localized changes in income and other outcomes, a feat not feasible with regional or country-level data. Thirdly, the provision of direct evidence on the mechanisms through which tourism affects socioeconomic outcomes, such as increased temporary contracts and changes in the sectoral composition of municipalities. Lastly, it stands as the first study estimating the long-term effects of tourism on a tourism-oriented economy, a crucial exploration given potential differences between short and long-term effects.

The paper is structured as follows: section 2 outlines the theoretical expectations behind the results. Section 3 delves into the Spanish institutional context, emphasizing the role of tourism in the economy. Section 4 describes the data and summarizes the primary variables used in the analysis. Section 5 details the empirical strategy, while section 6 presents the main findings and robustness checks. Finally, section 7 concludes the paper.

3.2. Theoretical expectations

The central hypothesis of this study posits that municipalities with tourism amenities, particularly those specializing in tourism, experience long-lasting impacts on their economic outcomes. Specifically, it is argued that tourism-specialized municipalities tend to exhibit lower long-term per capita income. This hypothesis suggests that the observed lower income levels are not incidental but stem from distinct mechanisms inherent to tourism specialization.

Drawing parallels with the natural resource boom literature, this analysis likens the tourism surge in Spain during the 1960s to a boom in commodity exports typically seen in economies rich in natural resources. Regions endowed with accessible natural amenities such as beaches, high temperatures, and sunny weather witnessed a transformative influx of tourism. This influx was akin to the discovery of new applications for these ‘natural resources,’ leading to a fundamental reshaping of local economies and fostering a shift towards service sector specialization.

As outlined in the natural resource theory (Corden, 1984; Corden and Neary, 1982), the availability of such natural resources tends to elevate marginal labor productivity in the tourism sector, thereby increasing labor demand. Given the labor-intensive nature of tourism and non-tradable sectors, employment in these specialized municipalities becomes heavily concentrated in the service sector. This shift curtails opportunities in traditional sectors like industry and agriculture, leading to a reduced production of tradable commodities. Conversely, it may boost the production of non-tradable commodities and services, alongside sectors like construction, which are closely linked to the tourism industry.

While tourism can bolster a country's economy in the short term, enhancing overall economic conditions, its long-term impact remains a subject of debate. Initially, tourism can drive structural economic shifts, particularly in economies reliant on primary sector activities, redirecting activity towards manufacturing and services. Benefits include increased foreign exchange earnings, job creation, local investments, exploitation of economies of scale, and dissemination of technical knowledge (Song et al., 2012).

However, some studies highlight potential drawbacks, notably the phenomenon of "Dutch Disease," where tourism's low productivity may displace more technologically advanced sectors, impeding broader economic development (Copeland, 1991; Holzner, 2011). A tourism-centric focus can skew export composition, leading to a decline in manufacturing – a sector typically associated with innovation and technological progress (Capó et al., 2007). This shift can restrict growth potential, stifling human development and productivity (Torvik, 2001). Gylfason (2001a, 2001b) further underscores this by showing how the exploitation of natural resources, often reliant on unskilled labor, can dampen long-term growth due to lower training levels.

Additionally, the tourism industry is vulnerable to demand fluctuations, impacted by external shocks such as terrorist attacks, pandemics, natural disasters, and political instability. These events can drastically affect tourist arrivals, revenue, and the economic stability of tourism-dependent countries (Enders and Sandler, 1991; Enders, Sandler, and Parise, 1992; Neumayer, 2004; González and Surovtsev, 2020).

Furthermore, critiques of the tourism sector focus on the nature of the employment it creates, often characterized by temporariness, low skill requirements, and consequently, low wages (Parrilla et al., 2007; Sequeira and Maçãs Nunes, 2008; Arezki, 2009; Brida et al., 2016). This industry frequently offers poor working conditions, unfavorable hours, and limited career progression, predominantly employing part-time, low-skilled workers, or those on temporary contracts, often at minimal wages. While job opportunities may surge during peak seasons, tourism areas typically experience high unemployment during off-peak periods.

Human capital accumulation models indicate that the increased attractiveness of low-skill jobs due to tourism may discourage further education, at least in the short term (Black et al., 2005b; Charles et al., 2015; Angelopoulos et al., 2017). With low-skill jobs offering immediate income, the opportunity cost of completing high school or pursuing higher education rises, thereby diminishing the perceived value of further education.

3.3. The Spanish context

Spain serves as an exceptionally suitable case study for investigating the impact of tourism on local income due to its significant reliance on tourism as a key economic driver. In comparison to other OECD countries, Spain stands out with a higher percentage of GDP and employment attributed to tourism. According to the OECD Tourism Statistics, on average, tourism constitutes 12.4% of Spain's GDP and employs 12.9% of its population, while the OECD average remains at 4.4% of GDP and 6.9% of employment.

Notably, coastal tourism dominates the tourism sector in Spain, with a staggering 70% of tourism activity concentrated in coastal regions, particularly within the 464 coastal municipalities. Despite representing only 16% of the total population, these coastal municipalities play a pivotal role in generating substantial tourism-driven economic activity.

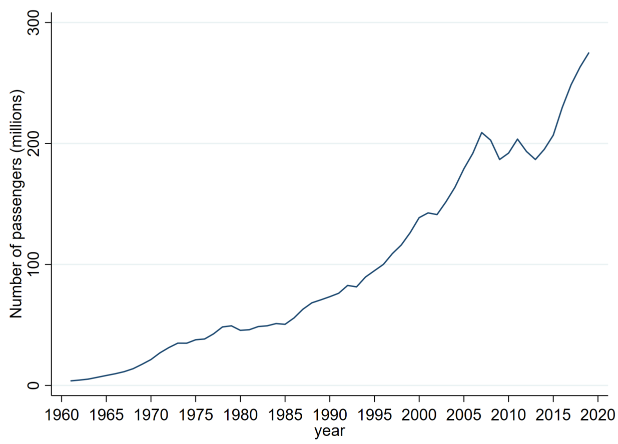
The development of tourism in Spain, particularly beach-and-sun tourism, is widely acknowledged to have started with the 'Plan de Estabilización' (Stabilisation Plan) of 1959. This plan, which included a significant devaluation of the national currency, marked Spain's return to the international markets after the dictatorship and the end of extreme autarchic policies. The effects of this plan on European tourists seeking sun and beach tourism in Spain were immediate and contributed to a sharp increase in tourism during the 1960s (Sánchez-Sánchez, 2001).[3]

Between 1997 and 2007, Spanish tourism experienced a second strong development period characterized by continuous growth in the number of international tourists (Albaladejo et al., 2020). This growth was primarily driven by exogenous changes in European regulations, which led to significant changes in the tourism industry in Spain including improvements in air accessibility, transportation, and

[3]The Stabilization Plan of 1959 refers to a set of economic measures that were implemented by the Spanish government from 1959 onwards. The main objective of the plan was to liberalize the Spanish markets and mark a shift away from the country's previous policies that aimed at achieving autarky. The plan was aimed at stabilizing the economy, reducing inflation, and increasing foreign investment. It involved the devaluation of the national currency, which helped to make Spanish exports more competitive on the international market.

the emergence of new forms of accommodation. Furthermore, this led to the introduction of low-cost airlines and the opening of new air routes. Thus, travel to Spain became more accessible and affordable for tourists from all over Europe. The number of international tourists visiting Spain increased significantly, particularly in coastal regions where sun and beach tourism are popular. Despite the increase in the number of tourists starting in the late 1990s, the real jump in growth did not occur until 2002, continuing until 2007, with the number of exits also increasing significantly in 2003 (Mason et al., 2016).[4]

Figure 3.1.: Growth trends in international tourist arrivals to Spain: 1961-2019



Notes: (2) Data represents the annual number of international tourists arriving in Spain. Two distinct phases of growth are visible: the tourism boom of the 1960s and accelerated growth in the late 1990s and early 21st century. (2) Source: National Statistic Institute (INE).

Specifically, the data in Figure 3.1 shows the numbers of international tourists arriving in Spain from 1961 to 2019, indicating a consistent growth in international arrivals since the 1960s, with several growth phases. The figure reveals two distinct periods of intense growth, the first of which was driven by the tourism boom of the 1960s. The second period was characterized by an accelerated growth at the end of the 1990s and the first years of the 21st century.

[4]The impact of these exogenous shocks after the 1997 varied depending on the origin and destination of tourists. It made it easier for tourists from Europe to travel to Spain, which resulted in a significant increase in the number of visitors from European countries. As a result, coastal municipalities, which were popular destinations for beach tourism, were among the most affected by the change in European regulations.

According to the Statistics on Tourist Movements at the Borders (FRONTUR), the number of foreign visitors was less than 5 million in the 1960s and almost no visitors in 1950. The majority of these visitors are from Europe, accounting for 75 percent of foreign tourists, followed by US Americans (4 percent) and Russians (1.5 percent).[5] International visitors make up roughly 70 percent of the tourism activity in Spain.

3.4. Data

Hereafter, I describe the main datasets used in the analysis. The data covers 22 coastal provinces and 2,658 municipalities (437 are onshore and 2,221 are inland, at different distances from shore) and spans the period 1960-2019. To construct the dataset, I digitized several historical records, including agricultural censuses, education censuses, and the 'Spanish Market Yearbook' since the beginning of the tourism development. In this section, I describe the main data sources and present some descriptive statistics.

Tourism data. In the first empirical analysis, we gauge tourism growth spanning from 1960 to 2019, utilizing data derived from the local business tax base associated with tourism-related activities. These activities encompass various accommodations such as hotels, motels, hotel-apartments, inns, boarding houses, guest houses, camp sites, and apartments. It also accounts for the income generated by restaurants, cafeterias, and bars. The local business tax in Spain, known as IAE ('Impuesto sobre Actividades Económicas', or 'Licencia fiscal' before 1992) is a presumptive tax that relies on proxies of business income, such as number of workers, power capacity, building surface, or, in the case of accomodation, number of rooms. This information has been instrumental in tracking the tourist activity of each Spanish municipality since the 1960s. The data is sourced from the 'Spanish Market Yearbook - Banesto' for the period 1965-1990 and from the 'Spanish Economic Yearbook – La Caixa' for the period 1995 to 2019. This information is available for municipalities with more than 1,000 inhabitants. Summary statistics for this variable can be found in Table A3.1 in the appendix.

To construct instruments for per capita tourism growth, I create variables related to the existence and extent of sandy beaches by utilizing aerial photos along with Geographic Information System (GIS) data. Specifically, I rely on aerial photo data from the National Geographic Institute (PNOA histórico 1956-57), featuring digital aerial orthophotographs covering the entire Spanish territory and captured in 1957. This enables the measurement of the length, width, and surface area of all

[5] Britons account for 21%, Germans 13%, French 13%, Italians 5,4%, Dutch 4,4% and Belgium 3%.

beaches before tourism development took place. An excerpt of the aerial photo data is presented in Figure 3.2. Additional GIS data layers, including the administrative shape file of municipality boundaries and the positioning of the Spanish coastline, are sourced from the geostatistics division of the National Geographic Institute. The average length and width of beaches in the sample are 3700 meters and 522 meters, respectively, as illustrated in Table A3.1, panel A.

Figure 3.2.: Aerial view of Spanish coastline prior to tourism development, 1956

Notes: (1) This orthophotograph provides a snapshot of the Spanish coast and beaches before the tourism industry's development. (2) Source: PNOA Americano Serie B for 1956.

In the second empirical analysis, I focus is on the sub-sample of municipalities with beaches. To measure tourism in this sub-sample, I use the number of overnight stays in hotels at the municipal level. This data is obtained from the National Statistics Institute and covers the period from 2003 to 2019. However, it is only available for a limited number of municipalities, which is 96 out of the 437 municipalities with beaches. To address this data limitation, I use the share of residents from tourist-source countries in each municipality to predict the number of overnight stays in hotels. This information is obtained from the Continuous Population Statistics of National Institute of Statistics (INE) from 1996 to 2019. Further details can be found in section 4.2.

Outcome variable. The primary outcome of interest is the per capita gross income in a municipality, obtained from the 2019 Household Income Distribution Map of the National Statistics Institute (INE). For years preceding 2004, there is a lack of available data on per capita income at the municipal level. To address this, I follow the approach of Parellada (1992) and estimate municipal per capita income through Ordinary Least Squares (OLS) regressions for the years 1965, 1990, and 1996. This estimation utilizes various economic indicators at the geographical level, including the number of trucks, the number of other firm-owned vehicles, the number of telephone lines, the number of commercial licenses excluding food, and the number of bank branches. The estimated per capita income serves as a control for the initial level of per capita income.

Regarding other key outcomes, various variables are sourced from the Public State Employment Service (SEPE), such as worker affiliation by economic activity and contract type (permanent or temporary contract) in 2019, and from the population census in 1990. Municipal-level years of schooling for 1990, 2000, and 2011 are derived from the population census. Additionally, data from the Continuous Register Statistics provides the number of residents by nationality at the municipal level from 1996 to 2019, allowing for the measrement of residents from tourist countries in a municipality. These variables are summarized in Panel B of Table A3.1.

Control variables. To account for potential municipal differences in socio-economic development before tourism development, I control for the level of agricultural production in 1960 using data from the 1962 and 1990 agricultural censuses. Population data from the population censuses since the beginning of the 20th century and education data from the 1930 census are also included. Geographical variables, such as soil quality, altitude, ruggedness, surface, temperature, rainfall, and hours of sun, are incorporated as additional controls. These control variables are summarized in Panel C of Table A3.1. The average population in 1950 was 5,138 inhabitants, with over 47 percent of the adult population unable to read or write in 1930. On average, almost 32 percent of the municipality's land was used for agricultural production in 1962.

3.5. Empirical analysis

The growth of tourism is closely intertwined with various factors that contribute to overall economic development. Factors that enhance a municipality's appeal to tourists, such as picturesque landscapes, favorable climates, or historical landmarks, can also make the city more attractive for residents and businesses in general. Additionally, advancements in accessibility, achieved through the construction of new highways and airports, have the potential to stimulate growth not only in tourism but also in other industries.

Therefore, accurately gauging the impact of tourism on per capita income necessitates identifying a source of external variation in tourism attractiveness. This paper employs two distinct approaches for this purpose. The first approach focuses on the evolution of tourism since its initial surge in the 1960s, utilizing cross-sectional variations in tourism exposure linked to amenities like beaches and weather for identification. The second method examines a subsequent wave of tourism development in the 1990s, employing a shift-share analysis that combines the percentage of residents from countries that are major sources of tourists in each municipality with the growth rate of tourists from these countries throughout Spain.

3.5.1. Long-term analysis from 1965-2019: beach amenities

In this subsection, I present the first approach, which is aimed at capturing the long-term economic consequences of tourism on local economic outcomes across Spanish municipalities. To do so, I exploit cross-sectional variation in tourism attractiveness, related to the presence of sandy beaches. I begin by documenting the effects of growth in tourism per capita on municipality-level per capita income in the current cross-section of Spanish municipalities, and estimate the following regression:

$$y_{i,2019} = \alpha \Delta Tourismpc_{i,1965-2019} + X'_{i,0}\beta + \gamma_p + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of interest in per capita terms (e.g. log per capita gross income in 2019) in municipality i in 2019. The variable $Tourismpc_{i,1965-2019}$ measures the growth per capita in municipality i, and in year 1965-2019.[6] I also include a number of municipal-level control variables, $X_{i,0}$, measured prior to the start of the period, and that I described in the previous section, and province fixed effects γ_p. One noteworthy aspect of this approach is its ability to capture the long-term impacts of tourism on the local economy, starting from its inception.

The primary challenge in studying the causal impact of tourism on economic outcomes lies in the likelihood that tourist inflows are correlated with various factors that independently affect economic outcomes, apart from their impact on tourism. This correlation could result in a biased Ordinary Least Squares (OLS) estimation of the equation (1).

To address this issue and exploit plausibly exogenous variation in tourism attractiveness across Spanish municipalities, I adopt the argument put forth by Fabert and Gaubert (2019). They posit that tourism activity is largely determined by the quality of specific local natural amenities, with the presence of a sandy beach being a key factor. Given that the Spanish tourism boom in the 1960s predominantly affected municipalities with easily accessible natural resources like beaches, high

[6]Notice that the tourism variable is not measured in logs, in order to address the challenges arising from the presence of zeros in the variable in the year 1965. During this year, many of the municipalities exhibited zero tourism activity. This variable is also very skewed, so in order to interpret the estimated coefficient we will multiply the coefficient by the median value of the sample, meaning that the results will indicate the effect percent change in per capita income as of 2019 for a municipality that during the period 1965-2019 has experienced an increase in tourism specialization equal to the median.

temperatures, and sunny weather, I use the presence of a beach in a municipality as an instrument for per capita tourism growth.

To compare municipalities with and without tourism amenities at the outset of the tourism boom in the 1960s, I focus on those with beaches and inland municipalities. Since nearly the entire Spanish coast has a beach (except for 2%), I choose inland municipalities as the control group. I control for various socioeconomic characteristics predating the onset of tourism, as detailed in the data section, to ensure comparability between coastal and inland municipalities.

In the first stage, I estimate the following first stage equation:

$$\Delta Tourismpc_{i,1965-2019} = \mu \mathbb{1}(Beach_i) + X_{it}'\eta + \gamma_p + \varepsilon_{it} \quad (2)$$

where $\mathbb{1}(Beach_i)$ is a dummy variable that takes on a value of one if a municipality has a beach, and zero otherwise.

Finally, I am able to estimate the second stage equation to measure the causal relation of interest:

$$y_{i,2019} = \gamma \Delta Tourism\widehat{pc_{i,1}}965-2019 + X_{it}'\theta + \pi_p + \xi_{it} \quad (3)$$

where $\Delta Tourism\widehat{pc_{i,1}}965-2019$ is the predicted value of the growth in tourism per capita obtained after from expression (2). Two fundamental assumptions underlie Specification (3): (a) the influence of a beach on municipality-level economic outcomes is confined to its impact on local tourism activity, and (b) the beach's effect on tourism activity is primarily limited to the municipality in which it is situated.

The identifying assumption in specification (3) is that the presence of a beach affects municipality-level economic outcomes relative to other municipalities only through its impact on local tourism activity. The main coefficient of interest, β, captures the marginal effects of an increase in per capita tourism growth on income per capita. To ensure a meaningful interpretation of this coefficient, it is essential to control for the determinants of long-run development that correlate with tourism.

The rationale supporting assumption (a) in our context stems from the comparison of coastal and inland municipalities situated not far away. This proximity ensures a shared set of characteristics between the two groups. To verify any remaining differences, Table A3.2 examines the correlates of tourism, presenting OLS estimates derived from regressing beach and per capita tourism growth against each variable separately. Standardized-beta coefficients and corresponding p-values are also reported. Results reveal that the presence of a beach and higher per capita tourism growth correlate with various geographic and demographic characteristics measured before the tourism surge. On average, beach municipalities exhibit smaller populations, lower illiteracy rates, and reduced agricultural production compared to inland counterparts. They also feature higher altitudes, average temperatures, more extensive surface area, and less rugged terrain. These covariates are incorporated into the main specifications at baseline to mitigate omitted variable concerns. Additionally, "entropy balancing" is employed to reweight municipalities, enhancing covariate balance between the treatment and control groups.

To further ensure the validity of assumption (a), a measure of beach tourism attractiveness considers not only the presence of a beach but also its length, width, temperature, hours of sun, and precipitation. Table A3.3 displays the correlation between tourism and the variables used to construct beach attractiveness, revealing positive associations between higher beach length, width, more hours of sun, higher temperatures, and increased tourism levels.

In addressing assumption (b), placebo analyses examine the impact of beaches on the prior decadal evolution of municipal population, the only crucial outcome variable with available historical data. The analysis excludes municipalities in large urban areas (with over 50,000 residents) to mitigate concerns related to higher tourism attractiveness and economic activity in proximity to major economic centers.

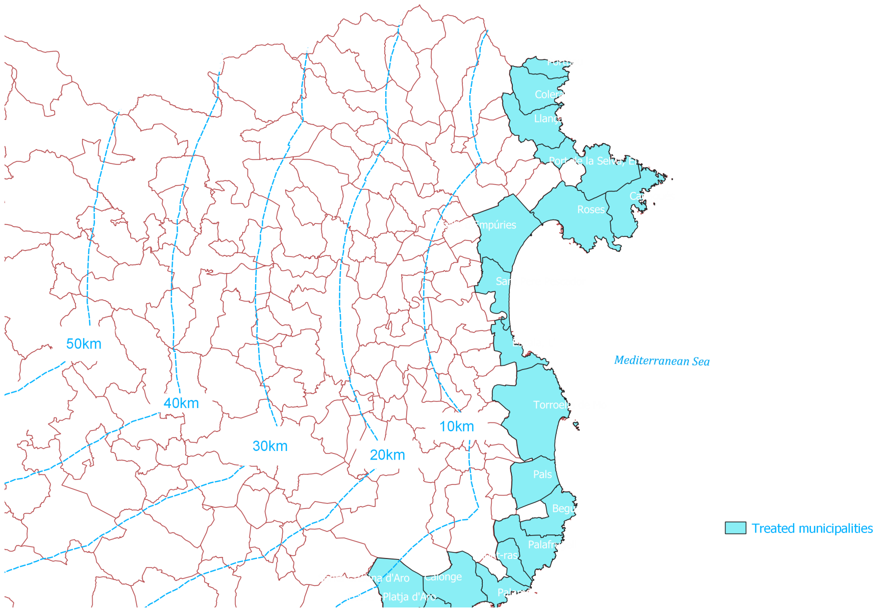
To confirm assumption (b), municipalities within 10km of the treated one are excluded from the control group. Robustness checks involve altering the distance cutoff for excluding nearby municipalities in Figure 3.3, where treated municipalities are depicted in blue and control municipalities in white. Various robustness checks explore modifying the control group by adjusting the distance cutoff for nearby municipality exclusions.

3.5.2. Second wave of tourism development from 1996-2019: shift-share analysis

The paper employs a second empirical approach to investigate the medium-term causal impact of tourism. This method involves analyzing a subsequent wave of tourism development in Spain during the late 1990s and early 2000s. This period witnessed notable changes in the tourism industry, including enhanced air accessibility and the introduction of new accommodation options, significantly influencing the influx of tourists to Spain. Given that the primary draw for tourism in Spain is sun-and-beach tourism, the analysis specifically centers on a sample of 437 municipalities with beaches.

To assess the effects of an exogenous rise in tourist numbers, an effective approach

Figure 3.3.: Location of treated and control municipalities in coastal provinces



Notes: Treated municipalities (colored in blue) have beaches, while control municipalities (colored in white) do not. Distance cutoffs are applied for excluding nearby municipalities in the robustness section. This map serves as a sample representation of the distribution of treated and control municipalities along the coastal provinces.

involves estimating the impact of tourism activity at the municipality level. This can be achieved by utilizing the per capita number of overnight stays in tourist accommodations, a commonly employed measure of tourism activity. The estimation of the baseline specification is as follows:

$$y_{i,2019} = \alpha + \beta \Delta OvernightStays_{i,1997-2019} + X'_{i,0}\beta + \gamma_p + \varepsilon_{it} \quad (4)$$

where $y_{i,2019}$ is the outcome of interest in per capita terms (e.g. log per capita gross income) in municipality i in t. The variable $\Delta OvernightStays_{i,1997-2019}$ measures the change in the per capita number of overnight stays before and after the second wave of tourism development in municipality i. The vector $X'_{i,0}$, includes a number of municipal-level controls that I described in the previous section.

This strategy comes with certain limitations, namely the availability of overnight stay data for a restricted number of beach municipalities, totaling only 96 out of the 437 considered. Moreover, it's crucial to acknowledge that tourists don't select destinations randomly; rather, they are drawn to areas that are with specific characteristics. Consequently, a simplistic comparison between high and low overnight stays in beach municipalities may yield a biased estimate of the tourism impact.

To overcome the limitations associated with using per capita overnight stays as a proxy for tourism activity, a shift-share analysis is employed to examine the effects of an exogenous increase in tourist numbers. This method combines information about the distribution of overnight stays by source country at the onset of the period with information about the national growth rate of overnight stays by source country. Because we do not have detailed information of overnight stays by source country at the municipality level, we rely on information regarding the number of residents of such nationalities in each municipality in the base year.

The shift-share analysis leverages early settlement patterns of residents from tourist countries, using this information to forecast the current distribution of tourists. Essentially, the shift-share analysis operates on the premise that tourists might transition into becoming residents in a municipality after visiting, potentially extending their stay. Furthermore, tourists are more inclined to visit areas in Spain where people from their home country reside.This analytical approach is closely aligned with a common instrument in immigration literature, which relies on historical settlement patterns of immigrants by country of origin to predict the current geographic distribution of the immigrant population (Cortes, 2008; Sá, 2015).. The rationale is akin to immigration dynamics, where networks play a pivotal role in influencing the location choices of new immigrants, aiding in the job search process and assimilation into a new culture (Card, 2003; Jaeger et al., 2018, among others). Similarly, tourists may depend on information and recommendations from fellow nationals when deciding on travel destinations or on property purchases.

In Table A3.4 in the appendix, it is evident that eleven nationalities consistently account for over 70% of tourist inflows to Spain across all observed years. Given this observation, my initial analysis focuses on examining the correlation between the number of residents by nationality at the municipality level and the corresponding number of overnight stays by nationality. This investigation is conducted for the 96 municipalities for which overnight stay data is available, and Figure A3.1 in the appendix visually presents the correlation. Notably, the figure illustrates a robust and statistically significant correlation between the number of overnight stays by nationality and the presence of residents from tourist source countries in a municipality.

Building on this correlation, I utilize the share of residents of nationality $j$from a tourist country in each beach municipality at time t_0 (that is in 1997) and the rate of growth of overnight stays by country at the national level to forecast the increase in the number of overnight stays from time t_0 to time t (that is during the period 1997-2019). To achieve this, I construct the shift-share instruments in the following manner. First, the forecasted increase in the number of overnight stays per capita in municipality i can be expressed as a summation of the forecasted increase for each source country j:

$$\widehat{\Delta O}_{i,t-t_0} = \sum_j \widehat{\Delta O}_{i,j,t-t_0} \quad (5a)$$

Second, the increase for each source country j in municipality i can be expressed as the product of the (estimated) share of overnight stays from this country going to municipality i in t0 and the growth of overnight stays from this country at the national level:

$$\Delta\widehat{O}_{i,j,t-t_0} = \left(\frac{\hat{O}_{i,j,t_0}}{\hat{O}_{i,j,t-t_0}}\right) \times \widehat{\Delta O}_{j,t-t_0} \quad (5b)$$

Third, because we do not have information on the distribution of overnight stays by country and municipality, we proxy the municipal share of overnight stays for each source country with the municipal share of residents from this country. We have discussed the plausibility of this assumption above.That is, we assume:

$$\frac{\hat{O}_{i,j,t_0}}{\hat{O}_{j,t_0}} \approx \frac{T_{i,j,t_0}}{T_{j,t_0}} \quad (5c)$$

Now, after plugging (5b) and (5c) into (5a), and dividing both sides of the equation by the population of the municipality in t_0, we get:

$$\Delta\widehat{Overnight}\ stays\ pc_{i,t-t_0} =$$

$$\sum_j Tourist\ residents\ pc_{i,j,t_0} \times \left(\frac{\Delta Overnight\ stays_{j,t-t_0}}{Overnight\ stays_{j,t_0}}\right) \times \left(\frac{Over.\ stays_{j,t_0}}{Tourist\ res._{j,t_0}}\right) \quad (5d)$$

This expression says that the increase in the number of overnight stays per capita can be predicted by a summation over all tourist source countries of the product between three terms: the number of tourist residents per capita from this country in the municipality, the national growth rate of overnight stays for this country, and the national ratio between overnight stays and number of tourist residents for this country.This computation is feasible with the available information. We will use information for nine key nationalities—Germans, Britons, French, Belgians, Dutch, Italians, Portuguese, Japanese, and Americans. These nationalities hold significant importance in terms of magnitude, and data for them is available throughout the entire study period.

The estimation now advances by estimating the following reduced-form equation:

$$y_{i,2019} = \alpha\Delta\widehat{Overnight}\ stays\ pc_{i,1997-2019} + X'_{i,0}\beta + \gamma_p + \varepsilon_{it} \quad (6)$$

Notice that one advantage of the above shift-share specification is that it can be estimated for all 437 beach municipalities, and not just for the ones with information regarding overnight stays. Using the 437 allows a fairer comparison of the result of this analysis with those of the first analysis performed in the paper and increases the external validity of the exercise. It is true, however, that the first-stage equation and, therefore also the 2SLS equation, can only be estimated with the reduced sample of 96 municipalities for which we do have that information. However, as we will show later, the coefficient of the first stage (estimated with these 96 municipalities) is very close to one, telling us that the coefficient of the reduced form and of the 2SLS should be very similar. Nevertheless, despite we are focusing for the main analysis in the larger sample, we will also report the results obtained in the smaller one, which are, by the way, very similar. This will also allow us to compare the 2SLS results with the OLS ones, that can only be obtained with the smaller sample.

I do some additional checks to validate the shift-share analysis. As argued by Goldsmith-Pinkham et al. (2018), in the shift-share framework, identification primarily stems from the 'shares.' The 'share' component imparts predictive power to the instrument by leveraging the tendency of new tourists from a specific nationality to choose municipalities where more residents from their country are settled. If the initial shares of residents from tourist countries across municipalities are uncorrelated with current changes in the outcomes of interest, this identification strategy effectively isolates the causal effect of actual tourism inflows on the variables in question.

By concentrating on the period immediately following the exogenous changes in regulation, we can more confidently attribute any alterations in tourism inflows to the exogenous shock rather than pre-existing developments in the municipalities. To ensure the validity and robustness of the shift-share analysis, I conduct several placebo tests in the subsequent section. These tests scrutinize the correlation between the shift-share and local income and population growth before the second wave of tourism development. Additionally, following the method of Goldsmith-Pinkham et al. (2020), I calculate Rotemberg weights for each nationality to further assess the validity of the identifying variation. These weights reveal which nationalities entering the instruments are driving the results. This approach helps evaluate whether the shift-share accurately captures the causal effect of tourism inflows on the variables of interest.

As argued in Goldsmith-Pinkham et al. (2018), in the shift-share setting identification mostly comes from the 'shares'. The "share" component provides predictive power to the instrument as it exploits the fact that new tourists of a given nationality tend to go to the municipalities where more residents from their country are settled. If the initial shares of residents from tourist countries across municipalities is uncorrelated with current changes in outcomes of interest then this identification strategy identifies the causal effect of actual tourism inflows on the variables of interest. By focusing on the period immediately following the exogenous changes in regulation, we can more confidently attribute any changes in tourism inflows to the exogenous shock rather than to pre-existing developments in the municipalities. In order to ensure the validity and robustness of the shift-share analysis, I run several placebo tests in the following section. These tests examine the correlation between the shift-share and local income and population growth prior to the second wave of tourism development. In addition, I follow the method of Goldsmith-Pinkham et al. (2020) and compute Rotemberg weights for each nationality to further assess the validity of the identifying variation. These weights indicate which nationalities entering the instruments are deriving the results. By doing so, I assess whether the shift-share accurately captures the causal effect of tourism inflows on the variables of interest.

3.6. Results

The results of the empirical analysis are as follows. First, I establish a descriptive association between the growth of per capita tourism and beach availability. Leveraging the first wave of tourism development, I demonstrate the impact of tourism per capita growth on municipalities' income, supported by robustness tests. Second, by exploiting the second wave of tourism development, I showcase the impact of overnight stays growth on municipalities' income, along with corresponding robustness tests. Both identification strategies yield similar results, indicating a significant impact of tourism specialization on the long-term economic outcomes of municipalities in Spain.

3.6.1. Long-term analysis from 1965-2019: beach amenities

Main results. The initial focus is on the early stages of tourism development in the 1960s, particularly examining the role of beach amenities in Spanish municipalities. The core findings are summarized in Table 3.1.

In Panel B, Columns 5 and 6 of Table 3.1, the first-stage regression results indicate a robust correlation between the presence of a beach and per capita tourism growth. Specifically, municipalities with beaches show, on average, 1.6 times higher per capita tourism growth than inland counterparts within coastal provinces. The F-statistic tests validate the strength of the beach presence as an instrumental variable in this context.

To account for potential pre-tourism developmental differences, several socio-economic and geographical factors are controlled for, including population size in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, initial levels of tourism per capita, along with variables like soil quality, altitude, ruggedness, surface area, temperature, rainfall, and sunshine hours. Province-level fixed effects and clustering are also employed where indicated.

Panel A of Table 3.1 presents the impact of per capita tourism growth on local income levels. Columns 1 and 2 show ordinary least squares (OLS) estimates for 2019 income levels, while columns 3 and 4 report the reduced-form impact of beach presence in 1965 on 2019 income levels. Columns 5 and 6 offer instrumental variable (IV) estimates, where per capita tourism growth is instrumented using beach presence.

The results reveal that beach presence correlates with lower income levels in 2019, a trend echoed in both the reduced-form and IV estimates. The IV estimates, particularly in columns 5 and 6, are more pronounced than the OLS results, suggesting the effectiveness of the instrumental variable approach in addressing endogeneity issues. Specifically, the results in column 6 reveal that municipalities experiencing the median increase in tourism per capita during the period 1965-2019 have 22,2% lower per capita income in 2019.[7] Hence, it can be conclude that tourism specialization exerts a significant, and negative, effect on municipalities' per capita income.

Table 3.1.: Impact of tourism growth on income per capita: long-term analysis (1960-2019)

	Panel A. Dep. variable: Income per capita 2019					
	OLS	OLS	RF	RF	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Per capita tourism	-0.004*	-0.004***			-0.051***	-0.085***
growth (1965-2019)	(0.010)	(0.010)			(0.008)	(0.018)
$\mathbb{1}$(Beach)			-0.071***	-0.137***		
			(0.015)	(0.025)		
Mean dep. var.	9.49	9.49	9.49	9.49	9.49	9.49
	Panel B. First stage: Dep. variable: per capita tourism growth					
$\mathbb{1}$(Beach)					2.093***	1.598***
Kleibergen-Paap rk					38.64	35.890
LM F-stat.					[16.38]	[16.38]
Observations	1,196	1,196	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y	Y	Y
Province FE	N	Y	N	Y	N	Y

Notes: (1) Panel A reports the estimated effect of per capita tourism growth on income level using three different regression methods: OLS, RF, and IV. The dependent variable in all cases is the log of income in 2019. Panel B reports the first-stage regression results, where the independent variable is $\mathbb{1}$(Beach), a dummy variable equal to one if the municipality has a beach and zero otherwise. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

[7] The median increase in tourism per capita during the 1965-2019 period is 2.62.

In furthering the analysis, I introduce the variable *Beach attractiveness*, which quantifies the quality and allure of a beach. This variable takes into account several factors such as the length and width of the beach, derived from 1957 aerial photos, as well as climatic elements like temperature, sunshine hours, and precipitation. The interaction of *Beach attractiveness* with the binary variable *Beach* is then incorporated into Equation (2).

Table A3.5 presents the results of this extended analysis, which incrementally integrates various beach features across five columns. The first column accounts for the standardized beach length in kilometers, followed by the addition of sunshine hours in the second column. The third column includes the beach width in kilometers, while the fourth and fifth columns add temperature and precipitation, respectively.

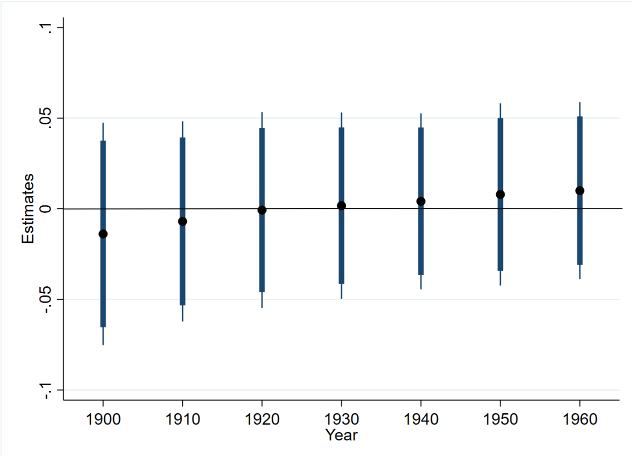
The instrumental variable (IV) results from this analysis suggest a nuanced conclusion: not only does the presence of a beach but also its favorable geographical and climatic characteristics, contribute to a negative impact on the income levels of municipalities. The first-stage F statistic, when incorporating these additional beach features, is observed to be slightly lower compared to the analysis focusing solely on the presence of a beach. This indicates that while beach presence remains a primary determinant of tourism attractiveness, other factors such as beach length and hours of sunshine, as highlighted in column 2, also significantly influence tourism attractiveness and, consequently, income levels.

Robustness. While Table 3.1 indicates a strong negative effect of tourism on local income levels, a potential concern arises regarding the direct influence of beach and related geographical and climatological features (such as extended beach lengths, high temperatures, more sun hours) on the local economy. These features could potentially affect the residential choices of Spanish residents, not just through their impact on tourism development, but also as direct amenities influencing employment and population dynamics relative to inland municipalities.

To address this concern, Figure 3.4 presents a placebo falsification test using the same sample of municipalities, but during a period preceding the emergence of beach tourism as a significant economic factor in Spain. This analysis utilizes a long time series of population census data spanning from 1900 to 1960. However, due to data limitations at the municipal level before 1960, the historical census data is restricted to municipality population figures.

The results displayed in Figure 3.4, which report the instrumental variable (IV) estimates, show insignificant point estimates regarding the effect on municipality populations before 1960. This finding suggests that the presence of a beach, used as an instrument for per capita tourism growth, is unlikely to be capturing locational fundamentals that directly influence residential choices along the coastline in a significant manner. Therefore, the negative impact of tourism on income levels observed in the study is less likely to be confounded by direct amenity effects unrelated to tourism.

Figure 3.4.: Placebo falsification test: impact of tourism growth on municipality populations prior to the tourism boom (1900-1960)



Notes: (1) The graph shows the IV estimation of the impact of per capita tourism growth instrumented by the presence of a beach on population. The dependent variable is the population expressed in logs. (2) The standard errors used in the estimation are robust and clustered at the province level.

Furthermore, the analysis presented in Table 3.1 is conducted with a carefully selected sample. To mitigate the influence of commuting patterns and ensure comparability in initial socio-economic and geographical characteristics, the control group (interior municipalities) excludes those located within less than 10 km or more than 50 km from beach municipalities.

In assessing the robustness of this distance criterion, Figure A3.2 in the appendix presents a sensitivity analysis. This analysis explores the impact of tourism growth on local income by adjusting the distance cutoffs for the selection of control municipalities. The treated municipalities in this analysis are consistently those with a beach, while the control municipalities are classified into six different distance ranges: (1) 20-50 km, (2) 20-40 km, (3) 20-30 km, (4) 30-50 km, (5) 30-40 km, and (6) 40-50 km.

This sensitivity analysis demonstrates consistent results. Across all specifications, regardless of the distance cutoffs applied to the control municipalities, the impact of tourism on local income remains consistent. Specifically, the analysis indicates that per capita tourism growth consistently has a negative impact on income levels, with the effect's magnitude being similar across all six distance ranges. This result suggests that the observed impact of tourism on income is robust and not significantly influenced by the selection of control municipalities based on varying distance criteria.

In Figure A3.3 in the appendix, another robustness check is performed, examining the impact of tourism growth on 2019 income levels by varying the definitions of the treated groups. This analysis reveals that when municipalities within 10 km of beach municipalities are considered the treated group, and those more than 20 km away as controls, a modest negative impact of per capita tourism growth on income is observed. However, in other configurations, where larger distances define the treated group, the effects are found to be insignificant.

To further understand these dynamics, Table A3.7 investigates the economic structures of municipalities situated around 10 km from beach municipalities compared to those more than 20 km away. The lack of significant differences in the economic structures of these groups, as indicated by the results in Table A3.7, suggests a uniformity in economic composition across these distances.

A plausible explanation for the observed negative effect in the 10 km range is the presence of commuting patterns. Workers residing in municipalities close to beach areas likely commute to work in beach municipalities, thus experiencing similar economic impacts as residents living directly in the beach municipalities. Consequently, the negative influence of tourism growth on income levels in municipalities within a 10 km radius of beach municipalities could predominantly affect these commuting workers. This observation reinforces the rationale behind excluding municipalities within 10 km of beach municipalities from the control group in the initial analysis.

In addition to the above analyses, I also conducted a sub-sample analysis categorizing the main sample into five groups based on surface area. This segmentation aimed to investigate how the dynamics between tourism specialization and income levels differ across municipalities of varying sizes. The findings, presented in Figure A3.4, offer nuanced insights into this relationship.

The analysis reveals that the negative impact of per capita tourism growth on income levels is both significant and negative for municipalities classified as medium, large, a combined category of medium and large, and across all municipalities collectively. However, an interesting deviation is observed in small municipalities (with a surface area up to 20 km^2), where the negative impact of per capita tourism growth, although present, does not reach statistical significance.

This pattern suggests that larger and medium-sized beach municipalities, likely with more substantial tourism activities and a more developed tourism industry, experience a more pronounced and statistically significant impact on income. In contrast, smaller beach municipalities, which may have lower levels of tourism activity and limited capacity to accommodate large numbers of tourists, seem to experience a less pronounced negative impact on local income levels.

Finally, the study examines whether the impact of tourism on income levels was more pronounced in the initial period (1965-1996) or in the later period (1996-2019). A closer analysis of the data shows that the median growth in per capita tourism from 1965 to 1996 was 1.5. This growth correlates with a 5.7% decrease in income levels by the end of 1996. In contrast, the median growth in per capita tourism for the later period from 1996 to 2019 was notably higher, at 4.16. This increased level of tourism growth is associated with a more substantial 16.6% reduction in income levels by 2019. These results highlight a more pronounced negative impact of tourism on income levels in beach municipalities during the later period (1996-2019) compared to the earlier period (1965-1996). This trend suggests that as tourism activity intensified in beach municipalities over time, its detrimental effects on local income levels became increasingly severe.

3.6.2. Second wave of tourism development from 1996-2019: shift-share analysis

Main results. I now present the results focusing on the second period of tourism development (1996-2019). Ideally, the analysis would begin with first-stage regressions showing the relationship between predicted and actual overnight stays. However, due to data constraints, such information is available for only a subset of 96 out of 437 beach municipalities.

Before delving into the reduced form results for all 437 beach municipalities, I examine the association between predicted and actual overnight stays for this subsample of 96 municipalities. As illustrated in Figure A3.5, a positive and significant correlation is observed between these two variables. Moreover, Table A3.6 presents the first-stage regression results, demonstrates high correlation between the predicted growth in overnight stays, as outlined in Equation (5), and the observed change. The coefficients are consistently close to 1 across all specifications, with standard F-statistic tests confirming the strength of the instrument in the context of these specifications.

Table 3.2.: Impact of overnight stays growth on income per capita: medium-term analysis (1996-2019)

	Income level				Population level	
	Reduced form		Placebos			
	2019	2019	1990	1990	1990	1990
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Predicted overnight stays (1996-2019)	-0.057*** (0.016)	-0.071*** (-0.071)	-0.002 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.000 (0.000)
Mean dep. var.	9.49	9.49	9.49	9.49	9.49	9.49
Observations	437	437	437	437	437	437
All Controls	Y	Y	Y	Y	Y	Y
Province FE	N	Y	N	Y	N	Y

Notes: (1) The table shows the reduced form estimates of the impact of tourism growth measured by the predicted number of overnight stays and instrumented by the shift-share explained in section 4 on income and population levels. The unit of observation is the municipality. Columns (1) and (2) show the impact of an increase in the number of overnight stays during the period 1997-2019 on income level in 2019. Columns (3) and (4) show the impact of an increase in the number of overnight stays during the period 1997-2019 on income level in 1990. Columns (5) and (6) show the impact of an increase in the number of overnight stays during the period 1997-2019 on population level in 1990. (2) Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

With the relationship between predicted and actual overnight stays established for the subsample of Spanish beach municipalities, I then explore the medium-term impact of per capita tourism on local income growth. I estimate specification (6) with log income per capita in 2019 on the left-hand side, and present the Reduced Form results in Table 3.2.

The Reduced Form estimation indicates a negative association between the increase in overnight stays following the second period of tourism development and per capita income. Specifically, column (2) in Table 3.2 reveals that municipalities asof 2019, a municipality experiencing a median increase in tourism per capita over the period has a per capita income a 21,5% lower. [8]

Consequently, the shift-share analysis corroborates that tourism specialization has a significant and negative effect on municipalities' per capita income. Table A3.6 in the appendix further supports this conclusion, presenting the OLS, reduced form, and IV regression results for the impact of tourism measured by the actual number of overnight stays, using the subsample of 96 municipalities. These coefficients align with the findings in Table 3.2, consistently indicating a negative effect of tourism specialization on income per capita.

[8]The median per capita tourism growth during the period 1997-2019 is 3.03.

Robustness. Table 3.2 illustrates a strong negative effect of the increase in overnight hotel stays on local income levels in municipalities with beaches. A key concern in this analysis is the validity of the shift-share instrument. This instrument's identification relies on initial shares of residents by nationality, which are presumed to reflect differential, exogenous exposure to the second wave of tourism development.

However, since these predetermined shares are equilibrium outcomes influenced by tourism attractiveness, there's a possibility they could correlate with income levels during that period. The literature on shift-share analysis, particularly Goldsmith-Pinkham et al., (2020), posits that the validity of this approach hinges on the assumption that initial shares are exogenous to changes in income, rather than to the initial income levels themselves.

To test this crucial assumption, I follow the methodology proposed by Goldsmith-Pinkham et al. (2020), beginning with the computation of Rotemberg weights for different nationalities, based on the framework established by Rotemberg (1983) and Andrews et al. (2017). These weights help identify which nationalities predominantly influence the shift-share results. In this analysis, the five most significant nationalities are Germans, Britons, French, Belgians, and Portuguese.

Following Goldsmith-Pinkham et al.'s (2020) recommendation, I examine the correlation between the initial nationality shares and potential confounders. The confounders considered include population growth (1997–2019), education level in 1990, unemployment level in 1990, and agricultural and manufacturing production in 1990. The results of this correlation analysis are reported in Table A3.9 in the appendix. Reassuringly, this robustness check indicates that the initial nationality shares are not significantly related to these potential confounding factors.

An essential aspect of verifying the validity of the shift-share analysis is to investigate its correlation with local income growth prior to the onset of the second wave of tourism development. This examination is crucial to ensure that the shift-share instrument is not influenced by pre-existing economic conditions.

In Table 3.2, the correlation between local income levels in 1990 (for municipalities with beaches) and the predicted growth in overnight stays from 1997 to 2019 (calculated using the shift-share method) is presented. The results suggest that there is no significant correlation between the shift-share instrument and the pre-treatment local income levels. This finding is critical as it implies that the shift-share analysis is not capturing pre-existing income trends that could confound the results. Furthermore, the study extends this examination to include population growth. The correlations between the shift-share instrument and population growth, as indicated in columns (3) and (4), are also found to be statistically insignificant. This lack of significant correlation adds an additional layer of validation to the shift-share analysis. It supports the conclusion that the shift-share instrument is not confounded by pre-treatment population growth trends.

3.6.3. Mechanisms

To understand the mechanisms driving the negative impact of tourism on income levels, I gathered detailed municipal-level data encompassing aspects like worker affiliation, contract type, and education level. The data reveals that an increase in per capita tourism growth correlates with an increase in temporary contracts and a decline in educational attainment. It also indicates a shift in the workforce from industrial and agricultural sectors to those associated with tourism.

Industrial Mix. The first mechanism explored is the economic structure of the municipality. Table 3.3 provides insights into tourism's impact on employment across various sectors, including manufacturing, construction, real estate, hotels, food-service, and agriculture. Unfortunately, sector-specific employment data at the municipal level is available only from 1990 onwards. In Table 3.3, I present findings using the first empirical strategy, examining the impact of per capita tourism growth on the level of outcomes in 2019. Per capita tourism growth is instrumented with $\mathbb{1}$(Beach), a dummy variable indicating the presence of a beach in the municipality.

Table 3.3.: Effects of tourism specialization on employment by sector: 1965-2019

	Employment by industry			
	Agriculture	Manufacturing	Construction and real estate	Hotels and foodservice
	IV	IV	IV	IV
	(1)	(2)	(3)	(4)
Per capita tourism growth (1965-2019)	-0.179* (0.221)	-0.621*** (0.192)	0.787** (0.745)	0.464*** (0.155)
Mean dep. var	14.384	18.336	9.077	15.968
Kleibergen-Paap rk LM F-stat.	19.343 [16.38]	19.343 [16.38]	19.343 [16.38]	19.343 [16.38]
Observations	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: (1) The table shows the estimates of the effect of per capita tourism growth on employment shares in different sectors using the IV method. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

The instrumental variable (IV) estimates suggest that an increase in per capita tourism growth is associated with a rise in employment in construction, real estate, hotels, and food services, and a decline in manufacturing and agricultural employment. This pattern highlights the tourism sector's role in reallocating resources toward service activities and away from traditionally traded sectors like industry and agriculture.

More specifically, the findings indicate that municipalities with higher per capita tourism growth in 2019 experienced a 17.9% decrease in agricultural employment and a 62% decrease in manufacturing employment. Conversely, there was a 78% increase in construction and real estate employment and a 46% increase in hotel and food-service related employment.

The increase in tourism activity has led to a resource shift from industry to services, as noted by Copeland (1991). According to the National Spanish Institute, in 2019, manufacturing workers earned 34 percent more gross annual income than workers in the tourism sector. This finding aligns with several studies in developed countries that have consistently found tourism workers to be among the lowest paid in any industry (Riley, Ladkin, and Szivas, 2002; Muñoz-Bullón, 2009; Santos and Varejao, 2007; Dogru et al., 2019).

The results obtained from the first empirical strategy are further supported by the reduced form results of the second empirical strategy, utilizing the shift-share analysis. In this approach, per capita tourism growth is measured using predicted overnight stays and is instrumented through the shift-share methodology outlined in section 4. With employment data by sector available at the municipal level from 1990, we are able to assess the growth in employment by sector between 1990 and 2019.

Table 3.4 presents these findings for the 437 municipalities with beaches. The results indicate that doubling the number of overnight stays during this period led to significant shifts in sectoral employment: there was a 27% decrease in agricultural employment and a 25% decrease in manufacturing employment. Concurrently, there was a notable increase in employment in tourism-related sectors, with a 26% rise in construction and real estate services, and an 11% increase in hotels and food-services employment.

Table 3.4.: Sectoral employment impacts of overnight stays growth: 1990-2019

	Employment change by industry			
	ΔAgriculture	ΔManufacturing	ΔConstruction and realestate	ΔHotels and foodservice
	RF	RF	RF	RF
	(1)	(2)	(3)	(4)
Predicted overnight stays	-0.272**	-0.253*	0.262**	0.112**
growth (1996-2019)	(0.106)	(0.09)	(0.102)	(0.052)
Observations	437	437	437	437
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: (1) The table shows the RF estimates of the effect of per capita tourism growth on employment shares growth in different sectors. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Labor Market. The second mechanism investigated to understand the observed negative impact on overall income levels is the nature of employment contracts in the tourism sector. The hypothesis posits that when a municipality specializes in tourism, there might be a shift toward more temporary contracts and fewer stable jobs, reflecting the seasonal nature of tourism employment.

To examine this hypothesis, Table 3.5 presents the results of the impact of tourism on the share of temporary contracts in 2019, utilizing both empirical strategies. Column (1) shows the results from the first empirical strategy, where per capita tourism is the independent variable, instrumented by the presence of a beach in a municipality. Meanwhile, column (2) aligns with the second empirical strategy, with the growth in overnight stays as the independent variable.

The findings indicate that municipalities that experienced a median increase in tourism per capita over the analysis period exhibited a significant rise in the proportion of temporary contracts. When applying the first empirical approach, which is based on accessibility to beach amenities, there was a 46% increase in temporary contracts. Conversely, using the second approach, grounded in shift-share analysis, the increase was 17%. This trend aligns with the observations of González and Surovtseva (2020), who noted that a shift toward tourism-related employment often entails a transition from permanent to more temporary contracts, leading to reduced employment stability. This shift in the nature of employment contracts can be attributed to the inherent seasonality of the tourism industry, where peak seasons necessitate a surge in temporary staffing, followed by periods of reduced employment demand during off-peak times.

Table 3.5.: Tourism's impact on employment contract types

	Share of temporary contracts 2019	
	Beach amenities analysis	Shift-share analysis
	IV	RF
Per capita tourism growth (1965-2019)	0.177**	
	(0.075)	
Predicted overnight stays growth (1996-2019)		0.058*
		(0.015)
Mean Dep. Var	0.943	0.943
Observations	1,196	437
All Controls	Y	Y
Province FE	Y	Y

Notes: (1) The table shows estimates of tourism on the share of temporary contracts. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Education Level. Another mechanism potentially contributing to lower income levels in tourism-specialized municipalities is the lower educational attainment among the workforce. This trend could result from the opportunity cost associated with higher education and the decreasing perceived returns from such education. Table 3.6 presents estimates of tourism's impact on education across different specifications.

Columns (1) to (3) utilize per capita tourism growth, instrumented by the presence of a beach in a municipality, as the independent variable. In contrast, column (4) employs growth in overnight stays as the variable of interest. The dependent variable in column (1) is the level of tertiary education in 2011, while in column (2), it is for 2001. Column (3) examines the secondary education level in 1990. Lastly, column (4) assesses the growth in tertiary education between 1991 and 2011.

The analysis reveals that municipalities with more intensive tourism specialization tend to have lower education levels in the years 2011, 2001, and 1990, as shown in columns (1) to (3). Specifically, a higher increase in per capita tourism growth correlates with lower tertiary education levels among the 26-35 age population in 2011 and 2001. A similar trend is observed for secondary education in 1990.[9]

Utilizing the second empirical strategy, column (4) examines the impact of tourism growth on tertiary education growth. It is observed that in municipalities with beaches, a doubling of the number of overnight stays from 1997 to 2019 is associated with a 1% lower growth rate in tertiary education. This finding aligns with Oliver

[9] Note that in 2011 and 2001, the focus is on tertiary education, as secondary education became mandatory in Spain after 1990.

and Villalonga (2019), who discovered that Spanish regions specializing in sectors requiring a high level of unskilled labor tend to discourage students from continuing their education, leading to higher dropout rates and fewer students completing non-compulsory secondary education stages. In contrast, regions focusing on sectors demanding medium or highly skilled workers encourage young students to remain in the education system, as employment opportunities in these areas often require a minimum level of education.

Table 3.6.: Tourism's impact on educational attainment across years

	Beach amenities analysis			Shift-share analysis
	College education		High school education	Δ Tertiary education
	2011	2001	1990	1991-2011
	IV	IV	IV	RF
	(1)	(2)	(3)	(4)
Per capita tourism growth (1965-2019)	-0.018***	-0.013***	-0.013***	
	(0.006)	(0.004)	(0.015)	
Predicted overnight stays growth (1996-2019)				-0.001**
				(0.004)
Mean Dep. Var	0.150	0.196	0.455	0.05
Kleinbergen-Paap rk LM F-stat.	17.82	14.41	11.92	
Observations	1,196	1,196	1,196	437
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: (1) The table presents estimates of the impact of tourism on education. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

3.7. Conclusions

Sun and beach tourism plays a crucial role in the economy of many countries. However, this economic sector is expected to suffer significant damage due to climate change, such as the projected loss of beaches, thermal comfort, water restrictions, and extreme weather events (Becken, 2020). Nevertheless, governments continue to invest in and promote tourism as a key sector in the national economy. The economic implications of tourism specialization are mixed and the economic evidence of the long-term impact continues to be scarce. With this paper, I aim to fill this current gap in the literature by analyzing the lasting impact of tourism specialization on per capita income.

In this study, I focus on Spain, one of the most popular tourist destinations in the world, where the tourism sector accounts for a substantial share of total employment, especially in coastal regions. To address the endogeneity underlying tourism specialization and economic development, I employ two distinct empirical strategies — one based on geographical and natural amenities and the other leveraging a shift-share analysis.

The first leverages geological variation along the Spanish coastline to construct instrumental variables for tourism specialization, including the existence and attributes of beaches and other weather features. This approach confirms that accessible natural resources, particularly beaches, are instrumental in determining the causal impact of tourism on income levels. The second methodology, a shift-share analysis, utilizes the distribution of residents from tourist-source countries in beach municipalities, capturing the effects of the second wave of tourism development.

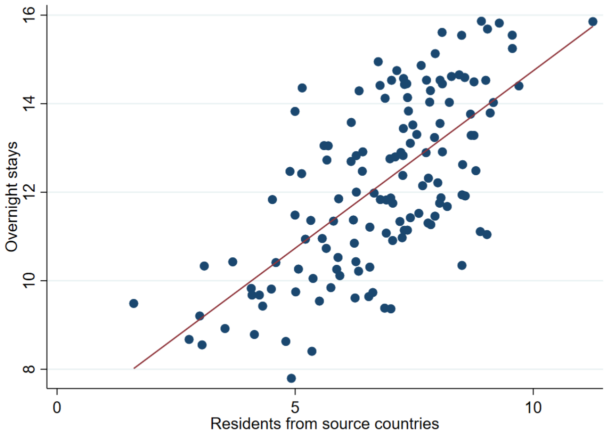
Employing these identification strategies, the results demonstrate that municipalities experiencing pronounced tourism specialization exhibit a substantial decline in per capita income. Specifically, a municipality undergoing an increase in tourism per capita, commensurate with the median level observed across the sample, witnessed a 22% reduction in per capita income by the year 2019. The study also reveals key mechanisms driving this impact: a shift towards temporary employment contracts, a reallocation of labor from industry and agriculture to tourism-related sectors, and lower educational attainment.

This study provides robust empirical evidence of the localized, long-term economic consequences of tourism activity. As current tourism policies often prioritize boosting local tourism appeal, these findings highlight the need for a nuanced understanding of tourism's long-term implications compared to other economic sectors. Future assessments of tourism's role in Spain's overall economic activity and employment should consider the indirect countrywide effects, including the chain reactions in the broader economic system triggered by the tourism industry's demand for production inputs and labor.

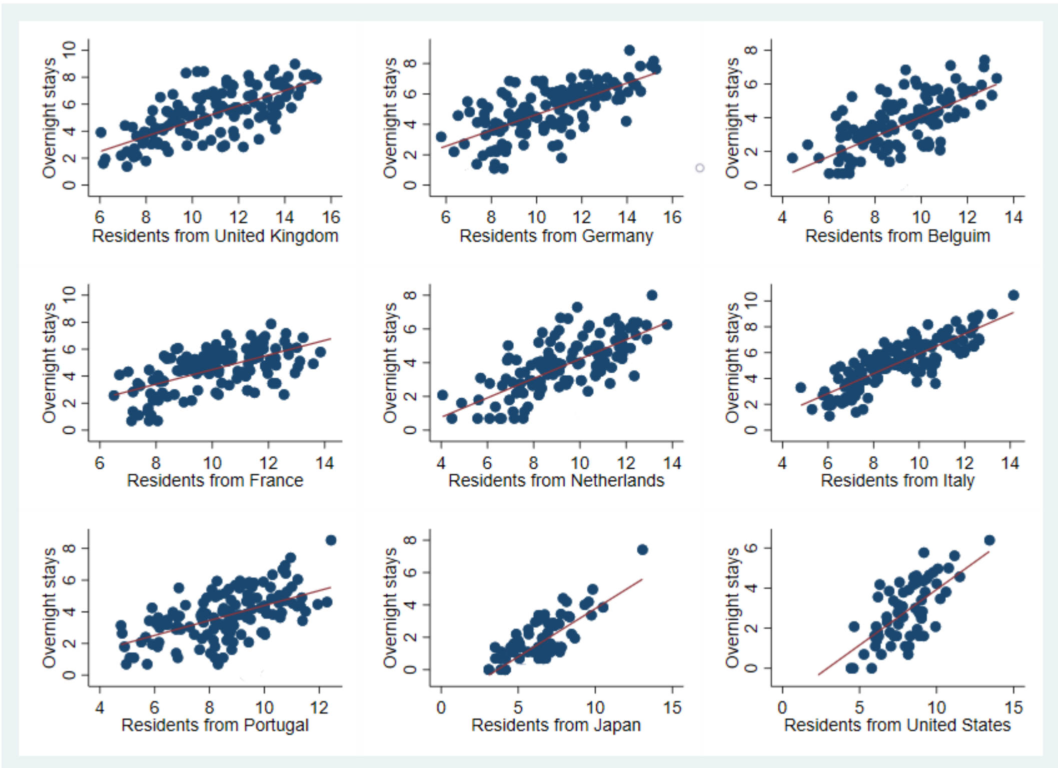
3.8. Appendix

Figure A3.1.: Correlation between overnight stays and resident nationalities in spanish municipalities

Panel A:

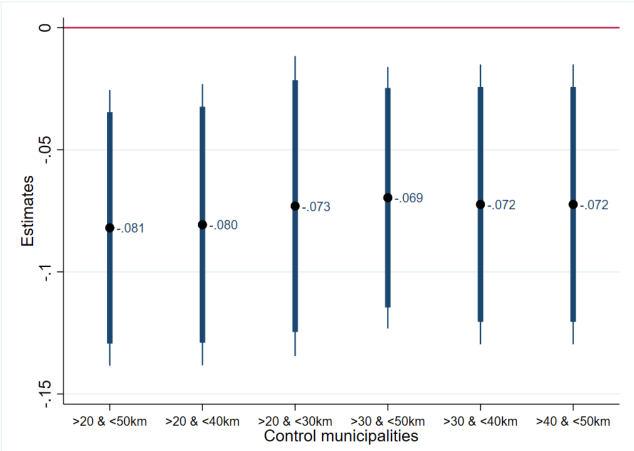


Panel B:



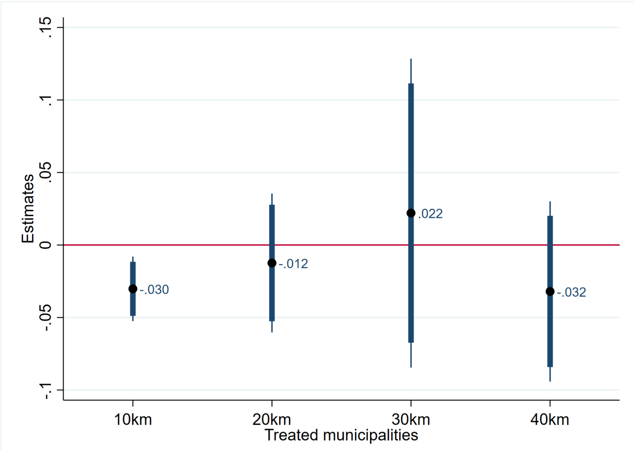
Notes: (1) This Figure shows the correlation between the per capita number of residents in a municipality and the per capita number of overnight stays in a municipal by nationality for the main tourist nationalities.

Figure A3.2.: Impact of tourism growth on income: varying distance cutoffs for control municipalities



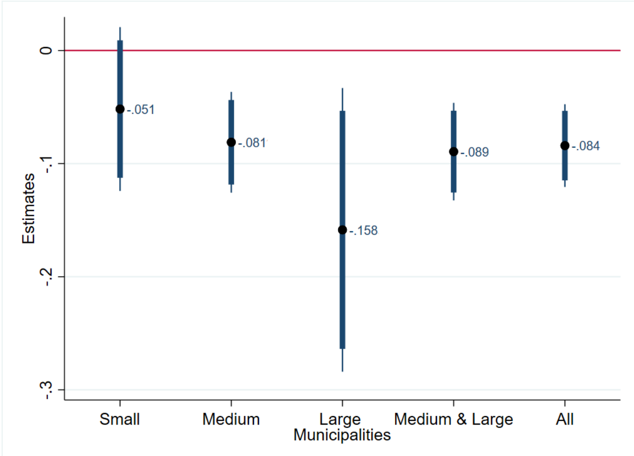
Notes: The graph presents the IV estimates of the impact of tourism growth on income level in 2019 using IV regression, where per capita tourism is instrumented by a dummy variable $\mathbb{1}$(Beach), which equals one if the municipality has a beach and zero otherwise. The unit of observation is the municipality. The graph shows the point estimates and 95% confidence intervals of the impact of tourism growth on income for different ranges of control municipalities, based on their distance to the treated (beach) municipalities. The control municipalities are classified into six distance ranges: (1) 20-50 km, (2) 20-40 km, (3) 20-30 km, (4) 30-50 km, (5) 30-40 km, and (6) 40-50 km.

Figure A3.3.: Impact of tourism growth on income: varying distance cutoffs for treated municipalities



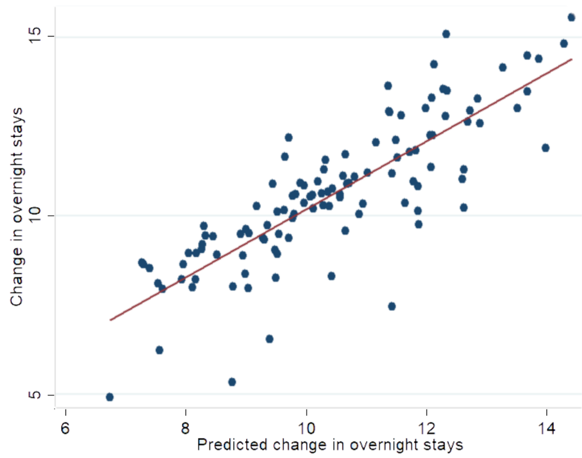
Notes: (1) This Figure presents the results of a sensitivity analysis examining the impact of tourism growth on income level in 2019 in different specifications. The analysis considers different treatment groups: (1) municipalities less than 10km far away from beach municipalities as the treated group, and municipalities >20km far away as controls; (2) municipalities 20km far away from beach municipalities as the treated group, and municipalities >30km far away as controls; (3) municipalities 30km far away from beach municipalities as the treated group, and municipalities >40km far away as controls; and (4) municipalities 40km far away from beach municipalities as the treated group, and municipalities more than 40km far away as controls.

Figure A3.4.: Impact of tourism growth on income levels: a sub-sample analysis by municipality size



Notes: (1) This graph shows the estimated impact of tourism growth on income level for different sizes of municipalities based on their surface area. The sample is split into five groups: small municipalities (with surface area up to 20 km²), medium municipalities (with surface area between 20 and 90 km²), large municipalities (with surface area over 90 km²), medium and large municipalities combined, and all municipalities. The estimates are obtained using IV regression, where per capita tourism is instrumented by a dummy variable equal to one if the municipality has a beach and zero otherwise.

Figure A3.5.: Visual correlation between predicted and realized overnight stays in beach municipalities



Notes: (1) The unit of observation is the municipality. The dependent variable is log growth of realized overnight stays for the subsample of municipalities of realized overnight stays data available. Predicted overnight stays is calculated as described in section 4. (2) Robust standard errors are in parentheses and clustered at province level when adding province fixed effects. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A3.1.: Descriptive statistics

Variables	count	mean	sd	min	max
Panel A					
Tourism per capita 1965	1196	13.82	88.56	0.00	2054.70
Tourism per capita 2019	1196	36.94	202.23	0.00	4454.32
Beach length	437	3691.88	4603.29	0.00	47320.96
Beach width	437	522.19	722.79	0.00	5873.76
Beach area	437	353507.99	668916.52	13.00	6641338.00
Share German residents 1996	437	0.01	0.02	0.00	0.18
Share Briton residents 1996	437	0.02	0.04	0.00	0.27
Share Dutch residents 1996	437	0.00	0.01	0.00	0.12
Share French residents 1996	437	0.00	0.01	0.00	0.05
Share Portuguese residents 1996	437	0.00	0.00	0.00	0.02
Share American residents 1996	437	0.00	0.00	0.00	0.07
Share Japanese residents 1996	437	0.00	0.00	0.00	0.00
Share Belgian residents 1996	437	0.00	0.01	0.00	0.07
Share Italian residents 1996	437	0.00	0.00	0.00	0.01
Panel B					
Log Predicted Income per capita 1965	1196	2.98	1.28	1.11	5.03
Log Predicted Income per capita 1990	437	2.71	0.95	1.80	4.82
Log Predicted Income per capita 1997	437	3.75	0.31	1.38	12.27
Log Income per capita 2019	1196	9.47	0.24	8.93	10.41
% Employed in manufacturing 2019	1196	15.55	15.43	0.00	84.22
% Employed in agriculture 2019	1196	18.34	19.71	0.00	80.28
% Employed in construction 2019	1196	9.14	5.69	0.00	49.44
% Employed in construction 2019	1196	9.14	5.69	0.00	49.44
% Employed in real estate 2019	1196	0.50	0.75	0.00	12.50
% Employed in hotels and foodservice 2019	1196	9.97	8.23	0.512	70.76
% Employed in ancillary services 2019	1196	2.69	3.07	0.00	55.48
% Employed in manufacturing 1990	437	1.3	0.9	0.00	9.5
% Employed in agriculture 1990	437	0.13	0.118	0.00	0.57
% Employed in construction 1990	437	12.78	11.7	0.23	85.9
% Employed in hotels and foodservice 1990	437	14.3	0.83	0.39	53.8

(Continued)

Table A3.1.: Descriptive statistics

Variables	count	mean	sd	min	max
% Secondary education age 25-36, 1991	1196	0.57	0.15	0.01	0.96
% Tertiary education age 25-36, 1991	437	0.07	0.04	0.01	0.26
% Tertiary education age 25-36, 2001	1196	0.17	0.08	0.07	0.49
% Tertiary education age 25-36, 2011	1196	0.22	0.12	0.00	0.86
% of temporary contracts	1196	0.94	0.06	0.38	1.00
% of immigrant workers	1196	0.05	0.05	0.00	0.43
Panel C					
Population 1950	1196	5138.00	5172.00	165.00	58768.00
Population 1990	437	9492.32	8570.71	57.00	50466.00
Temperature	1196	14.95	2.11	7.00	19.90
Precipitations	1196	6.75	2.76	1.06	15.57
Anual hours of sun	1196	2649.75	426.97	1600.00	3500.00
Altitude	1196	4.33	3.62	0.03	24.06
Ruggedness	1196	1.17	1.11	0.00	8.30
Surface	1196	84.13	110.25	0.73	969.01
Soil quality	1196	8.32	1.13	0.95	10.00
Agricultural production 1950	1177	0.32	0.17	0.00	0.95
Agricultural production 1990	437	0.14	0.12	0.00	0.57
Illiteracy rate 1930	1196	47.83	14.20	14.39	88.41
Illiteracy rate 1990	437	1.80	0.99	0.00	4.46

See Section 4 for a description of the datasets

Table A3.2.: Correlates of tourism

Variables	Coef. (1)	Std. Error (2)	Std. Beta. (3)	P-value (4)
Panel A: Beach				
Temperature	0.090***	(0.007)	0.281	(0.000)
Precipitations	-0.019	(0.029)	-0.022	(0.497)
Hour of sun	-0.075***	(0.011)	-0.204	(0.000)
Altitude	0.093	(0.076)	0.039	(0.224)
Ruggedness	-3.229***	(0.176)	-0.419	(0.000)
Surface	0.166*	(0.068)	-0.067	(0.015)
Soil quality	-0.046	(0.074)	-0.019	(0.531)
Share agricultural production 1962	-0.523***	(0.059)	-0.289	(0.000)
Illiteracy rate 1930	-0.034*	(0.020)	-0.050	(0.094)
Log population 1960	-0.451***	(0.052)	0.234	(0.000)
Panel B: Per capita tourism growth 1965-2019				
Temperature	0.006***	(0.001)	0.172	0.000
Precipitations	-0.023***	(0.004)	-0.221	0.000
Hours of Sun	0.004***	(0.000)	0.101	0.000
Altitude	0.014	(0.013)	0.050	0.313
Ruggedness	-0.085***	(0.029)	-0.093	0.003
Surface	0.021***	(0.006)	0.074	0.000
Soil Quality	0.019***	(0.006)	0.069	0.001
Share of Agricultural Production in 1962	-0.001	(0.003)	-0.007	0.667
Illiteracy Rate in 1930	0.005***	(0.001)	0.072	0.000
Log Population in 1960	-0.011***	(0.004)	-0.051	0.006

Notes: This table shows the correlates of Beach and per capita tourism growth. The unit of observation is the municipality. There are 1,196 observations included in each regression. Column (1) shows the point estimate for the regressor of interest. Column (2) shows the corresponding standard error. Column (3) corresponds to the standardized-beta and column (4) to the corresponding p-value.

Table A3.3.: Correlations between beach attractiveness metrics and tourism levels

VARIABLES	All beach municipalities (1)	Beach municipalities used in the analysis (2)
Beach length	0.591**	0.529*
	(0.270)	(0.291)
Beach width	0.825***	0.991***
	(0.188)	(0.207)
Temperature	0.773**	0.627*
	(0.303)	(0.339)
Precipitation	-0.228	-0.0893
	(0.263)	(0.298)
Hours of sun	0.218***	0.245***
	(0.188)	(0.165)
Observations	461	437

Notes: Each row reports the OLS estimate of regressing per capita tourism in 2019 on the variable indicated in each row separately. The unit of observation is the municipality. Column 1 shows the estimates for all municipalities with beach. Column 2 shows the estimates for the municipalities used in the main sample as explained in section 3.5. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3.4.: Major nationalities contributing to Spanish tourism

2000		2010		2019	
Germany	0.234	United Kingdom	0.190	United Kingdom	0.185
United Kingdom	0.207	Germany	0.182	Germany	0.129
France	0.092	France	0.108	France	0.106
United States	0.063	Italy	0.071	United States	0.061
Italy	0.062	United States	0.051	Italy	0.057
Netherlands	0.038	Netherlands	0.037	Netherlands	0.037
Belgium	0.035	Portugal	0.035	Portugal	0.029
Portugal	0.035	Belgium	0.027	Belgium	0.024
Japan	0.029	Sweden	0.018	Japan	0.020
Switzerland	0.018	Rusia	0.018	Irlanda	0.019
Sweden	0.016	Japan	0.017	Rusia	0.018
Argentina	0.015	Switzerland	0.016	Sweden	0.017
Cumulative	0.828		0.771		0.703

Notes: The table displays the composition of international tourist inflows in 2000, 2010 and 2019. Only twelve countries with largest tourist inflows are displayed. The data source is the National Statistic Institute data on international arrivals to Spain.

Table A3.5.: Instrumenting tourism growth with beach characteristics: Impact on municipal income levels

	Panel A. IV, Dep. variable: Income level 2019				
	(1)	(2)	(3)	(4)	(5)
	length	+sun hours	+width	+temperature	+precipitations
Per capita tourism	-0.050**	-0.067***	-0.049**	-0.049**	-0.051***
growth (1965-2019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)
Mean dep. var.	9.49	9.49	9.49	9.49	9.49
	Panel B. First stage: Dep. variable: per capita tourism growth				
$\mathbb{1}$(Beach)	1.933***	3.480***	3.172***	3.173***	3.172***
Beach attractivness	1.028***	1.773***	0.150***	0.150***	0.147***
Kleibergen-Paap rk	27.172	27.596	27.636	27.752	26.987
LM F-stat.	[19.93]	[19.93]	[19.93]	[19.93]	[19.93]
Observations	1,196	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y

Notes: (1) Panel A reports the IV estimates of tourism growth using IV regression, where per capita tourism is instrumented by $\mathbb{1}$(Beach), a dummy equal to one if the municipality has beach and zero otherwise plus *Beach attractiveness* which measures beach quality and panel B reports the First stage regressions. *Beach attractiveness* in column 1 is the standardized km of beach length, column 2 adds standardized hours of sun, column 3 adds standardized km of beach width, column 4 adds standardized temperatures and column 5 adds standardized precipitations. The unit of observation is the municipality; (2) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A3.6.: Instrument validity for shift-share analysis: Overnight stays growth and municipal income

	Panel A. Dep. variable: Income level 2019		
	OLS	RF	IV
	(1)	(2)	(3)
Δ Realized overnight stays (1996-2019)	-0.014** (0.005)		
Δ Predicted overnight stays (1996-2019)		-0.029*** (0.001)	-0.029*** (0.009)
Mean dep. var.	9.49	9.49	9.49
	Panel B. First stage: Dep. variable: Predicted overnight stays		
Realized overnight stays			0.966***
Kleibergen-Paap rk LM F-stat.			16.86 [16.38]
Observations	96	96	96
All Controls	Y	Y	Y
province FE	Y	Y	Y

Notes: (1) Panel A reports the estimates of tourism growth measured by overnight stays as explained in section 3.4 on income level using the OLS, reduced-form and IV regressions results where the dependent variable is the log of income in 2019 and panel B the First stage where the independent variable is the realized overnight stays. The unit of observation is the municipality; (2) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level when adding province fixed effects. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A3.7.: Economic structure comparison: Proximity to beach municipalities. Treated vs. Control (10km vs. >20km cutoffs)

	Employment change by industry			
	Agriculture	Manufacturing	Construction	Hotels and foodservice
	(1)	(2)	(3)	(4)
Per capita tourism growth (1965-2019)	-0.109 (0.107)	-0.103 (0.087)	0.100 (0.102)	0.009 (0.053)
Observations	1452	1452	1452	1452
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: (1) The table shows the reduced form estimates of the effect of per capita tourism growth on employment shares in different sectors. The treated municipalities are those that are within a 10km radius of the beach municipalities, while the control municipalities are those that are further away, at least 20km. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table A3.8.: Comparative impact of tourism growth on income levels: a sub-period analysis (1965-1996 vs 1996-2019)

	Income level	
	1996	2019
	(1)	(2)
Per capita tourism growth (1965-1996)	-0.038**	
	(0.028)	
Per capita tourism growth (1996-2019)		-0.040***
		(0.010)
Kleibergen-Paap rk LM F-stat.	18.88	19.52
	[16.38]	[16.38]
Observations	1,196	1,196
All Controls	Y	Y
Province FE	Y	Y

Notes: (1) The table shows estimates of tourism on income level. Columns (1) and (2) use per capita tourism as the independent variable, which is instrumented by the existence of a beach in a municipality. Column 1 uses the dependent variable the income level in 1996 and Column 2 the income level in 2019. The unit of observation is the municipality. (2) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table A3.9.: Instrument validity: correlation of initial nationality shares with potential confounders

Nationality	Germans	Britons	French	Dutch	Portuguese
Population change	0.009	0.0318	0.0033	0.0065	-0.0002
Education level	0.000	-0.0022	-0.0010	-0.0014	-0.0000
Unemployment	-0.0006	0.0018	-0.0005	-0.0004	0.0000
Manufacturing industry employment	0.0105	0.0297	-0.0123	0.0984	-0.0162
Agriculture industry employment	-0.0119	-0.0550	-0.0059	-0.0098	0.0011

Notes: The table shows the correlations of the initial nationality shares across municipalities of the five tourist nationalities with the highest Rotemberg weights with several potential confounders. Changes refer to the period 1997-2019; Education, unemployment and shares by industry employment refer to the year 1990.

Table A3.10.: Other outcomes

	Unemployment level in 2019	Unemployment growth (1997-2019)	Population growth (1960-2019)	Population growth (1997-2019)
	(1)	(2)	(3)	(4)
Per capita tourism growth	0.002		0.806***	
	(0.002)		(0.354)	
Predicted overnight stays growth		0.461		0.038
		(0.322)		(0.063)
Mean dep. var.	0.07	0.51	0.97	0.22
Observations	1196	437	1196	437
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: (1) The table shows the the impact of tourism growth on the specified outcomes. The unit of observation is the municipality; (3) All columns control for pre-determined socioeconomic and geographic variables: population in 1950, illiteracy rate in 1930, income in 1965, agriculture production in 1960, tourism per capita in 1965, soil quality, altitude, ruggedness, surface, temperature, rainfall and hours of sun; (4) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

4. The Impact of Housing Costs on Labor Market Trajectories

4.1. Introduction

Over the past decade, many developed economies have experienced sustained house price increases. In the last two decades, housing prices have increased by roughly 46% in OECD countries (OECD, 2022). In France, prices have soared by more than 59% during this period (OECD, 2022) raising a real affordable housing concern, particularly, in large urban areas. Several studies have shown that as housing becomes unaffordable in the core of the urban area, workers are compelled to relocate to more affordable areas, often in the suburbs (Ong and Leishman, 2020; Bailey and Minton, 2018; Hulchanski, 2011). This creates a spatial mismatch between housing and employment opportunities that can have profound implications for labor market outcomes, given that an individual's location determines the amenities they experience, the network they create, and their opportunities to accumulate human capital.

The classical models by Rosen (1974) and Roback (1982) suggest that in an equilibrium state, wages should fully offset the variations in housing costs across urban areas. This theory is grounded in the observation that while the dense population in large urban areas drives up housing prices, firms in these urban areas can afford higher wages due to the benefits of agglomeration economies. Consequently, these models assume workers are perfectly mobile and enjoy equal real wages and indirect utility in different urban areas. When considering heterogeneous agents (high-skilled and low-skilled) with non-homothetic preferences and an income elasticity of demand for housing below one, the Rosen-Roback model implies that low-skilled workers, who face a higher housing cost burden, should be compensated more relative to high-skilled workers in expensive areas (Kim, Liu andamp; Yezer, 2009). Yet, this prediction doesn't fully align with the situation in France. Studies by Schmutz, Sidibé, and Vidal-Naquet (2021) indicate that low-skilled workers in France have limited mobility across urban areas. Moreover, recent evidence suggests these workers receive disproportionately lower compensation for increased housing costs compared to their high-skilled counterparts, a trend consistent with findings in other countries (Diamond and Moretti, 2021; Albouy et al., 2016).

In this study, we hypothesize that the relatively greater housing cost burden on low-skilled workers compared to high-skilled workers could significantly influence their labor market trajectories. We investigate the effects of housing prices on the disparities in labor market outcomes between high-skilled and low-skilled workers in French urban areas. Employing a between-city methodology, we use variations in housing costs across urban areas to examine their impact on the likelihood of obtaining a long-term contract and the probability of promotion, particularly focusing on the differences between these skill groups. [1]

Building on earlier observations that low-skilled workers are comparatively under-compensated for rising housing costs, this study hypothesizes that variations in housing prices across urban areas causally influence the labor market disparities between high-skilled and low-skilled workers. Elevated housing prices tend to exacerbate residential location discrepancies in relation to central employment hubs. For low-skilled workers, this often results in a greater disconnection from urban area centers. Such a disconnection can lead to reduced access to job opportunities, less intensive job search, as well as limited spatial search scopes due to extended commuting distances. These factors collectively create spatial dynamics that potentially have a detrimental impact on the labor market prospects of low-skilled workers.[2] Similarly, insights derived from standard search and matching models, such as those proposed by Mortensen and Pissarides (1999), suggest that a higher housing cost burden may compel low-skilled workers to hasten their job search. This is due to the increased utility of being employed versus unemployed under financial strain. Such urgency to find employment might result in low-skilled workers accepting sub-optimal job matches, leading to lower-quality job placements. In contrast, high-skilled workers, who face less pressure from housing costs, can afford to be more selective in their job search, potentially securing better matches.

[1]An urban area is a group of touching municipalities encompassing an urban center providing at least 10,000 jobs, and adjacent suburban regions where at least 40% of the employed residents work within the urban core or its neighboring municipalities. More details will be given in Section 4.

[2]The disconnection between jobs and job-seekers has been explored through the literature under the so-called spatial mismatch hypothesis. Higher distance to the employment center implies a strong downward trend in jobs among low-income workers due to higher commuting search costs, higher anticipated commuting costs, information decay in job opportunities, and reduced spatial search horizon (Zenou, 2009; Patacchini and Zenou, 2005; Gobillon et al., 2013; Manning and Petrongolo, 2017; Schmutz and Sidibé, 2019; Schmutz et al., 2021; Flemming, 2020; Balgova, 2022)

While the existing literature primarily explores the relationship between housing prices and aggregate labor market outcomes, there is limited research on how housing costs impact individual career paths and, more importantly, how these effects differ among workers of varying skill levels. To bridge this gap, in our research, we attempt to answer the following questions: How do housing cost differences across urban areas affect labor market disparities between high-skilled and low-skilled workers? And how might these differences exacerbate existing inequalities in long-term contract and promotion opportunities between skill levels?

To track individual labor trajectories, we use an extensive employer-employee matched dataset that provides detailed records of individuals' employment histories. Focusing on the 2010-2018 period, our data includes 3,724,905 observations of 463,785 individuals across the 151 largest French urban areas. Through a between-city analysis, we investigate disparities between high-skilled and low-skilled workers in various urban areas, while considering potential within-city factors contributing to these disparities.

To accurately assess the impact of housing costs on labor market disparities by skill levels across urban areas, two critical challenges must be addressed. First, the issue of spatial sorting arises, as individuals, particularly those with higher skills, tend to non-randomly select larger urban areas. This spatial sorting can introduce potential selection bias into local labor market indices. To mitigate this bias, we employ a two-step approach analogous to Combes et al. (2008). In the first stage, we estimate local urban area labor market indices by skill level, accounting for both observable and non-observable individual characteristics. This reduces the influence of spatial sorting across urban areas. Specifically, we regress the outcome variables, the probability of having a long-term contract and promotion probabilities, against worker characteristics, worker fixed effects, and urban area-skill fixed effects.[3] This allows us to obtain long-term contract and promotion probability indices by skill level for each urban area. In the second stage, we regress the estimated urban area-skill fixed effects, particularly the differences between high-skilled and low-skilled workers, on housing costs and urban area controls.

The second challenge lies in the endogeneity between housing and labor markets. Housing prices often rise in urban areas with dynamic labor markets, making it challenging to isolate the pure effect of housing costs on labor outcomes. To address this potential endogeneity, we employ inverse housing supply elasticities derived by

[3]Given the particularity of the definition of promotion a preliminary step is added in the empirical strategy to accurately retrieve individual fixed-effects from a Mincerian standard wage equation. More details will be given in section 5.

Chapelle et al. (2023) as an instrumental variable for housing costs. These elasticities provide a measure of housing supply constraints, that are known to increase housing costs, and that therefore help in disentangling the separate influences of housing costs, and labor market dynamism on labor market outcomes.

Our results show that rising housing costs increase disparities in securing long-term employment contracts between high-skilled and low-skilled workers. Specifically, a 10% increase in housing costs correlates with a 0.14 percentage point increase in this disparity. This indicates that higher housing costs disproportionately affect the job stability of low-skilled workers compared to their high-skilled counterparts, aligning with mechanisms outlined in urban and labor economics models.

Further, our sector-based heterogeneity analysis reveals that low-skilled workers consistently face greater disadvantages than high-skilled workers in the service, industry, and commerce sectors. In addition, low-skilled women are more affected than high-skilled women by the impact of housing cost fluctuations on long-term employment contracts. Age is also an influencing factor; low-skilled workers over 34 years old experience increased job insecurity relative to their high-skilled counterparts as housing costs rise. However, our study does not find a significant impact of housing prices on promotion probabilities, suggesting that other factors may be more influential in determining career advancement opportunities across different skill levels.

The structure of this paper is as follows: section 2 situates our study within the existing body of literature. In section 3, we detail the data used in our analysis and offer some descriptive statistics. Section 4 presents stylized facts highlighting the pronounced impact of housing costs on low-skilled workers and outlines the theoretical framework of our study. Section 5 describes our empirical design and the methodologies employed to address endogeneity concerns. Our empirical results are discussed in section 6. Finally, we draw conclusions in section 7.

4.2. Literature review

This paper contributes to several strands of literature. First, this work is related to a large literature that studies the interactions between the housing and the labor market. Previous studies have explored the relationships between local housing markets and labor markets. For instance, Bover et al. (1989) demonstrated a correlation between regional differences in house-price-to-earnings ratios and unemployment rates in the United Kingdom. Johnes and Hyclak (1994) used data from the southeast region of the UK to analyze the impact of housing prices on short-term regional employment adjustments using a system of equations. In a separate study, Johnes and Hyclak (1999) examined four U.S. metropolitan areas and employed an error correction model to illustrate the significant influence of housing prices on the size of the local labor force.

More recent research has made noteworthy contributions to this field. Saks (2008) investigated how housing supply regulations affect housing and labor market dynamics in various U.S. metropolitan areas. She argued that government regulations, such as land-use policies, can reduce the flexibility of housing supply, leading to changes in housing price distribution and patterns of labor migration. Consequently, areas with constrained housing supply experience lower employment growth. Saks provided empirical evidence supporting this hypothesis.

Similarly, Zabel (2012) employed a comprehensive model that considered in-migration, out-migration. His findings indicated that positive demand shocks result in increased migration and labor turnover in high-housing-cost areas. However, the resulting employment and wage increases are comparable between high and low-housing-cost areas. Fort et al. (2013) observed more substantial declines in net employment growth for young and small businesses in states where housing prices experienced significant declines. Chakrabarti and Zhang (2014) investigated the relationship between housing affordability and employment growth, finding that unaffordable housing hampers local employment expansion.

Several papers find that as housing becomes unaffordable in central urban areas, workers are forced to relocate to more affordable regions, often on the fringes of cities (Ong and Leishman, 2020; Bailey and Minton, 2018; Hulchanski, 2011). In the recent period, several studies suggest that low-skilled are more affected by higher housing costs than high-skilled. Recent increases in housing costs make it more difficult for low-skilled workers to afford to live in high-employment opportunity areas (Diamond, 2016; Broxterman and Yezer (2015); Moretti, 2013; Gyourko et al. 2013; Hulchanski 2011). As a result, there is a growing skill divergence between high-cost and low-cost cities (Berry and Glaeser, 2005). This spatial mismatch between affordable housing and job locations leads to increased worker commuting burdens and local labor shortages, an issue that has been highlighted in several studies (Gobillon and Selod 2011; Brueckner and Zenou, 2003; Brueckner and Martin, 1997; Coulson et al., 2001; Patacchini and Zenou, 2005; Oi, 1976, among others)

This paper makes two contributions to the existing literature. First, we provide a skill-specific analysis, focusing on high-skilled and low-skilled workers. While previous studies have explored the impact of housing costs on labor markets in general, this study narrows down to understanding how these costs affect individuals with varying skill levels. This granularity offers insights into how housing disparities can impact existing inequalities. Second, we examine the impact of housing prices on different labor market outcomes, such as job stability and promotion. Prior studies have primarily focused on the relationship between housing prices and aggregate labor market outcomes, such as unemployment and labor mobility. To our knowledge, the question of whether housing costs increase the disparities in job stability and promotion between high-skilled and low-skilled workers has not been explicitly addressed. That said, Charruau and Epaulhard (2022) study the effect of agglomeration economies on promotion but they do not differentiate across skill levels nor consider the the costs of agglomeration.

Besides, this study is connected to the broader work that studies the benefits of bigger cities. Cities have long been seen as engines of economic growth, and there is a growing body of research studying the relationship between city size and economic opportunity. One of the key findings from the literature is that workers in larger cities are more likely to be employed in high-productivity industries and to earn higher wages (Roca and Puga, 2017). This is likely due to the fact that larger cities have a wider variety of industries and businesses, which creates more opportunities for workers to specialize in their skills and to find jobs that match their abilities. Additionally, larger cities tend to have higher levels of agglomeration economies, which can lead to higher productivity and wages for workers (Koster and Ozgen, 2021).

However, the benefits of city size are not evenly distributed. Workers at the lower end of the skill distribution in big cities often get paid less than necessary to offset their greater housing costs. Diamond and Moretti (2021) highlight the challenges faced by low-skilled workers in expensive cities, finding a significant negative relationship between consumption and cost of living for this group. This implies that high-cost cities offer a lower standard of living compared to more affordable cities for low-skilled workers, while their presence, due to extreme-skill complementarity, is essential to the functioning of large cities (Eeckhout et al. 2014). Additionally, less-educated workers in urban areas may face barriers to employment due to their lower geographical mobility (Hornbeck and Moretti, 2022). This study makes a significant contribution to the literature by providing a comprehensive and nuanced analysis of how city size shapes labor market disparities between high-skilled and low-skilled workers across a range of urban areas in France.

4.3. Data

In this section, we detail the classification of workers in terms of skills, the unit of analysis, and the data used in the analyses.

4.3.1. Definitions

Defintion of high-skilled and low-skilled workers. In this paper, we define workers' skill levels based on the socio-professional categories defined by the French National Institute of Statistics and Economic Studies (INSEE). The INSEE defines six broad socio-professional categories that reflect occupations and positions on the labor market: (i) farmers, (ii) independent workers (craftsmen, shopkeepers, and business owners), (iii) senior executives and higher intellectual occupations, (iv) intermediate occupations, (v) low-level white-collar workers, and (vi) blue-collar workers. We define high-skilled workers as senior executives and individuals in higher intellectual occupations, while low-skilled workers encompass low-level white-collar positions and blue-collar jobs.[4] Table 4.3.1 below outlines the decile distribution of hourly net wages in general and by these two skill levels on a representative sample of French workers. We can see that net hourly wages for low-skilled workers tend to fall within the lowest deciles, while those for high-skilled workers are predominantly found in the 9th decile. This highlights that comparisons between low-skilled and high-skilled individuals can equate to comparisons between low-income and high-income workers.

Table 4.3.1.: Distribution of net hourly wage in general and by skill level

	D1	D2	D3	D4	D5	D6	D7	D8	D9	Mean	SD
	10%	20%	30%	40%	50%	60%	70%	80%	90%		
Hourly wage	8.7	9.6	10.5	11.5	12.7	14.2	16.4	19.8	25.6	15.7	10.9
High-skilled	14.6	17.1	19.2	21.2	23.1	25.1	27.6	31.8	40.6	26.6	16.9
Low-skilled	8.3	8.9	9.4	9.9	10.6	11.3	12.2	13.4	15.6	11.5	4.2

Source: Panel of active individuals from the EDP (2010-2018). The sample includes 3,724,905 observations, 151 urban areas, and 463,785 individuals.

Definition of urban areas. This analysis employs French urban areas, defined in 1999, as the basis for characterizing cities and representing local labor markets and housing markets.[5] These urban areas, as delineated by INSEE, comprise a cluster of contiguous municipalities (communes) centered around an urban core with significant employment opportunities (comprising a minimum of 10,000 jobs in larger urban areas) and adjacent suburban regions where at least 40% of the employed residents work within the urban core or its neighboring municipalities. As it will be outlined in section 4.3, we will focus in our analysis on the main French urban areas.

[4] Further details will be given on their distribution in our estimation sample in section 4.3.

[5] We use the 1999 urban areas denomination mainly due to the inverse supply elasticities measured at this level in Chapelle et al. (2023) that we used in our empirical approach. There does not also seem to be a significant value-added in using the newest denominations, especially when we stick to the main French urban areas, the 1999 delimitation still reflecting well today's local labor and housing markets.

4.3.2. Employment data

To follow the labor market trajectories of French workers, this paper relies on data sourced from the *Panel d'actifs of the Echantillon Démographique Permanent,* which is generated and maintained by INSEE, covering the period from 2010 to 2018.

The *Echantillon Démographique Permanent* is a longitudinal panel that was initiated in 1968. Its sample is composed of individuals born on specific reference days within a given year. Initially, this sample included individuals born between the 1st and the 4th of October (spanning 4 days) every even year. Since 2002, and specifically for the *panel d'actifs,* it also encompasses individuals born between the 2nd and 5th of January, as well as the 1st and 4th of April (totaling 16 days) each year. As a result, this panel represents approximately 4% of the French active population.

The *panel d'actifs* within the EDP gathers information pertaining to the professional salaried activities of these individuals. This information is derived from the *Panel Tous Salariés* (DADS-DSN), which is a comprehensive employer-employee matched dataset that tracks individuals across both the public and private sectors on an annual basis for as long as they remain employed. The data available in the *panel d'actifs* of the EDP is aggregated at an annual level. This means that for each year, the dataset retains information on the individual's overall employment activity. However, it only preserves the characteristics of the individual's main job.[6]

The Panel d'actifs provides information on the nature and localization of the main job each year. This includes details such as the number of hours worked, the annual net wage, the specific workplace at the municipality (commune) level, the type of employment contract (whether it is long-term or short-term), the sector of activity or industry, the occupation level (socio-professional category), and the terms of employment (whether it is full-time or part-time). Additionally, it captures several worker's characteristics, including age, gender, and place of residence at the municipality level.

In our analysis, we focus on several key outcome measures. For contract type, we categorize individuals as having long-term contracts if they are engaged in open-ended contracts (*contrat à durée indéterminée*). Regarding wage calculations, we determine the net hourly wage by dividing an individual's annual net earnings

[6]The Panel d'actifs of the EDP defines the main job as the job with the longest pay period in a particular year.

(adjusted to constant euros) by their total annual working hours. As for promotions, we define it as a binary variable (0/1), indicating whether an individual experienced at least a 35% increase in hourly wage during the period from 2010 to 2018.[7] This variable captures advancements in individuals' careers within the labor market.

4.3.3. Housing prices data

To measure housing prices at the urban area level, we leverage the comprehensive housing transaction dataset known as *Demande de Valeurs Foncières (DV3F)*. This dataset is compiled by the *Centre d'études et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement (CEREMA)* and covers the period from 2010 to 2021.

DV3F is an exhaustive dataset that records all second-hand housing transactions and land transactions taking place in France, as reported by land registries.[8] Of particular significance to our study are the detailed information on the second-hand dwelling sales. This information includes not only the sale price but also crucial transaction details, such as the transaction date and comprehensive characteristics of the dwelling. These characteristics encompass various aspects, including, among others, the dwelling's location, total surface area (measured in square meters), construction period, number of rooms, and the composition of the dwelling (including features like bathrooms, cellars, balconies, and the presence of parking facilities).

As explained in details in the appendix in Section B, we compute housing price indexes at the urban area level by employing a hedonic regression model on housing transactions from 2010 to 2018 across 151 French urban areas. To capture the nuances of housing prices, we control for various factors such as dwelling type, construction period, surface area, number of rooms, distance from urban centers, and quarter of the year. The resulting urban area housing price indexes effectively reflect fluctuations in housing prices across urban areas over time. Notably, the regression results in A4.16 indicate that houses consistently command higher prices, while factors like modern construction periods, larger surface areas, more rooms, proximity to urban centers, and transactions occurring in the third and fourth quarters of the year were positively associated with housing prices.

[7]We use 35% as a threshold as it corresponds to the 8th decile in the distribution of increases in hourly wage in our sample of analysis. It is also one of the alternative definitions of promotion used in Charruau and Epaulhard (2022) on a similar panel of French workers.

[8]Only 2% of the transactions observed in DV3F correspond to new dwellings.

Finally, to adequately account for potential variations in labor market dynamics across urban areas, our analysis incorporates several control variables. These include the total number of jobs, the unemployment rate, and the distribution of employment across different sectors: agriculture, construction, industry, commerce, transport, service sectors, and public administration, in each urban area present in the analysis. We sourced this data from various censuses available through the INSEE online database.

4.3.4. Descriptive statistics

Our analysis spans the years 2010-2018. To ensure data quality, we refined the sample by removing individuals with missing location data, outliers in hourly wages, those outside the 15-60 age range at first observation, retirees in certain economic activities, and individuals with extremely low annual hours or hourly wages. Our focus is on three skill levels within the private sector: high-skilled (senior executives), low-skilled (blue-collar and white-collar workers), and intermediaries (intermediate occupations). The final sample includes only urban areas with available inverse supply elasticities, leading to a dataset of 3,724,905 observations from 463,785 individuals across the 151 largest French urban areas. This represents 72% of the urban population in 2010.[9]

Table A4.1 presents the descriptive statistics for individuals in the estimation sample. A majority of the workers present in the sample are aged between 30 and 49 years old, while our comprises 57.7% of men vs. 42.3% of females, thus reflecting the lower extensive labor supply of women. In terms of occupations and skill level, 60% of the individuals in the sample are low-skilled workers (27.1% blue-collar workers and 32.9% low-level white-collar workers), high-skilled represent 19.7% of the sample while intermediaries correspond to 20.3% of the individual observed.[10] In terms of urban area types, we can see that as we restrict our analysis to the 151 largest French urban areas, we have 60% of the observed individuals that reside in urban areas with more than 500,000 inhabitants, with 33.8% of the individuals in the Paris urban area which we know, concentrates an important part of jobs in France. Table A4.2 outlines the distribution of skills by urban area size in the analysis. Unsurprisingly, we can see that the larger the urban area, the higher the share of high-skilled in the active population. If we now move to the characteristics of the main job held, we can see that a significant portion of the sample (77.3%) holds full-time employment while the majority of the workers (80.7%) have open-ended contracts.

[9] For more details on sample restrictions, see C.

[10] Figures that are close to the INSEE data regarding the distribution of jobs by socio-professional categories in the French labor market. https://www.insee.fr/fr/statistiques/6453692?sommaire=6453776

Regarding hourly net wage distribution, Table A4.3 shows that the mean hourly wage stands at 14.4 euros with a first decile at 8.3 and a 9th decile at 23.4 euros. Moreover, the main sector of employment for individuals present in the sample is the service sector (62.3%), underlining the tertiarisation of the French economy that started long ago. We can therefore say that our sample is representative of the French labor market and the demographic composition of individuals in the 151 largest French urban areas.

Table A4.4 provides a summary of statistics related to housing prices and the labor market characteristics of the 151 urban areas present in our analysis over the 9 years period. The unemployment rate in these areas stands on average at 14%. Similarly to the descriptive statistics for individuals, most urban areas exhibit a higher concentration of jobs in the service and commerce sectors.

4.4. Stylized facts and theoretical framework

4.4.1. Stylized facts

This paper is motivated by several key observations indicating that low-skilled workers are disproportionately impacted by rising housing costs, particularly in recent years.

Fact 1: The recent rise in housing prices. Figure A4.1 provides a backdrop for our analysis, depicting the housing market trends in France from 2010 to 2021. This figure shows yearly housing price indexes, highlighting a notable increase in housing costs during this period. The surge is especially pronounced in the Paris urban area and other major urban centers such as Lyon, Marseille-Aix-en-Provence, Bordeaux, Montpellier, Toulouse, Toulon, and Nantes.[11]

Fact 2: Divergence in wage growth between skill levels. Table 4.3.1 elucidates the initial wage disparities between different skill levels. It shows that the average net hourly wage for low-skilled workers (€11.5) is significantly lower than that for high-skilled workers (€26.6). Additionally, Figure A4.2 tracks the progression of average hourly wages for these two groups from 2010 to 2018. Using a base index of 100 in 2010, the upper figure reveals that high-skilled workers experienced a 15%

[11]These trends are derived from the DV3F housing transactions dataset, which includes 5,404,885 housing transactions across the 151 urban areas (1999 denomination) under study. To gauge the overall housing market trends, we conducted a hedonic price regression (as in equation 4.6) with year-fixed effects to generate general annual house price indexes. We also calculated the general log(mean price) and log(mean price per m2) annually, which yielded similar findings. A parallel analysis was performed for different types of urban areas, using three subsets of data: (i) the Paris urban area (1,523,403 transactions); (ii) other large urban areas (1,042,641 transactions); and (iii) the remaining 143 urban areas (2,838,831 transactions).

wage increase over this period, compared to a 10% increase for low-skilled workers. Consequently, as depicted in the lower figure, there has been a noticeable rise in the skill wage ratio since 2016.[12] These observations, coupled with the increase in housing prices (Fact 1), paint a picture of a widening economic gap that may further exacerbate the challenges faced by low-skilled workers.

Fact 3: Disparities in housing costs and wage compensation across skill levels. In France, low-income workers typically spend a larger portion of their income on housing compared to higher-income groups. For instance, the lowest income quintile allocates about 30% of their budget to housing, compared to only 17% for the highest quintile, a disparity that has been growing over time (OECD Affordable Housing Database, 2020). According to the standard Rosen-Roback models that account for heterogeneous agents, this higher housing cost burden implies that low-skilled workers should receive *relatively* higher wage compensation than high-skilled workers in areas with higher living costs. This would mean that the skill wage ratio (high-skilled to low-skilled) should decrease as housing costs rise, indicating a negative coefficient in the relationship between skill wage ratios and housing prices.

Figure A4.3 is derived from yearly regression analyses on a panel of 151 French urban areas, covering the period from 2010 to 2018. In these regressions, for each year (denoted as t), we model the equation $Y_j = \alpha + \beta Housing\ costs_j + \varepsilon_j$, where Y_j is the logarithm of the average hourly wage (for high-skilled workers, low-skilled workers, or their ratio) in urban area j. The term $Housing\ costs_j$ represents the housing price index in urban area j. The figure visualizes the β coefficients from these regressions, highlighting the relationship between the wage outcomes Y and the housing price indexes in various urban areas. The objective of these regressions is to examine the temporal evolution of hourly wages for high-skilled and low-skilled workers in relation to fluctuating housing prices across different urban areas.[13]

Analysis of Figure A4.3's upper panel reveals that both high-skilled and low-skilled workers tend to receive some wage compensation in urban areas with higher housing costs, as indicated by the positive coefficients. However, a distinction is observed in the compensation patterns across skill levels. High-skilled workers show a consistent upward trend in their hourly wages in response to rising housing costs over time. In contrast, the wages of low-skilled workers remain relatively stable in

[12] These averages were calculated using data from the Panel d'actif of the EDP over 2010-2018, representing a sample of 3,724,905 observations across the 151 largest French urban areas. The hourly wages, adjusted for inflation, are expressed in constant euros. We constructed hourly wage indices with 2010 as the base year.

[13] This analysis utilizes data from 151 French urban areas, as classified in 1999. Housing price indexes are calculated as described in equation 4.6 using DV3F data. Meanwhile, the average hourly wages per skill level in each urban area from 2010 to 2018 are sourced from the Panel d'actifs of the EDP, with adjustments for inflation to maintain constant value over time.

relation to housing prices during the same period. This disparity becomes markedly significant from 2016 onwards, leading to a positive and statistically significant (at the 1% risk level) coefficient in the skill wage ratio versus housing prices relationship, as shown in the lower panel of Figure A4.3. These findings suggest that while high-skilled workers receive financial compensation commensurate with increasing housing costs, the same is not true for low-skilled workers.[14] Despite bearing a heavier housing cost burden, low-skilled workers appear to be less compensated than high-skilled workers for rising housing costs in recent years.

Fact 4: Reduced mobility among low-skilled workers across urban areas. The final stylized fact concerns the well-documented lower mobility rates of low-skilled workers across urban areas in France, compared to their high-skilled counterparts. This trend, consistent since 2013 as reported by Schmutz et al. (2021), is evident in Census data from INSEE, which shows that the mobility rate for low-skilled workers across urban areas is around 3%. In contrast, the highest mobility rates are observed among young high-skilled individuals. This limited mobility among low-skilled workers can induce various frictions, including challenges related to relocating for employment or housing purposes.

4.4.2. Theoretical framework

Building from these previous stylized facts that tend to show that the low-skilled suffer more from high housing costs over the recent period, we hypothesize in this paper that urban areas with higher housing costs exhibit larger disparities in labor market outcomes between high-skilled and low-skilled workers than urban areas with lower housing costs.

To make the link between high housing costs and labor market disparities between skills, let's consider the following simple initial hypothetical setting where we have one urban area A and two types of workers: high-skilled and low-skilled. Due to different factors, urban area A exhibits high housing costs, and we assume that these high housing costs increase the gap in labor market outcomes between high-skilled and low-skilled in this urban area. It is well-established that housing costs influence location choices within urban areas. According to classical urban economics models (Fujita, 1989), when the income elasticity of marginal transport costs is higher than the income elasticity of housing consumption, wealthier individuals tend to reside closer to the central business district (CBD), which concentrates on job opportunities, while lower-income individuals settle in the outskirts. In urban area A, the high housing prices therefore imply a larger distance to the CBD for the low-skilled than the high-skilled. This spatial segregation of low-skilled workers away from the

[14] For detailed regression results related to Figure A4.3, refer to Table A4.5 in the Appendix.

urban area center gives rise to spatial mechanisms that may hinder their labor market outcomes. Consistent with urban labor economics models, low-skilled workers in urban area A, compared to their high-skilled counterparts, face reduced access to information about job opportunities (information decay), conduct less intensive job searches, and limit their spatial search horizon and the set of acceptable employment matches (Zenou, 2009; Patacchini and Zenou, 2005; Gobillon et al., 2013; Manning and Petrongolo, 2017; Schmutz and Sidibé, 2019; Schmutz et al., 2021; Flemming, 2020; Balgova, 2022). At the aggregate level, and following standard search and matching labor economics models, the higher housing cost burden for low-skilled workers in urban area A pushes them, due to the increased utility of employment relative to unemployment, to find more rapidly a job, potentially leading to sub-optimal firm-worker matches, and therefore, jobs of lower quality (Mortensen and Pissarides, 1999).

If we now add an urban area B with lower housing costs and two skill levels to this initial setting and go to the between-city level, the (i) within-city spatial mechanisms and (ii) aggregated labor market mechanisms at play in urban area A implies that labor market disparities across skill levels will be more pronounced in urban area A. In fine, our empirical approach corresponds to a *between-city* analysis that amounts to compare labor market disparities between high-skilled and low-skilled in urban area A to labor market disparities in urban area B, drawing on the two previously mentioned types of mechanisms.

It's important to recognize that while our hypothesis, which posits specific within-city mechanisms based on established urban labor models, guides our analysis across different cities, we must acknowledge our limitations in directly testing or validating these mechanisms within our empirical framework.

4.5. Empirical strategy

Our study focuses on analyzing the disparities in labor market outcomes between high-skilled and low-skilled workers in French urban areas. Following the methodology of Combes et al. (2008), we adopt a two-step approach. We start by calculating indices for local labor market outcomes, specifically focusing on the likelihood of securing a long-term contract and promotion opportunities, categorized by skill level. Then, we examine the influence of housing prices on these labor market skill indices across different urban areas.

A key aspect of our analysis involves addressing two significant challenges in identification. As highlighted by Combes et al. (2008) and Roca and Puga (2017), there is a non-random distribution of individuals across urban areas, with more skilled individuals often gravitating towards larger urban areas. Individuals sort across urban areas based on observable and unobservable characteristics, and such spatial sorting could introduce a bias in our labor market indices. To mitigate this, we control for both observable and unobservable factors influencing individuals' location choices. Another challenge is the potential endogeneity between housing and labor markets. Urban areas with dynamic labor markets often have rising housing prices, creating a complex interplay between the two sectors. Our task is to separate the direct impact of housing prices on labor market outcomes from the effects of a dynamic labor market.

4.5.1. Labor market disparities and spatial sorting across urban areas

We first turn to elaborate on the two-step approach for defining local labor market indices and handling spatial sorting.

The first step of our methodology is focused on creating skill-specific labor market indices for urban areas. These indices are carefully adjusted to account for both observable and unobservable individual characteristics, thus mitigating the impact of spatial sorting across these areas. This adjustment allows us to more accurately discern labor market disparities that are not influenced by individual traits or their choice of location.

To exemplify our approach, consider the likelihood of obtaining a long-term contract as a measure of labor market outcomes. In this first step, we conduct a regression analysis using a panel dataset of individuals. The equation used for this analysis is as follows:[15]

$$LT_{isjt} = \alpha + \gamma X_{isjt} + \tau_i + UA_{jts} + \varepsilon_{isjt} \tag{4.1}$$

We define LT_{isjt} as a binary variable indicating whether individual i with skill level s in urban area j at time t has a permanent contract. The variable X_{isjt} includes time-varying individual attributes like age, age squared, and employment sector, all of which can influence this probability. The individual fixed-effects, τ_i, serve two critical functions. Firstly, they account for both constant observed and unobserved individual characteristics, such as ambition and ability, thereby controlling for

[15] It's important to note that, in this first step, intermediaries remain in the sample to capture the overall effect of the urban area.

endogeneity in the outcome linked to these traits within the urban area. Secondly, they help mitigate the bias from spatial sorting by accounting for the fact that certain individual traits, correlated with securing long-term contracts, might also influence their choice of urban area.

We also introduce UA_{jts}, representing urban area-year-skill fixed effects. These are crucial in our analysis, as they provide localized, time-specific indices on the likelihood of securing long-term contracts, categorized by skill level. These effects, capturing unique aspects of each urban area at each time point for different skill groups, have been adjusted for both observed and unobserved individual characteristics. They will be the dependent variables in our second-step analysis.

In the second step, we shift our focus to a more aggregated level of analysis. Here, we use the urban area-year-skill fixed effects, determined in the first step, as our dependent variables. Our primary interest lies in examining the disparities between high-skilled and low-skilled workers. We conduct regression analyses to understand how differences in these fixed effects, based on skill level, relate to housing costs and other control variables specific to urban areas. These controls are selected to reflect the broader labor market dynamics within each urban area.

We estimate the following equation on a panel dataset comprising 151 urban areas, observed over a period of nine years from 2010 to 2018:

$$\hat{\mu}UA_{jt_{high}} - \hat{\mu}UA_{jt_{low}} = \rho + \beta Housingcosts_{jt} + \sigma Z_{jt} + \eta_t + \psi_{jt} \tag{4.2}$$

In this second step, $\hat{\mu}UA_{jt_{high}}$ represents the estimated local labor market index for high-skilled workers in urban area j at time t, indicating their probability of securing a long-term contract. The term $\hat{\mu}UA_{jt_{high}} - \hat{\mu}UA_{jt_{low}}$ measures the difference in these long-term contract indices between high-skilled and low-skilled workers within the same urban area and time frame. This difference quantifies the disparity in long-term contract opportunities between the two skill groups in a specific urban area.

$Housing\ costs_{jt}$ refers to the housing expenses in urban area j at time t, where β is the coefficient of interest. These costs are represented by local housing price indexes, computed as per Equation 4.6 explained in the appendix. To tackle potential endogeneity issues related to housing costs, we will use the inverse of housing supply elasticities unique to each urban area as instruments, a method detailed in the following subsection.

Finally, Z_{jt} includes various control variables at the urban area level, aimed at reflecting local labor market conditions. These variables encompass the number of jobs, unemployment rates, and the sectoral composition of employment, including sectors such as agriculture, construction, industry, commerce, transport, services, and public administration.

In the second step of our analysis, we handle a panel of urban areas with varying population sizes across different years. These variations in population can affect the precision of our dependent variable estimates, as larger urban areas typically yield more accurate estimations. To counter this, we implement a weighting scheme where weights are assigned to each urban area-year observation based on the population size of urban area j at time t. This approach helps to balance the differences in precision across observations, giving greater weight to those with lower error variances and more substantial data. This method aligns with the strategy suggested by Moretti (2004).

Additionally, considering that the dependent variable in Equation 4.2 is derived from an earlier step and might not represent the 'true' value, we refine our standard errors using bootstrap methods. This adjustment aims to mitigate potential uncertainties, thereby ensuring more robust and reliable estimates in our analysis.

4.5.2. Addressing spatial sorting for promotion outcome

We define a promotion as a 35% increase in hourly wage over an 8-year period.[16] In this context, promotion is considered a singular 'event' in an individual's labor market trajectory. Given that promotion is a binary outcome compute over different yearly observations, equation 4.1 cannot be analyzed using panel data but rather through a cross-sectional setting, where only one observation per individual is retained.

Acknowledging the significance of τ_i in our two-step approach for mitigating the potential spatial sorting, we incorporate an additional step for the promotion outcome following the method of Charruau and Epaulhard (2022). This involves extracting the individual fixed effects from a standard micerian wage equation. To this end, we estimate the following equation using a panel of individual data:

$$w_{isjt} = \upsilon + \chi X_{isj} + \tau_i + UA_{js} + \kappa_t + \zeta_{isjt} \tag{4.3}$$

where w_{isjt} represents the log annual hourly wage of individual i with skill level s in urban area j at time t. The variable X_{isjt} includes time-varying individual characteristics that influence wages, such as age, age squared, and sector of employment. The term UA_{js} denotes urban area-by-skill fixed effects, capturing unique wage characteristics in urban area j for skill group s. This term, for example, reflects the tendency for high-skilled workers to earn higher wages in larger urban areas. κ_t accounts for time trends, and ζ_{isjt} is the error term.

[16]This threshold aligns with the 8th decile in the distribution of changes in hourly wages.

The crucial parameter in equation 4.3 is τ_i, the individual fixed effects. It encompasses the impact of unobserved factors (e.g., talent, ambition) and observed but constant characteristics (e.g., education level, gender) on wages. To mitigate potential spatial sorting bias, we use $\hat{\tau}_i$ as a proxy for individual fixed effects in the promotion probability analysis, as part of the first step in our two-step approach. Following this, we implement the same two-step approach as detailed in our analysis of long-term contract probabilities.

Given that our setting is cross-sectional, we manage mobility across urban areas by censoring the data. Specifically, we retain only those observations where the individual i is predominantly located in a particular urban area, to accurately compute the probability of promotion.

4.5.3. Instrumental variable: inverse housing supply elasticities

Given the possibility that housing prices may be influenced by factors affecting local labor markets, a direct OLS estimation in equation 4.2 might introduce bias. This bias could cause the coefficient β to inaccurately reflect labor market conditions. To address this issue, we adopt an instrumental variable (IV) approach, using inverse housing supply elasticities at the urban area level in France, as calculated by Chapelle et al. (2023). This method is supported by a substantial body of literature that employs housing supply elasticities and variations in housing supply as instrumental variables (e.g., Campbell and Cocco, 2007; Mian et al., 2013; Aladangady, 2017; Kaplan et al., 2020; Guren et al., 2020).

Housing supply elasticity measures the responsiveness of housing stock to demand changes. Inelastic supply means that demand shifts predominantly affect prices rather than the quantity of housing. In their study, Chapelle et al. (2023) determine inverse housing supply elasticities for French urban areas, considering factors like geographical constraints (such as flood-prone areas, steep slopes, and topological limitations), land-use regulations, and demographics. They specifically examine the role of geographical and regulatory constraints in amplifying housing prices following demand shocks.[17]

[17] Geographical constraints include the proportion of land unsuitable for construction near the central business district (CBD), while regulatory constraints are indicated by building permit refusal rates. A higher inverse housing supply elasticity signifies more substantial constraints and limited housing supply.

To counter the potential endogeneity of local housing prices in our model, we use these constraints as exogenous instruments. They are assumed to be uncorrelated with labor market dynamics and have a direct impact on housing prices. This strategy enables us to isolate the influence of housing prices from the effects of dynamic labor markets. For this purpose, we employ a two-stage least squares (2SLS) regression method, beginning with the following first stage:

$$Housingcosts_{jt} = \zeta + \lambda invsupelast_j \times t + \omega Z_{jt} + \phi_t + \iota_{jt} \tag{4.4}$$

with $invsupelast_j \times t$ the inverse supply elasticity in urban area j interacted with a time t effect that accounts for the effect of general housing market trends.
In the second stage, we rewrite equation 4.2 with the instrumented housing costs:

$$\hat{\mu}UA_{jt_{high}} - \hat{\mu}UA_{jt_{low}} = \rho + \beta \widehat{Housingcosts}_{jt} + \sigma Z_{jt} + \eta_t + \psi_{jt} \tag{4.5}$$

4.6. Results

In this section we present the results. We first underline the relationship between housing prices and the inverse housing supply elasticities before turning to the long-term contract and promotion model results.

4.6.1. Instrumental variable

Table A4.6 presents the first-stage instrumental variable (IV) regression. We employ the inverse housing supply elasticity interacted with year dummies as an instrument and regress it on the housing price index at the urban area level each year. The results in column one demonstrate, as expected, a robust positive relationship between the inverse housing supply elasticity and house prices. The F-statistic strongly rejects the hypothesis of a weak instrument, indicating the instrument's validity. This finding aligns with prior studies that have established the power of this instrument in predicting house prices (Saiz, 2010; Mian and Sufi, 2011, among others).

4.6.2. Analysis of the impact of housing costs on long-term contracts

Main results. In the first step of our two-step approach, we estimate urban-area-skill-year fixed-effects, which serve as indices for the yearly likelihood of holding a long-term contract across different skill levels and urban areas. These fixed-effects are calculated in relation to a reference individual as defined in equation 4.1: one with average age, average unobserved ability, and employed in the industry sector. By using these fixed-effects, we can compare the probability of being in an open-ended contract for this reference individual across varying skill levels, urban areas, and years. We set the baseline for these fixed-effects at the level of low-skilled individuals in Toulon in 2010, assigning them a fixed-effect coefficient of 0. For example, a positive fixed-effect coefficient for high-skilled workers in Paris in 2010 would indicate a higher likelihood of these individuals having a long-term contract compared to our reference group of low-skilled workers in Toulon in 2010.[18]

The distribution of long-term contract indexes (the estimated urban-area-skill-year fixed effects) is displayed in Table A4.8 for both high-skilled and low-skilled workers, along with the difference between the two groups. The long-term contract indexes are divided into deciles (D1 to D9) and provide insights into the variation in contract types across the sample. High-skilled workers tend to have higher long-term contract indexes compared to low-skilled workers across all deciles. This suggests that high-skilled individuals generally have a higher likelihood of having long-term contracts. This is why we observe that the differences between the long-term contract indexes of high-skilled and low-skilled workers (HS−LS) are positive across all deciles.

We now turn our attention to our second step to analyze the impact of housing costs on long-term employment contracts. The estimation results are presented in Table 4.6.1. This table first examines how differences in housing costs across urban areas can influence the likelihood of both high-skilled and low-skilled, of having long-term contracts. It, more importantly, underlines how the disparity in the probability of having long-term contracts between the two skill levels is affected by urban areas differences in housing costs.

Housing price index (Price Index) represents the direct measure of housing prices, while the instrumented housing price index (Instrumented Price Index) is derived using inverse housing supply elasticities interacted with year dummies to treat endogeneity. The estimated urban-area-skill-year fixed effect for high and low-skilled is denoted as HS LT index and LS LT index, respectively.

Columns (1)-(3) display the results obtained through ordinary least squares (OLS) regression, while columns (4)-(6) present the instrumental variable (IV) results. Notably, we observe a reversal in the sign of the house price index coefficients when we switch from OLS to IV estimation. This change can be attributed to the instrumental variable method, which helps to adress endogeneity issues in the analysis.

[18]The reference group of low-skilled individuals in Toulon in 2010 was selected because they represent the average likelihood among low-skilled individuals in our sample for holding long-term contracts.

The results presented in columns (4) and (5) of the table offer contrasting insights into the impact of housing costs on high and low-skilled individuals. In column (4), the analysis reveals a negative but non-significant effect of housing costs on the probability of high-skilled workers in securing long-term contracts. This suggests that rising housing costs do not have a statistically discernible impact on the job stability of high-skilled workers. On the other hand, the findings in column (5) paint a different picture of low-skilled workers. Here, the estimated coefficient indicates that higher housing costs are significantly associated with a reduced probability of low-skilled workers in obtaining long-term contracts.

In columns (3) and (6) of the table, we focus on the dependent variable representing the difference between the high-skilled and the low-skilled. This variable captures the disparity in the probability of obtaining long-term employment contracts between high-skilled and low-skilled workers. Our preferred specification, treating endogeneity, in column (6), indicates that increasing housing costs increases the disparity in long-term employment between the two skill levels. The housing price index coefficient is indeed positive and statistically significant.

Specifically, the coefficient in column (6) implies that a 10% increase in housing costs is associated with a 0.14 percentage point increase in the disparity in long-term contract probabilities between high-skilled and low-skilled workers. In other words, this implies that a one standard deviation increase in the price index increases the disparity between high-skilled and low-skilled by 0.458 percentage points.[19] This finding implies that rising housing costs have an effect on the job stability of low-skilled workers compared to high-skilled workers.

As outlined in section 4.4, such effect could be attributed to the mechanisms explained in urban labor economics models. As housing costs increase in an urban area, low-skilled workers may locate further away from employment centers than high-skilled thus having reduced access to information about job opportunities, less intensive job searches, and a reduced spatial search horizon due to anticipated long commuting distances. Low-skilled workers may therefore lack the networks and resources that are already more readily available to high-skilled workers, possibly hindering their probability of finding a long-term contract. Such effect could also be attributed to the urgency, driven by higher housing cost burdens for low-skilled, to find employment, and to therefore accept sub-optimal job matches. This is consistent with search and matching models in labor economics, which suggest that the need for immediate employment can lead to quicker but less ideal job matches. Indeed, for low-skilled workers facing high housing costs, the utility derived from being employed (even in a less ideal job) may outweigh the utility of remaining unemployed

[19]The standard deviation of the price index is 0.327.

and searching for a better match. This necessity-driven approach to employment can lead to more short-term contract work as opposed to stable, long-term employment. This situation is less prevalent among high-skilled workers who might have better financial cushions and can afford to be more selective in their job search.

Results for control variables in Table 4.6.1 also outline the effects of urban area labor market characteristics on the probability of having open-ended contracts and the disparity between skill levels. While differences in housing costs across urban areas tend to increase the long-term contract disparities between skill levels, a higher concentration of jobs in an urban area still benefits both skill levels and lowers their differences in finding long-term contracts. A 10% increase in the number of jobs in an urban area would therefore decrease by 0.04% the difference in long-term contract probabilities across skills. The unemployment rate coefficient in column (6) of Table 4.6.1 tends to outline that low-skilled suffer more than high-skilled in terms of the probability of finding a long-term contract in urban areas facing high-unemployment rates: a 10% increase in its share rising the long-term contract probability gap between the skills by 0.79%. This is not surprising as unemployment affects to a greater extent low-skilled workers, who with fewer job opportunities, and increased competition to find a job might turn themselves to less stable jobs.

Additional results: Gender. Table 4.6.2 presents the results of the effects of housing costs on the likelihood of obtaining long-term employment contracts disaggregated by gender. Columns (1)-(3) focus on male workers, while columns (4)-(6) address female workers. Within these, columns (1) and (4) examine high-skilled (HS) men and women, respectively, columns (2) and (5) look at low-skilled (LS) men and women, and columns (3) and (6) explore the disparities between HS and LS individuals within each gender.

For high-skilled men (1) and women (4), housing costs do not appear to influence the probability of obtaining long-term contracts, suggesting that this group may be less sensitive to fluctuations in housing costs. On the other hand, column (5) indicates that low-skilled women are adversely affected by high housing costs, while no significant effect is observed for low-skilled men (2). Columns (3) and (6) present the results of the gender-specific disparities between HS and LS workers. For men, the disparity is positive but falls short of statistical significance. In contrast, for women, this disparity is both positive and significant, implying that low-skilled women are particularly vulnerable to the impact of housing costs on job stability.

Table 4.6.1.: The effect of housing costs on the probability of long-term contracts

Dependent Variables:	HS LT index	LS LT index	HS index - LS index	HS LT index	LS LT index	HS index - LS indexx
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Price index	0.033***	0.023***	0.009***			
	(0.004)	(0.005)	(0.003)			
Instrumented Price index				−0.006	−0.020**	0.014**
				(0.008)	(0.009)	(0.006)
% jobs by economic sector						
Agriculture	−0.247**	−0.751***	0.504***	−0.562***	−1.102***	0.541***
	(0.118)	(0.122)	(0.093)	(0.120)	(0.142)	(0.110)
Industry	0.056*	0.091***	−0.035	−0.046	−0.023	−0.023
	(0.030)	(0.032)	(0.023)	(0.035)	(0.036)	(0.027)
Construction	−0.568***	−0.686***	0.118	−0.349***	−0.442***	0.092
	(0.111)	(0.118)	(0.087)	(0.124)	(0.117)	(0.094)
Commerce, Transport, Services	0.071***	0.100***	−0.029*	0.076***	0.105***	−0.004*
	(0.024)	(0.023)	(0.017)	(0.024)	(0.023)	(0.017)
Public admin.	−0.135***	−0.062**	−0.073***	−0.223***	−0.161***	−0.007***
	(0.023)	(0.025)	(0.017)	(0.034)	(0.036)	(0.022)
% Unemployment	0.113***	0.043	0.070**	0.036	−0.043	0.079***
	(0.035)	(0.041)	(0.030)	(0.038)	(0.046)	(0.031)
Log(employment)	0.005***	0.009***	−0.004***	0.010***	0.014***	−0.004***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Constant	−0.324***	−0.348***	0.024	0.080	0.103	−0.023
	(0.047)	(0.050)	(0.038)	(0.091)	(0.098)	0.065)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N urban areas	151	151	151	151	151	151
N years	9	9	9	9	9	9
Observations	1,359	1,359	1,359	1,359	1,359	1,359
R^2	0.998	0.997	0.268	0.998	0.997	0.267

Notes: The table presents the second step regression of the two-step approach as in equation (3) for columns (1)-(4) and as in equation (5) in columns (5)-(8). Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, ** and *** denote statistical significance at the 10, 5 and 1 % level.

Previous research highlights gender differences in preferences over commuting. Women favor shorter commutes due to family responsibilities, thus impacting negatively their labor market outcomes (Farré et al. 2023; Petrongolo and Ronchi, 2020; Le Barbanchon et al. 2019; among others). We can hypothesize that in a context of high housing costs, such patterns would be reinforced for low-skilled women therefore suffering from a "double" burden.

Additional results: Sectors. It is important to note that the majority of workers in our estimation sample (62%) are employed in the service sector, followed by commerce (16%) and industry (15%). As the construction and agriculture sectors represent a small amount of the workforce in the sample (7% and 0.15% respectively), the accurate computation of urban-area-skill-year fixed-effects could be questioned on such small sub-samples. This is why we perform the heterogeneity analysis on the three previously mentioned main represented sectors. Across diverse sectors of the economy, our findings consistently reveal substantial disparities in the impact of housing costs on the likelihood of obtaining long-term employment contracts across various economic sectors, with low-skilled workers experiencing clear disadvantages compared to their high-skilled counterparts. As in our main model, housing costs significantly impact the likelihood of securing long-term employment contracts, with high-skilled workers consistently being less impacted than their low-skilled counterparts. In the service sector (Table A4.10), high-skilled workers exhibit a positive disparity (0.020***) in their employment prospects compared to low-skilled workers, suggesting higher job instability for the latter group. A similar pattern emerges in the industry sector (Table A4.11), where we observe a positive disparity that further disadvantages low-skilled workers (0.031***). Finally, the commerce sector (Table A4.12) reinforces this trend, with a coefficient of 0.053*** indicating a positive disparity that increases the challenges faced by low-skilled workers in securing long-term contracts.

Additional results: Age. We also consider the age groups of workers. Table A4.13 focuses on workers between the ages of 15 and 34, the coefficient in column (3) is not statistically significant indicating that for this younger age group, housing costs do not have a significant differential impact on the disparities of the likelihood of obtaining long-term employment contracts between high and low-skilled individuals. We can therefore hypothesize that all young workers are affected similarly by housing costs at the beginning of their careers, with no differences across skills. On the other hand, in Table A4.14, which considers workers above the age of 34, we find in column (3) a significant and positive effect of housing costs on the disparities between high-skilled and low-skilled workers. This indicates that, as career path advance, and as housing costs increase, low-skilled older workers face higher job insecurity and lower probability of securing long-term contracts compared to their

Table 4.6.2.: The effect of housing costs on the probability of long-term contracts by gender

	Men			*Women*		
Dep var.	HS LT index	LS LT index	HS index - LS index	HS LT index	LS LT index	HS index - LS index
	(1)	(2)	(3)	(4)	(5)	(6)
Price index instrumented	0.001	−0.011	0.012	−0.012	−0.029**	0.017**
	(0.009)	(0.008)	(0.007)	(0.012)	(0.012)	(0.009)
% jobs by economic sector						
Agriculture	−0.682***	−1.113***	0.431***	−0.281	−0.955***	0.674***
	(0.120)	(0.149)	(0.127)	(0.197)	(0.215)	(0.164)
Industry	−0.038	−0.037	−0.001	−0.067	0.024	−0.091**
	(0.034)	(0.037)	(0.033)	(0.051)	(0.050)	(0.036)
Construction	−0.145	−0.236*	0.091	−0.956***	−0.948***	−0.008
	(0.110)	(0.121)	(0.117)	(0.191)	(0.181)	(0.137)
Commerce, Transport, Services	0.072***	0.068***	0.003	0.100***	0.180***	−0.079***
	(0.021)	(0.023)	(0.021)	(0.036)	(0.032)	(0.023)
Public admin.	−0.192***	−0.174***	−0.018	−0.277***	−0.141***	−0.136***
	(0.036)	(0.034)	(0.026)	(0.045)	(0.050)	(0.031)
Share unemployment	0.104***	−0.072*	0.176***	−0.094	0.018	−0.112**
	(0.038)	(0.044)	(0.037)	(0.070)	(0.068)	(0.050)
Log(employment)	0.010***	0.012***	−0.002	0.007***	0.015***	−0.008***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Constant	−0.079	−0.029	−0.050	0.121	0.077	0.044
	(0.097)	(0.095)	(0.082)	(0.133)	(0.138)	(0.093)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N urban areas	151	151	151	151	151	151
N years	9	9	9	9	9	9
Observations	1,359	1,359	1,359	1,359	1,359	1,359
R^2	0.999	0.999	0.234	0.999	0.999	0.121

Notes: The table presents the second step regression of the two-step approach as in equation as in equation (5) by gender. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, ** and *** denote statistical significance at the 10, 5, and 1 % level.

high-skilled counterparts. This could be explained by the fact that housing choices change over time: housing consumption increases, individuals try to access home ownership, while labor market trajectories diverge between skill levels, and tend to stabilize more for the high-skilled than the low-skilled.

4.6.3. Analysis of the impact of housing costs on probability of promotion

We now shift our focus to examine the impact of housing costs on the probability of promotion. The estimation results are presented in Table 4.6.3. This table explores how housing costs influence the likelihood of individuals, both high-skilled and low-skilled, being promoted. Columns (1)-(3) display the results obtained through ordinary least squares (OLS) regression, while columns (4)-(6) present the instrumental variable (IV) results. In column (6), the coefficient represents the difference in promotion probabilities between high-skilled and low-skilled workers. The coefficient is not statistically significant, indicating that there is no significant difference in the probability of promotion between high-skilled and low-skilled workers.

When dividing the sample by gender in Table 4.6.4, we do not observe any discernible impact of housing costs on the likelihood of promotion for either men or women. We therefore do not find any significant influence of housing prices on promotion probabilities, suggesting that other factors may play a more dominant role in shaping career advancement opportunities for workers of varying skill levels.

It is important to note that the promotion measure adopted in this study is one of several possible approaches. Future research could explore alternative measures of promotion to gain a more comprehensive understanding of the factors influencing career advancement across diverse skills.

Table 4.6.3.: The effect of housing costs on the probability of promotion

Dep var.	HS LT index	LS LT index	HS index - LS index	HS LT index	LS LT index	HS index - LS index
	(1)	(2)	(3)	(4)	(5)	(6)
Price index	−0.022	0.033***	−0.055***	−0.013	0.017	−0.031
	(0.017)	(0.007)	(0.017)	(0.031)	(0.015)	(0.033)
% jobs by economic sector						
Agriculture	−0.420	−0.094	−0.326	−0.356	−0.217	−0.139
	(0.403)	(0.220)	(0.368)	(0.435)	(0.236)	(0.408)
Industry	−0.023	0.055	−0.078	−0.001	0.015	−0.016
	(0.225)	(0.061)	(0.251)	(0.271)	(0.077)	(0.277)
Construction	−0.131	−0.286**	0.155	−0.165	−0.221	0.056
	(0.389)	(0.138)	(0.408)	(0.419)	(0.153)	(0.465)
Commerce, Transport, Services	0.021	−0.014	0.035	0.019	−0.010	0.029
	(0.218)	(0.054)	(0.256)	(0.264)	(0.069)	(0.268)
Public admin.	−0.140	0.029	−0.169	−0.122	−0.006	−0.116
	(0.218)	(0.057)	(0.248)	(0.263)	(0.080)	(0.272)
Share unemployment	−0.082	−0.018	−0.064	−0.073	−0.035	−0.038
	(0.113)	(0.055)	(0.115)	(0.111)	(0.057)	(0.117)
Log(employment)	0.011**	0.005**	0.006	0.010**	0.007**	0.003
	(0.004)	(0.002)	(0.004)	(0.005)	(0.003)	(0.005)
Constant	0.811***	0.324***	0.487	0.724*	0.490***	0.233
	(0.300)	(0.112)	(0.298)	(0.419)	(0.186)	(0.438)
N urban areas	151	151	151	151	151	151
R^2	0.560	0.882	0.127	0.558	0.874	0.106

Notes: The table presents the second step regression of the two-step approach without instrumenting for columns (1)-(4) and instrumented as in equation (8) in columns (5)-(8). Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, **, and *** denote statistical significance at the 10, 5, and 1 % level.

Table 4.6.4.: The effect of housing costs on the probability of promotion by gender

	Men			*Women*		
Dep var.	HS LT index	LS LT index	HS index - LS index	HS LT index	LS LT index	HS index - LS index
	(1)	(2)	(3)	(4)	(5)	(6)
Price index instrumented	−0.032	0.021	−0.053	0.029	0.013	0.015
	(0.037)	(0.018)	(0.041)	(0.053)	(0.020)	(0.049)
% jobs by economic sector						
Agriculture	−0.437	−0.210	−0.227	−0.121	−0.260	0.138
	(0.498)	(0.256)	(0.497)	(0.693)	(0.277)	(0.681)
Industry	0.005	0.047	−0.042	0.016	−0.002	0.018
	(0.242)	(0.096)	(0.293)	(0.388)	(0.100)	(0.335)
Construction	0.161	−0.196	0.357	−0.971	−0.276	−0.695
	(0.464)	(0.207)	(0.543)	(0.671)	(0.203)	(0.636)
Commerce, Transport, Services	0.050	−0.027	0.077	−0.047	0.009	−0.056
	(0.229)	(0.083)	(0.279)	(0.374)	(0.092)	(0.313)
Public admin.	−0.136	0.039	−0.175	−0.061	−0.054	−0.007
	(0.242)	(0.098)	(0.295)	(0.373)	(0.106)	(0.323)
Share unemployment	0.057	−0.066	0.123	−0.357*	0.036	−0.393*
	(0.133)	(0.063)	(0.144)	(0.212)	(0.082)	(0.215)
Log(employment)	0.016***	0.008**	0.009	−0.002	0.005	−0.008
	(0.006)	(0.003)	(0.006)	(0.008)	(0.003)	(0.008)
Constant	0.399	0.003	0.396	0.445	0.527**	−0.082
	(0.460)	(0.212)	(0.516)	(0.643)	(0.235)	(0.608)
N urban areas	151	151	151	151	151	151
R^2	0.48046	0.81539	0.11018	0.36211	0.79125	0.03380

Notes: The table presents the second step regression of the two-step approach as in equation (7) for columns (1)-(4) and as in equation (8) in columns (5)-(8). Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, **, and *** denote statistical significance at the 10, 5, and 1 % level.

4.7. Conclusions

In this study, we have empirically examined the relationship between housing prices and labor market disparities among high-skilled and low-skilled workers in urban areas across France. Using rich administrative panel data over the period (2010-2018), and a between-city analysis, we shed light on the impact of housing costs on two critical labor market outcomes: the likelihood of securing long-term employment contracts and job promotions. We use a methodology that allows us to remove the spatial sorting bias, as well as the potential reverse causality.

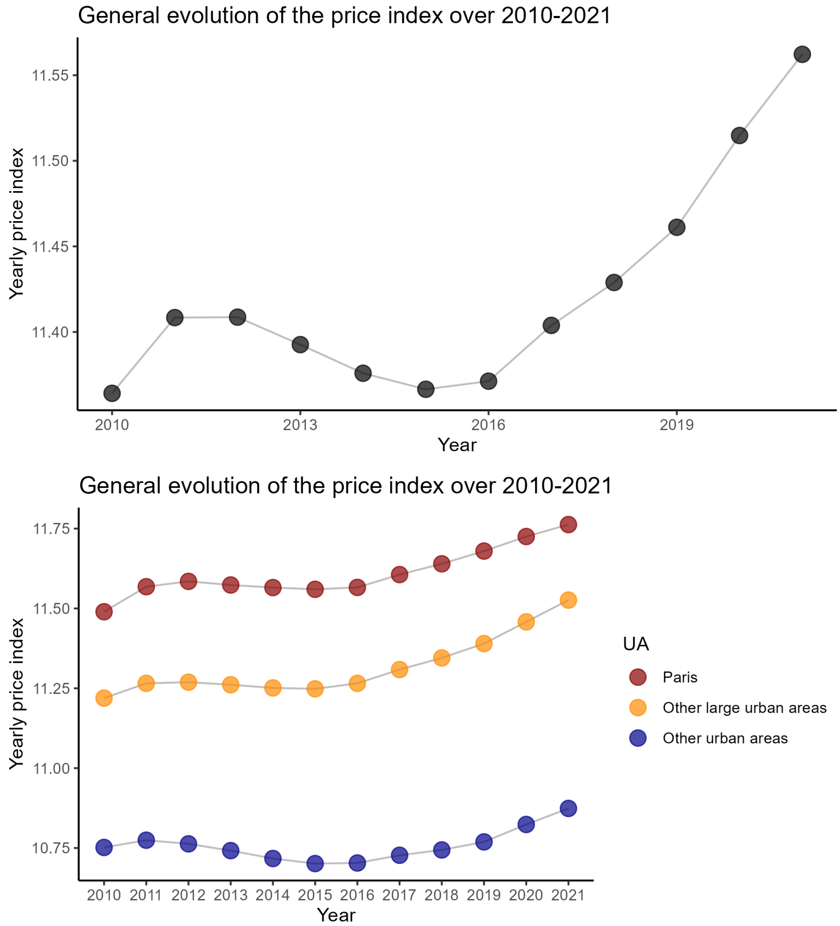
Our findings underscore the significance of housing price differentials across urban areas as a contributing factor to labor market disparities. Notably, we observed that these disparities are more pronounced in the context of long-term employment contracts, indicating that higher housing costs tend to increase inequalities in job stability between high-skilled and low-skilled individuals. Specifically, we find that a 10% increase in housing costs is associated with a 0.14 percentage point increase in the disparity in long-term contract probabilities between high-skilled and low-skilled workers. This implies that a one standard deviation increase in the price index increases the disparity between high-skilled and low-skilled by 0.458 percentage points. Across different economic sectors, low-skilled workers consistently experience disadvantages compared to their high-skilled counterparts when it comes to the impact of housing costs on long-term employment contracts. Moreover, age also plays a role, with low-skilled workers above 34 years facing higher job insecurity as housing costs increase further increasing the disparities in long-term employment contracts. It is additionally crucial to emphasize that these disparities in labor market outcomes were particularly pronounced among women, with low-skilled women being more affected by higher housing costs. This emphasizes the need for targeted policy interventions to address these issues and promote equitable employment opportunities.

We do not find a significant influence of housing prices on promotion probabilities, suggesting that other factors may play a more dominant role in shaping career advancement opportunities for workers of varying skill levels. Further research in this domain is needed to explore additional nuances and potential policy implications.

4.8. Appendix

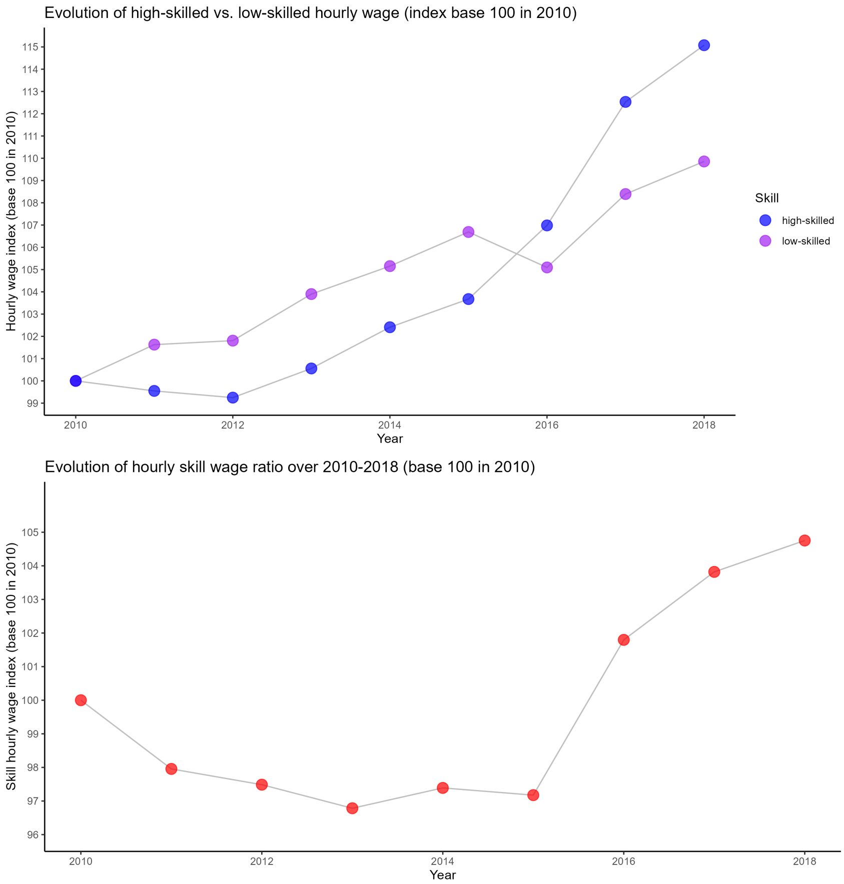
A. Figures and tables

Figure A4.1.: General evolution of housing prices (2010-2021)



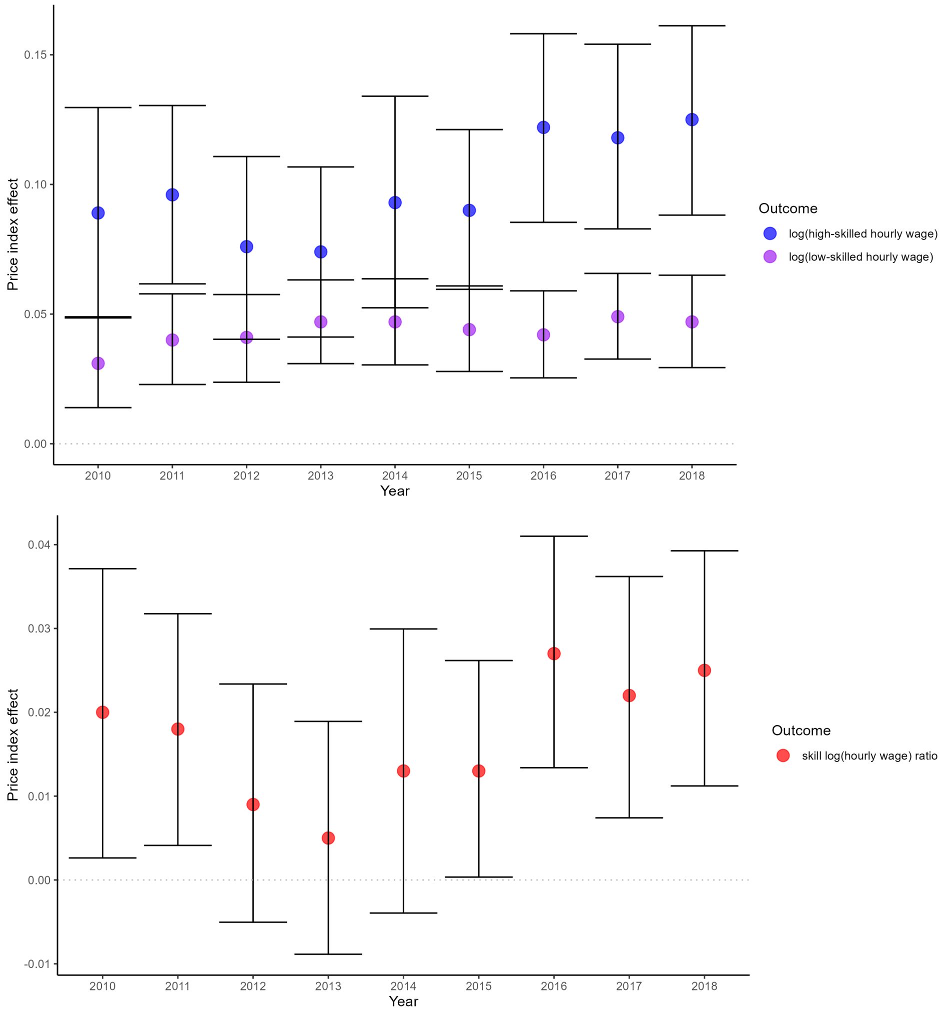
Notes: These figures depict housing price trends in France. The left graph illustrates the overall evolution of housing prices from 2010 to 2021 across 151 urban areas, while the right graph focuses on the evolution in specific French urban areas. Source: DV3F housing transaction data. Own computation.

Figure A4.2.: General evolution of hourly wages by skill level (2010-2018)



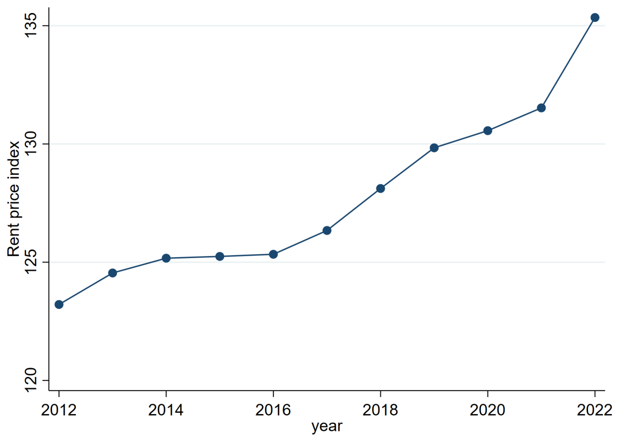
Notes: These figures illustrate the trends in high-skilled and low-skilled net hourly wages over the period 2010-2018. Source: Panel d'actifs of the EDP (2010-2018). Own computation.

Figure A4.3.: Wage compensation of housing costs across skill levels



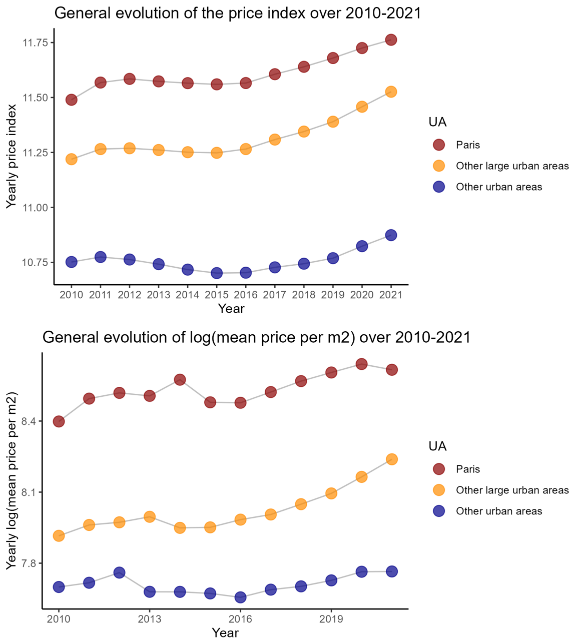
Notes: These figures illustrate the hourly net wage distribution evolution.

Figure A4.4.: Evolution of Rent Price Indices Over Time



Notes: The figure shows the evolution of the rent price index over 2000-2022 with 1998 as the base year. Source : INSEE, Indice de référence des loyers.

Figure A4.5.: Housing price evolution by French urban areas



Notes: The figures presented below offer a detailed analysis of the housing price trends by French urban areas from 2010 to 2021. Other large urban areas include: Lyon, Marseille-Aix-en-Provence, Bordeaux, Montpellier, Toulouse, Toulon and Nantes. Source: DV3F dataset.

Table A4.1.: Descriptive statistics (in %) on individuals present in the sample

Individual characteristics	
Age	
15-29 y.o	28.1
30-39 y.o	27.9
40-49 y.o	27.0
50-59 y.o	16.6
Above 60	0.4
Sex	
Male	57.7
Female	42.3
Occupations	
High-level occupations	19.7
Intermediate occupations	20.3
Low level white-collars	32.9
Blue-collar workers	27.1
Skill	
High-skilled	19.7
Intermediaries	20.3
Low-skilled	60.0
UA size	
Less than 100,000 inhabitants	10.3
100, 000 to 500,000 inhabitants	27.9
500, 000 to 1,000,000 inhabitants	11.6
More than 1,000,000 inhabitants (except Paris)	16.5
Paris urban area	33.8
Characteristics of main job held	
Terms of employment	
Full time	77.3
Part time	22.7
Type of contract	
Open-ended	80.7
Short term	19.3
Economic sector	
Agriculture	0.2
Industry	14.5
Construction	7.1
Commerce	15.9
Services	62.3
N individuals	463,785
N urban areas	151

Notes: This table provides descriptive statistics for the characteristics of individuals within the estimation sample.

Table A4.2.: Distribution of skills per urban area type

Urban area size	% high-skilled	% of low-skilled	% intermediaries
Paris urban area	28.53	52.03	19.44
$> 1,000,000$ inhabitants	19.98	57.81	22.21
$< 1,000,000$ inhabitants	17.54	61.72	20.74
$< 500,000$ inhabitants	12.95	66.52	20.53
$< 100,000$ inhabitants	10.76	69.92	19.32

Source: Panel of active individuals from the EDP (2010-2018). The sample includes 3, 724, 905 observations, 463, 785 individuals, and 151 urban areas.

Table A4.3.: Distribution of net wage, n° hours worked and net hourly wage

	D1	D2	D3	D4	D5	D6	D7	D8	D9	Mean	SD
	10%	20%	30%	40%	50%	60%	70%	80%	90%		
Net wage	5919	10,784	14,808	17,510	19,907	22,600	26,061	31,157	41,253	23,309	19,554
Hours worked	585	1040	1440	1670	1815	1820	1827	1908	2028	1547	567
Hourly wage	8.3	9.1	9.8	10.6	11.7	13.1	14.8	17.8	23.4	14.4	9.4

Source: Panel of active individuals from the EDP (2010-2018). The sample includes 463, 785 individuals, and 151 urban areas.

Table A4.4.: Desciptive Statistics on the urban areas present in the sample

Variables	Count	Mean	SD	Min	Max
Panel A					
Average House Price in 2010	1359	161566.20	41324.48	96398.2	316422.7
Average House Price in 2018	1359	169588.60	53395.56	88238.98	366066.8
Average House Price over (2010-2018)	1359	164440.60	48170.44	91291.13	327386.4
Average House Price per Square Metre in 2010	1359	1975.72	638.91	1090.34	4438.97
Average House Price per Square Metre in 2018	1359	1983.82	776.02	1013.8	5263.97
Average House Price per Square Metres (2010-2018)	1359	1976.48	729.80	1070.33	4942.71
Price Index in 2010	1359	10.78	0.30	10.26	11.8
Price Index in 2018	1359	10.76	0.35	10.11	11.93
Average Price Index over 2010-2018	1359	10.76	0.32	10.17	11.87
Inverse Housing Supply Elasticity 2000-2010	1359	1.75	0.59	0.42	3.68
Panel B					
Population	1359	281456.70	1017958	15187	1.25e+07
Active Population	1359	133593.20	513566	7363.46	6257414
Number of Unemployed	1359	17656.84	62830.18	730	792853
Unemployment to Active Population Ratio	1359	0.14	0.03	0.05	0.26
Agricultural Employment Share	1359	0.02	0.01	0	0.09
Industrial Employment Share	1359	0.14	0.06	0.04	0.37
Construction Employment Share	1359	0.07	0.01	0.03	0.13
Trade, Transport, and Services Employment Share	1359	0.43	0.05	0.29	0.57
Public Administration Employment Share	1359	0.34	0.05	0.19	0.57

Notes: The table shows descriptive statistics of housing at urban area level (Panel A) and the characteristics of urban areas (Panel B). There are 151 urban areas x 9 years observations (1,359).

Table A4.5.: Wage compensation across skill levels

Dep. variable	log (average high-skilled hourly wage) in UA								
	2010	2011	2012	2013	2014	2015	2016	2017	2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Price index in UA	0.089***	0.096***	0.076***	0.074***	0.093***	0.090***	0.122***	0.118***	0.125***
	(0.021)	(0.017)	(0.018)	(0.017)	(0.021)	(0.016)	(0.018)	(0.018)	(0.018)
Constant	2.163***	2.082***	2.307***	2.346***	2.154***	2.188***	1.863***	1.935***	1.882***
	(0.221)	(0.188)	(0.192)	(0.179)	(0.222)	(0.167)	(0.197)	(0.194)	(0.199)
R^2	0.112	0.170	0.107	0.118	0.120	0.184	0.227	0.225	0.234
Dep. variable	log(average low-skilled hourly wage) in UA								
Price index in UA	0.031***	0.040***	0.041***	0.047***	0.047***	0.044***	0.042***	0.049***	0.047***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Constant	2.032***	1.946***	1.948***	1.899***	1.912***	1.955***	1.968***	1.923***	1.957***
	(0.096)	(0.096)	(0.092)	(0.088)	(0.090)	(0.090)	(0.091)	(0.090)	(0.097)
R^2	0.078	0.122	0.131	0.182	0.174	0.159	0.142	0.189	0.155
Dep. variable	log(high-skilled hourly wage)/log(low-skilled hourly wage) in UA								
Price index in UA	0.020**	0.018**	0.009	0.005	0.013	0.013**	0.027***	0.022***	0.025***
	(0.009)	(0.007)	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	1.103***	1.116***	1.210***	1.252***	1.166***	1.157***	1.018***	1.075***	1.037***
	(0.094)	(0.076)	(0.078)	(0.076)	(0.092)	(0.070)	(0.075)	(0.078)	(0.076)
R^2	0.034	0.042	0.011	0.003	0.015	0.027	0.092	0.057	0.078
N urban areas	151	151	151	151	151	151	151	151	151

Notes: Log(average hourly net wage) of high-skilled, low-skilled, and skill wage ratio in the urban area explained by the price index in each urban area each year. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A4.6.: First-stage regression: instrumenting housing prices

Dep. var.:	Price index
	(1)
Constant	10.11***
	(0.1501)
Inverse supply elasticity x Year	
Inv. elast. × 2010	0.2015***
	(0.0265)
Inv. elast. × 2011	0.2114***
	(0.0260)
Inv. elast. × 2012	0.2139***
	(0.0256)
Inv. elast. × 2013	0.2073***
	(0.0253)
Inv. elast. × 2014	0.2085***
	(0.0250)
Inv. elast. × 2015	0.2093***
	(0.0254)
Inv. elast. × 2016	0.2116***
	(0.0256)
Inv. elast. × 2017	0.2108***
	(0.0261)
Inv. elast. × 2018	0.2035***
	(0.0267)
N urban areas	151
N years	9
Year FE	Yes
Observations	1,359
F-Stat.	55.5
R^2	0.921

Notes: This table presents the regression results using the inverse housing supply elasticity as an instrumental variable for housing prices. We control for the log(employment) in the urban area, the unemployment rate, as well as the employment share within several sectors, including agriculture, industry, construction, commerce, transport, services, and public administration.

Table A4.7.: First stage estimate of the two-step approach

Dep. var. Long term contract (0/1)	(1)
Age	-0.1699
	(341.3)
Age^2	-0.0005***
	(4.11×10^{-6})
Economic sector	
Agriculture	-0.2185***
	(0.0132)
Industry	Ref.
Construction	0.0701***
	(0.0036)
Commerce	-0.0501***
	(0.0025)
Services	-0.2710***
	(0.0023)
NA	-0.1540***
	(0.0411)
Fixed-effects	
UA x SKILL X YEAR	Yes
Individual	Yes
Clustered SE	Indiv. level
Observations	3,724,905
R^2	0.63433

Notes: The table presents the first step regression of equation (2). Industry is used as the reference category for the economic sector in the regression. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A4.8.: Distribution of LT contract indexes

	D1 10%	D2 20%	D3 30%	D4 40%	D5 50%	D6 60%	D7 70%	D8 80%	D9 90%	Mean	SD
High-skilled	0.104	0.292	0.498	0.698	0.904	1.113	1.321	1.519	1.705	0.906	0.555
Low-skilled	0.035	0.236	0.442	0.650	0.852	1.054	1.255	1.463	1.665	0.852	0.565
Difference HS−LS	0.011	0.025	0.035	0.043	0.053	0.061	0.071	0.083	0.104	0.055	0.038

Notes: This table displays the distribution of Long-Term (LT) contract indexes for both high-skilled and low-skilled workers, along with the difference between the two groups. The LT contract indexes are divided into deciles (D1 to D9) and provide insights into the variation in contract types across the sample. The table includes a total of 1,359 observations spanning 151 urban areas over 9 years.

Table A4.9.: Descriptive statistics on the independent variables in step 2 (equation 4.5)

Statistic	N	Mean	St. Dev.	Min	Max
Price index	1,359	10.756	0.327	10.108	11.934
% Unemployment	1,359	0.138	0.029	0.054	0.262
Log(employment)	1,359	10.860	1.019	8.935	15.850
% jobs by economic sector					
Agriculture	1,359	0.019	0.011	0.002	0.092
Industry	1,359	0.137	0.061	0.043	0.366
Construction	1,359	0.067	0.013	0.028	0.129
Commerce, Transport, Services	1,359	0.425	0.049	0.290	0.571
Public admin.	1,359	0.344	0.055	0.192	0.572

Notes: This table presents descriptive statistics for the independent variables used in the second step of the analysis (equation 4.5). The table includes information on the housing price index, unemployment rates, employment levels (logged), and the distribution of jobs across economic sectors. The data comprises observations from 151 urban areas over a span of 9 years, resulting in a total of 1,359 observations.

Table A4.10.: Impact of housing costs on long-term contracts in the service sector

Dep. var.	HS LT index (1)	LS LT index (2)	HS index - LS index (3)
Price index instrumented	−0.017	−0.037***	0.020***
	(0.010)	(0.009)	(0.008)
% jobs by economic sector			
Agriculture	−0.135	−0.834***	0.699***
	(0.153)	(0.182)	(0.140)
Industry	−0.063	−0.071*	0.008
	(0.042)	(0.042)	(0.035)
Construction	−0.484***	−0.346**	−0.138
	(0.137)	(0.146)	(0.124)
Commerce, Transport, Services	0.133***	0.100***	0.033*
	(0.028)	(0.025)	(0.020)
Public admin.	−0.255***	−0.173***	−0.082***
	(0.041)	(0.038)	(0.026)
Share unemployment	0.110**	0.079	0.031
	(0.047)	(0.055)	(0.039)
Log(employment)	0.013***	0.018***	−0.006***
	(0.002)	(0.002)	(0.002)
Constant	0.059	0.145	−0.086
	(0.104)	(0.106)	(0.081)
Year FE	Yes	Yes	Yes
N urban areas	151	151	151
N years	9	9	9
Observations	1,359	1,359	1,359
R^2	0.998	0.998	0.175

Notes: The table presents the second step regression of the two-step approach as in equation (5) for workers in the service sector. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, **, and *** denote statistical significance at the 10, 5, and 1 % level.

Table A4.11.: Impact of housing costs on long-term contracts in the industry sector

Dep. var.	HS LT index	LS LT index	HS index - LS index
	(1)	(2)	(3)
Price index instrumented	−0.00005	−0.031*	0.031***
	(0.013)	(0.016)	(0.010)
% jobs by economic sector			
Agriculture	−0.887***	−1.413***	0.525***
	(0.201)	(0.232)	(0.157)
Industry	0.052	−0.043	0.095***
	(0.046)	(0.054)	(0.033)
Construction	0.237	0.941***	−0.703***
	(0.175)	(0.192)	(0.121)
Commerce, Transport, Services	−0.141***	−0.154***	0.013
	(0.027)	(0.031)	(0.019)
Public admin.	−0.122***	−0.330***	0.208***
	(0.047)	(0.061)	(0.033)
Share unemployment	−0.075	−0.079	0.004
	(0.064)	(0.076)	(0.043)
Log(employment)	0.0001	0.002	−0.002
	(0.002)	(0.003)	(0.002)
2011	−5.468***	−5.462***	−0.006
	(0.005)	(0.006)	(0.004)
Constant	0.116	0.429**	−0.313***
	(0.147)	(0.176)	(0.104)
Year FE	Yes	Yes	Yes
N urban areas	151	151	151
N years	9	9	9
Observations	1,359	1,359	1,359
R^2	0.999	0.999	0.123

Notes: The table presents the second step regression of the two-step approach as in equation (5) for workers in the industry sector. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, **, and *** denote statistical significance at the 10, 5, and 1 % level.

Table A4.12.: Impact of housing costs on long-term contracts in the commerce sector

Dep. var.	HS LT index (1)	LS LT index (2)	HS index - LS index (3)
Price index instrumented	0.070***	0.017	0.053***
	(0.012)	(0.013)	(0.009)
% jobs by economic sector			
Agriculture	−0.274	−0.807***	0.533***
	(0.195)	(0.175)	(0.144)
Industry	0.206***	0.117**	0.089**
	(0.047)	(0.048)	(0.042)
Construction	−0.671***	−0.637***	−0.034
	(0.145)	(0.154)	(0.147)
Commerce, Transport, Services	0.014	0.077***	−0.062***
	(0.023)	(0.028)	(0.023)
Public admin.	0.071**	0.041	0.030
	(0.035)	(0.037)	(0.034)
Share unemployment	−0.125**	−0.328***	0.203***
	(0.054)	(0.061)	(0.043)
Log(employment)	−0.004*	0.005**	−0.009***
	(0.002)	(0.002)	(0.002)
Constant	−0.713***	−0.277*	−0.436***
	(0.142)	(0.142)	(0.103)
Year FE	Yes	Yes	Yes
N urban areas	151	151	151
N years	9	9	9
Observations	1,359	1,359	1,359
R^2	0.995	0.995	0.091

Notes: The table presents the second step regression of the two-step approach as in equation (5) for workers in the commerce sector. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, ** and *** denote statistical significance at the 10, 5, and 1 % level.

Table A4.13.: Impact of housing costs on LT contracts for young workers (Aged 15-34)

Dep. var.	HS LT index (1)	LS LT index (2)	HS index - LS index (3)
Price index instrumented	−0.002	0.001	−0.003
	(0.017)	(0.011)	(0.017)
% jobs by economic sector			
Agriculture	−0.982***	−1.295***	0.313
	(0.273)	(0.192)	(0.281)
Industry	−0.023	0.007	−0.030
	(0.069)	(0.049)	(0.073)
Construction	0.032	−0.298	0.330
	(0.230)	(0.183)	(0.242)
Commerce, Transport, Services	0.038	0.109***	−0.071*
	(0.044)	(0.037)	(0.039)
Public admin.	−0.250***	−0.118***	−0.132**
	(0.059)	(0.045)	(0.062)
Share unemployment	0.161**	−0.058	0.219***
	(0.080)	(0.057)	(0.076)
Log(employment)	0.017***	0.017***	0.0003
	(0.004)	(0.003)	(0.003)
Constant	−0.134	−0.298**	0.165
	(0.185)	(0.124)	(0.188)
Year FE	Yes	Yes	Yes
N urban areas	151	151	151
N years	9	9	9
Observations	1,359	1,359	1,359
R^2	0.737	0.852	0.113

Notes: The table presents the second step regression of the two-step approach as in equation (5) for workers between 15 and 30 years old. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, ** and *** denote statistical significance at the 10, 5, and 1 % level.

Table A4.14.: Impact of housing costs on LT contracts for older workers (Above 34)

Dep. var.	HS LT index	LS LT index	HS index - LS index
	(1)	(2)	(3)
Price index instrumented	−0.015*	−0.038***	0.024***
	(0.009)	(0.008)	(0.007)
% jobs by economic sector			
Agriculture	−0.360***	−0.942***	0.582***
	(0.137)	(0.168)	(0.111)
Industry	−0.014	−0.021	0.008
	(0.035)	(0.038)	(0.027)
Construction	−0.653***	−0.813***	0.160**
	(0.118)	(0.123)	(0.081)
Commerce, Transport, Services	0.107***	0.155***	−0.047***
	(0.020)	(0.019)	(0.014)
Public admin.	−0.191***	−0.194***	0.002
	(0.039)	(0.035)	(0.022)
Share unemployment	−0.086**	−0.090*	0.004
	(0.043)	(0.049)	(0.034)
Log(employment)	0.003*	0.008***	−0.005***
	(0.002)	(0.002)	(0.001)
Constant	0.188*	0.346***	−0.157**
	(0.103)	(0.101)	(0.074)
Year FE	Yes	Yes	Yes
N urban areas	151	151	151
N years	9	9	9
Observations	1,359	1,359	1,359
R^2	0.99766	0.99656	0.19278
Adjusted R^2	0.99763	0.99651	0.18315

Notes: The table presents the second step regression of the two-step approach as in equation (5) for workers above 30 years old. Coefficients in column (1) are reported with bootstrapped standard errors in parenthesis. *, ** and *** denote statistical significance at the 10, 5, and 1 % level.

Table A4.15.: Distribution of promotion indexes (period 2010-2018)

	D1	D2	D3	D4	D5	D6	D7	D8	D9	Mean	SD	Min.	Max.
	10%	20%	30%	40%	50%	60%	70%	80%	90%				
High-skilled	0.575	0.593	0.608	0.619	0.627	0.638	0.646	0.661	0.678	0.629	0.041	0.524	0.735
Low-skilled	0.703	0.711	0.719	0.724	0.728	0.732	0.737	0.743	0.749	0.726	0.02	0.673	0.788
Dif. HS−LS	-0.149	-0.127	-0.118	-0.109	-0.102	-0.093	-0.079	-0.065	-0.042	-0.098	0.042	-0.196	0.038

Notes: This table displays the distribution of promotion indexes for both high-skilled and low-skilled workers, along with the difference between the two groups. The promotion indexes are divided into deciles (D1 to D9) and provide insights into the variation in contract types across the sample. The sample includes 151 urban areas.

B. Computing housing price indexes

To obtain price indexes at the urban area level, we run the following hedonic regression model. This model is applied to a sample including all housing transactions that occurred between 2010 and 2018 within the 318 largest urban areas in France.

$$log(Y_{ijt}) = \alpha + \gamma X_{ijt} + \beta Dist_{ijt} + UA_{jt} + \varepsilon_{ijt} \tag{4.6}$$

The right-hand side variable Y_{ijt} corresponds to the sale price (*valeur foncière*) of a housing transaction i in a urban area j in year t. We control for the X_{ijt} characteristics of the dwelling/housing transaction that include: a dummy variable that equals 1 if it is a flat and 0 if it is a house, the construction period of the house or the flat (with the period 1961-1974 as reference), the ratio of the total surface in m^2 to the number of rooms, and quarter-time dummies (with quarter 2 as reference).

In addition, we incorporate the variable $Dist_{ijt}$ to control for the Euclidean distance between the centroid of the municipality where the sold dwelling is located and the centroid of the urban area (specifically, the centroid of the most populated municipality within that urban area). We include this variable under the assumption that greater distances from the urban area center are associated with reduced housing prices.

The key variable of interest in Equation 4.6 is UA_{jt}. This variable represents urban area-year fixed effects, and we use these as local-time housing price indexes. In their original levels, each urban area-year fixed effect corresponds to what would have been the constant of the model if that particular urban area $\times$ year was chosen as the reference point. In other words, an urban area $\times$ year fixed effect reflects the log(housing price) of a dwelling with reference characteristics, as specified in Equation 4.6, in that specific urban area and during a particular year. These urban area $\times$ year fixed effects thus provide us with local indexes that capture the fluctuations in housing prices across urban areas over time.

Table A4.16 presented below displays the regression results for Equation 4.6. As anticipated, we observe that the housing price of a dwelling tends to be higher when it is a house rather than a flat. Additionally, prices are positively influenced by factors such as the construction period, surface areas, and the number of rooms within a dwelling. On the other hand, the distance of a dwelling from the urban area center negatively affects its price, while housing transactions occurring in the 3rd and 4th quarters of the year are associated with higher prices.

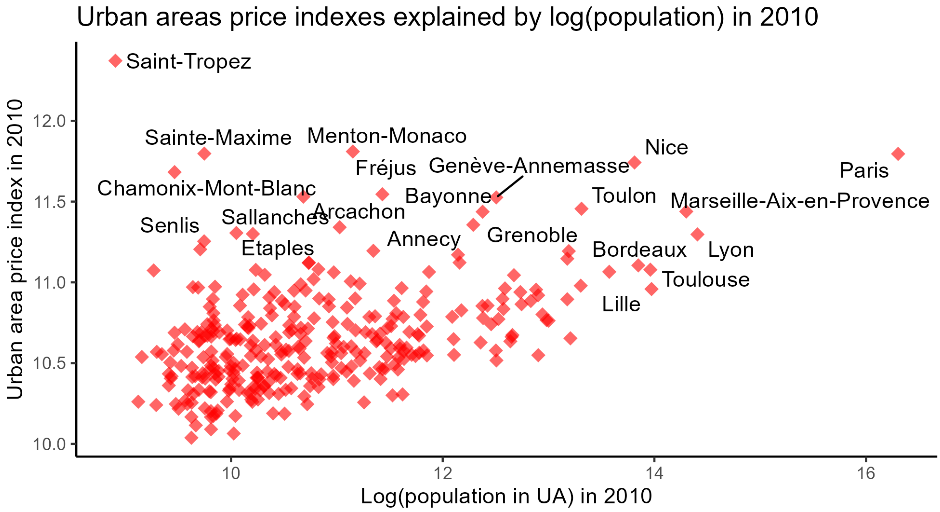
Table A4.16.: Hedonic price regression results

	log(sale price)	
	Coefficient	Std. Error
Flat (0/1) vs. house	-0.272***	(0.010)
Construction period		
< 1914	0.091***	(0.015)
1914-1944	0.055***	(0.004)
1945-1960	-0.008***	(0.003)
1961-1974	Ref.	Ref.
1975-1989	0.134***	(0.004)
1990-2012	0.238***	(0.005)
> 2013	0.262***	(0.008)
N° of rooms in the dwelling		
1 room	Ref.	Ref.
2 rooms	0.482***	(0.011)
3 rooms	0.790***	(0.012)
4 rooms	1.00***	(0.014)
More than 5 rooms	1.232***	(0.013)
Total surface (m^2)/N rooms	0.009***	(0.001)
Distance to the UA centre	-0.018***	(0.001)
Quarter time dummies		
Q1	-0.011***	(0.001)
Q2	Ref.	Ref.
Q3	0.027***	(0.001)
Q4	0.018***	(0.001)
UA x year FE	Yes	
Clustered SE	UA x year level	
Obs.	5,933,118	
N urban areas	318	
R^2	0.60	

Notes: The table presents regression results for a hedonic price model with the dependent variable being log(sale price). It includes various explanatory variables such as dwelling type, construction period, number of rooms, total surface area, distance to the urban area center, and quarter-time dummies. The table also accounts for urban area-year fixed effects and reports clustered standard errors at the urban area-year level.

Figure A4.6 illustrates the price indexes for the 318 largest urban areas in France in 2010, as computed using Equation 4.6. These indexes are plotted against the logarithm of their population in 2010. The purpose of these figures is to assess the meaningfulness of our price indexes. As expected, we can observe from this figure that housing price indexes generally exhibit a positive relationship with population size, with the rightmost point representing the urban area of Paris, where prices are notably higher. However, there are noteworthy exceptions to this trend. Housing prices are particularly elevated in tourist destinations like Saint Tropez and Sainte-Maxime, located in the upper-left corner of the graph, as well as in border urban areas such as Menton-Monaco and Genève-Annemasse.

Figure A4.6.: Relationship between urban areas' housing prices indexes in 2010 and log(population)



Notes: This figure depicts housing price indexes for 318 major urban areas in France in 2010, plotted against their 2010 population logarithm. The figure illustrates a positive association between housing prices and population.

C. Estimation sample

The sample used for our analysis is derived from the Panel d'actifs of the Echantillon Démographique Permanent (EDP), for which data is available until 2018. The construction of our initial sample involved several steps to obtain our estimation sample.

We first select data from the years 2010 to 2018, as the data for housing costs is only available from 2010. To clean our data, we eliminated individuals with missing information regarding their residential and workplace locations as well as observations with hourly wages considered outliers (below 0), which resulted in a more refined dataset. To focus our analysis on the working-age population, we excluded individuals below the age of 15 and above 60 in their first year of observation. We additionally excluded observations containing retired individuals engaged in specific economic activities, removed seasonal employment, and following Charruau and Epaulhard (2002) dropped observations with a yearly number of hours worked below 150, and below the hourly wage threshold of 6.5.

We also decided to keep three skill levels defined based on the INSEE socio-professional categorization: high-skilled (senior executives), low-skilled (blue-collar and white-collar), and intermediaries (intermediate occupations), as they represent the majority of the French labor force. We also choose to restrict our analysis to the private sector and therefore remove individuals from the public sector.[20]

[20]This choice was mainly driven by the categorization of individuals in the public sector in the Panel d'actifs of the EDP as individuals with "other types of contracts", therefore making difficult their categorization into long-term vs. short-term contracts.

Due to the instrumental variable that we use, we narrowed down the sample to include the 318 Urban Areas (UA) for which we have the inverse supply elasticities (see Chapelle et al. (2023)). We further refined our sample by removing urban areas with less than 30 high-skilled individuals per year. With a minimum threshold of 4 observations for each individual over the 2010-2018 period, this iterative process ultimately yielded a dataset comprising 3,724,905 observations with 463,785 individuals in the 151 largest French urban areas. These 151 urban areas represent 72% of the French population living in urban areas in 2010.

5. Concluding Remarks

This dissertation has explored two pivotal aspects of urban development, focusing on housing and tourism, and their implications for urban residents. As cities expand at an unprecedented pace, with an estimated 1.5 million new inhabitants added every week (United Nations, 2018), understanding these factors becomes increasingly crucial for ensuring the well-being of urban residents.

Chapter 2 delved into the influence of developers on local housing policies in California. Our research has brought to light the substantial influence of real estate developers on local housing policies. When developers are part of the city council, there's a noticeable increase in housing permits, suggesting a direct impact of their presence on housing supply. However, this effect, while significant, is not sufficient in the context of California's larger housing crisis.

Chapter 3 explored the long-term impact of tourism specialization on per capita income in Spanish municipalities. Our analysis uncovered a troubling trend: municipalities with the highest growth in tourism specialization now exhibit lower per capita income. This decline stems from several factors, including an increase in temporary job contracts, a decrease in industrial employment, and lower levels of educational attainment.

Chapter 4 examined the relationship between housing prices and labor market disparities among high-skilled and low-skilled workers in French cities. Our findings highlighted the exacerbating effect of rising housing costs on the inequality in job security between individuals with high and low skill levels. This suggests that housing costs can contribute to a widening gap in economic opportunities, further marginalizing low-skilled workers.

The research presented in this dissertation makes a substantial contribution to the field of urban economics. The first chapter, builds upon pivotal studies in urban economics and housing policy that have established the significant impact of land-use regulations on housing supply elasticity and affordability (Glaeser and Gyourko, 2003; Quigley and Rosenthal, 2005; Ihlanfeldt, 2007; Brueckner, 2009; Zabel and Dalton, 2011; Gyourko and Molloy, 2015; Jackson, 2018). These works consistently find that regulatory constraints can restrict land development, elevate construction costs, and limit housing diversity, thereby decreasing supply and elevating prices (Molloy, 2020). Distinguishing this study, it delves into the origins of these supply restrictions, offering a novel perspective on the development and enactment of urban policies. By examining the role of local politicians in shaping land use regulations and housing supply constraints, this research provides critical insights into the political and economic dynamics underpinning urban planning decisions. This approach not only enhances the understanding of regulatory impacts on housing markets but also contributes to a deeper comprehension of the intricate relationship between local governance, policy design, and urban economic outcomes.

The second chapter of this study, contributes to the existing body of literature on urban development and tourism. While there is no doubt that tourism has increasingly shaped the development of cities over the past decades, the literature exploring the comprehensive impact of city tourism on urban development is surprisingly limited. This study addresses this gap by delving into the long-term effects of tourism on urban areas.

The rapid expansion of tourism in cities has generated extensive debate regarding its economic consequences. Prior research, including studies by Garcia-López et al. (2020), Barron et al. (2021), and Franco and Santos (2021), has predominantly focused on its short-term impacts, particularly highlighting the negative repercussions on the housing market such as escalating housing prices and rents. Additionally, the influence of tourism on the well-being of residents and tourists, as explored by Farronato and Fradkin (2018), Almagro and Domınguez-Iino (2019), and Calder-Wang (2019), has been a subject of considerable interest. However, these studies primarily concentrate on short term effects, leaving a crucial aspect underexplored: the long-term impact of tourism on urban environments. This research fills this critical gap by being, to the best of my knowledge, the first study to examine the long-term consequences of tourism on the economic development. This approach not only broadens the understanding of tourism's role in urban dynamics but also offers valuable insights for future urban planning and policy-making in the context of growing tourism activities.

Chapter three of this thesis makes a noteworthy contribution to the urban economics literature by offering a detailed analysis of the interplay between housing and labor markets. The existing literature is extensive with studies exploring the interactions between jobs and job-seekers, particularly under the framework of the spatial mismatch hypothesis. This concept is explored in works by Brueckner and Zenou (2003), Patacchini and Zenou (2005), Gobillon et al. (2013), and Flemming (2020), among others. They focus on the link between employment and urban structure and identify adverse effects such increased commuting and job search costs, diminishing information about job opportunities, and a narrowed spatial range for job searches. Duranton and Puga (2004) highlight the positive effects of urban density in facilitating the match between economic agents, such as employers and employees. By building upon and synthesizing these concepts, chapter three enriches the understanding of how housing and labor markets interact and influence the disparities between high-skilled and low-skilled workers, especially in the context of urban environments.

To conclude, I want to summarize three main lessons that can be drawn from this dissertation. First, the role of real estate developers in shaping local housing policies cannot be overstated. Their influence is markedly evident in the amplification of housing supply, notably in the sector of multi-family residential units. This study underscores the necessity of policy interventions in the realm of zoning regulation and bureaucratic process optimization. Such interventions are pivotal, particularly in the context of facilitating housing projects targeting low and middle-income demographics. Optimizing these processes can significantly enhance the accessibility and affordability of housing, addressing a crucial need within urban population

Second, a prominent theme in this research is the disproportionate burden of housing costs borne by low-skilled workers. It is imperative for urban policy frameworks to integrate strategies that alleviate this burden. Constructing affordable housing in close proximity to employment centers emerges as a viable solution. Additionally, the implementation of housing subsidies for low-income workers and the promotion of mixed-use neighborhood developments can effectively diminish commuting expenses and times. These measures collectively serve to bridge the spatial disconnect between affordable housing locations and employment opportunities, fostering more equitable urban environments.

Third and finally, the interplay between tourism and urban economies, especially in tourist-heavy countries like Spain, is multifaceted. While tourism undeniably contributes to job creation, this study reveals its potential to precipitate lower per capita income, an increase in temporary employment contracts, and shifts in the municipal sectoral compositions. This could involve diversifying the local economy beyond tourism, investing in education and training to improve the skill sets of local workers, and implementing policies that mitigate the adverse effects of tourism, ensuring that the benefits of tourism are more evenly distributed among residents.

6. References

Allen, T., Fuchs, S., Ganapati, S., Graziano, A., Madera, R., and Montoriol-Garriga, J. (2021): “Urban Welfare: Tourism in Barcelona." Processed Dartmouth College.

Aladangady, A. (2017): “Housing wealth and consumption: Evidence from geographically linked microdata.” *American Economic Review*, 107(11), 3415–3446.

Albaladejo, I. P., González-Martínez, M. I., and Martínez-García, M. P. (2020): “A double life cycle in tourism arrivals to Spain: Unit root tests with gradual change analysis.” *Journal of Destination Marketing and Management* 18, 100497.

Albouy, D., Ehrlich, G., and Liu, Y. (2016): “Housing demand, cost-of-living inequality, and the affordability crisis.” *National Bureau of Economic Research 22816*, Technical report.

Allcott, H., and Keniston, D. (2018): “Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America.” *The Review of Economic Studies* 85(2), 695-731.

Almagro, M., and Domınguez-Iino, T. (2021): “Location sorting and endogenous amenities: Evidence from Amsterdam.” Processed Chicago Booth School of Business.

Andrews, I. , Gentzkow, M. , Shapiro, J.M. (2017): “Measuring the sensitivity of parameter estimates to estimation moments.” *The Quarterly Journal of Economics* 132 (4), 1553–1592.

Angelopoulos, K., Malley, J., and Philippopoulos, A. (2017): “Human capital accumulation and transition to skilled employment.” *Journal of Human Capital* 11(1), 72-105.

Anzia, S. (2011): “Election timing and the electoral influence of interest groups.” *Journal of Politics*, 73, 412-427.

Anzia, S. (2021): “Party and ideology in American local government.” *American Review of Political Science*, 24, 133-150.

Anzia, S. (2022): “Local interests: Politics, Policy, and Interest Groups in US City Governments.” U. of Chicago Press.

Arezki, R., Cherif, R. and Piotrowski, J., (2009): “Tourism specialization and economic development: Evidence from the UNESCO world heritage list.” International Monetary Fund.

Bailey, N. and Minton, J. (2018): “The suburbanisation of poverty in British cities, 2004-16: extent, processes and nature.” *Urban Geography*, 39(6), 892-915.

Balgova, M. (2022): “Moving to look or looking to move? The role of job search in migration decisions.” Processed IZA .

Barron, K., Kung, E., and Proserpio, D. (2021): “The effect of home-sharing on house prices and rents: Evidence from airbnb.” *Marketing Science,* 40(1):23–47.

Beach, B. and Jones, D. (2016): “Business as usual: Politicians with business experience, government finances, and policy outcomes.” *Journal of Economic Behavior and Organization*, 131, 292-307.

Becken, S., Whittlesea, E., Loehr, J., and Scott, D. (2020): “Tourism and climate change: evaluating the extent of policy integration.” *Journal of Sustainable Tourism* 28, 1603 - 1624.

Berry, C. R., and Glaeser, E. L. (2005): “The divergence of human capital levels across cities.” *Papers in Regional Science*, 84(3), 407-444.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005): “Why the apple doesn’t fall far: Understanding intergenerational transmission of human capital.” *American Economic Review* 95(1), 437-449.

Bonica, A. (2014): “Mapping the ideological marketplace.” *American Journal of Political Science*, 58(2), 367-386.

Bover, O., Muellbauer, J., and Murphy, A. (1988): “Housing, wages and UK labour markets (No. 268).” *CEPR Discussion Papers*.

Brida, J.G., Cortes-Jimenez, I. and Pulina, M. (2016): “Has the tourism-led growth hypothesis been validated? A literature review.” *Current Issues in Tourism*, 19(5), 394-430.

Bronzini, R., Ciani, E., and Montaruli, F. (2022): “Tourism and local growth in Italy.” *Regional Studies* 56(1), 140-154.

Broxterman, D.A and Yezer, A.M (2015): "Why does skill intensity vary across cities? The role of housing cost." *Regional Science and Urban Economics*, 55, 14–27.

Brueckner, J. K. (2009): *Government land use interventions: An economic analysis. In Urban land markets: Improving land management for successful urbanization* (pp. 3-23). Dordrecht: Springer Netherlands.

Brueckner, J. and Zenou, Y. (2003): "Space and unemployment: The labor market effects of spatial mismatch." *Journal of Labor Economics*, 21(1), 242–262.

Brueckner, J. K. and Martin, R. W. (1997): "Spatial mismatch: An equilibrium analysis." *Regional Science and Urban Economics*, 27(6), 693–714.

Brueckner, J. K., and Sridhar, K. S. (2012): "Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits." *Regional Science and Urban Economics*, 42, 1061–1067.

Brueckner, J., and Joo, M.S. (1991): "Voting with capitalization." *Regional Science and Urban Economics*, 21, 453-467.

Bughin, J., Manyka, J., and Woetzel, J. (2016): "A tool kit to close California's housing GAP: 3.5 million homes by 2025." McKinsey Global Institute.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016): " Agricultural productivity and structural transformation: Evidence from Brazil." *American Economic Review* 106(6), 1320-1365.

Calder-Wang, S. (2022): "The Distributional Impact of the Sharing Economy on the Housing Market," *American Economic Review,* forthcoming.

Calonico, S., Cattaneo, M. and Titiunik, R. (2014): "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal*, 144, 909-946.

Calonico, S., Cattaneo, M.D., and Farrell, M.H. (2019): "Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs." *Econometrics Journal*, 23(2), 192-210.

Calonico, S., Cattaneo, M.D., Farrell, M.H., and Titiunik, R. (2019): "Regression Discontinuity Designs Using Covariates." *Review of Economics and Statistics*, 101(3), 442-451.

Campbell, J. Y. and Cocco, J. F. (2007): "How do house prices affect consumption? evidence from micro data." *Journal of Monetary Economics*, 54(3), 591–621.

Capo, J., Font, A. R., and Nadal, J. R. (2007): "Dutch disease in tourism economies: Evidence from the Balearics and the Canary Islands." *Journal of sustainable Tourism* 15(6), 615-627.

Card, D. (2009): "Immigration and inequality." *American Economic Review* 99(2), 1-21.

Casella, A. and Palfrey, T. (2019): "Trading Votes for Votes. A Dynamic Theory." *Econometrica*, 87, 631-652.

Caselli, F., and Michaels, G. (2013): "Do oil windfalls improve living standards? Evidence from Brazil." *American Economic Journal: Applied Economics* 5(1), 208-38.

Cattaneo, M., Idrobo, N. and Titiunik, R. (2019): *A practical introduction to Regression Discontinuity Designs: Foundations.'* Cambridge University Press.

Cattaneo. M., Jansson, M. and Ma, X. (2018): "Manipulation testing based on density discontinuity." *Stata Journal*, 18(1), 234-261.

Causa, O., Woloszko, N., and Leite, D. (2019): "Housing, wealth accumulation and wealth distribution: Evidence and stylized facts." *OECD Economics Department Working Papers*, No. 1588.

Chakrabarti, R., and Zhang, J. (2015): "Unaffordable housing and local employment growth: Evidence from California municipalities." *Urban Studies*, 52(6), 1134-1151.

Chapelle, G., Eyméoud, J.B., and Wolf, C. (2023): "Land use regulation and the housing supply elasticity: evidence from France." Processed CY Cergy Paris Université.

Charruau, P., and Epaulard, A. (2022): "The Geography of Social Mobility During Working Life." Processed Jean Monnet University.

Chen, J. and Roth, J. (2023): "Log-like? Identified ATEs defined with zero-value outcomes are (arbitrarily) scale-dependent." Processed, Harvard U.

Combes, P. P., Duranton, G., and Gobillon, L. (2008): "Spatial wage disparities: Sorting matters!." *Journal of Urban Economics*, 63(2), 723-742.

Combes, P. P., Duranton, G., Gobillon, L., and Roux, S. (2012): "Sorting and local wage and skill distributions in France." *Regional Science and Urban Economics*, 42(6), 913-930.

Copeland B.R. (1991): “Tourism, Welfare and De-industrialization in a Small Open Economy.” *Economica* 58(232), 515-29.

Corden, M. (1984): “Booming sector and Dutch disease economies: Survey and consolidation.” *Oxford Economic Papers* 36, 359–380.

Corden, M. and Neary, P. (1982): “Booming sector and de-industrialization in a small open economy.” *The Economic Journal* 92, 825–848.

Cortes, P. (2008): “The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data.” *Journal of Political Economy* 116(3), 381-422.

Coulson, N., Laing, D., and Wang, P. (2001): “Spatial mismatch in search equilibrium.” *Journal of Labor Economics*, 19(4), 949–972.

Croes, R., and Semrad, K. (2018): *Destination competitiveness.”* The SAGE Handbook of Tourism Management: Applications of theories and concepts to tourism 77-90.

D’Costa, S., and Overman, H. G. (2014): “The urban wage growth premium: Sorting or learning?.” *Regional Science and Urban Economics*, 48, 168-179.

De Benedictis-Kessner, J. and Warshaw, Ch. (2016): “Mayoral partisanship and municipal fiscal policy.” *Journal of Politics*, 78(4), 1124-1138.

De Benedictis-Kessner, J., Jones, D. and Warshaw, Ch. (2023): “How partisanship in cities influences housing policy.” *American Journal of Political Science*, forthcoming.

Deller, S. C., Lledo, V., and Marcouiller, D. W. (2008): *Modeling regional economic growth with a focus on amenities.* In Review of Urban and Regional Development Studies: Journal of the Applied Regional Science Conference 20(1), 1-21. Melbourne, Australia: Blackwell Publishing Asia.

Diamond, R. (2016): “The determinants and welfare implications of us workers’ diverging location choices by skill: 1980-2000.” *American Economic Review*, 106, 479–524.

Diamond, R., and Moretti, E. (2021): “Where is standard of living the highest? Local prices and the geography of consumption.” *National Bureau of Economic Research 29533*.

Dogru, T., McGinley, S., Line, N., and Szende, P. (2019): “Employee earnings growth in the leisure and hospitality industry.” *Tourism Management* 74, 1-11.

Dong, Y. (2019): "Regression discontinuity designs with sample selection." *Journal of Business and Economic Statistics*, 37(1), 171-185.

Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014): "Spatial sorting." *Journal of Political Economy*, 122(3), 554-620.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and Snyder, J. M. (2015): "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races." *American Journal of Political Science*, 59(1), 259-274.

Enders, W. and Sandler, T., 1991: "Causality Between Transnational Terrorism and Tourism: The case of Spain." *Studies in Conflict and Terrorism*, 14(1), 49-58.

Enders, W., Sandler, T. and Parise, G.F., 1992: "An Econometric Analysis of the Impact of Terrorism on Tourism." *Kyklos*, 45(4), 531-554.

Faber, B., and Gaubert, C. (2019): "Tourism and economic development: Evidence from Mexico's coastline." *American Economic Review* 109(6), 2245-93.

Farré, L., Jofre-Monseny, J., and Torrecillas, J. (2023): "Commuting time and the gender gap in labor market participation." *Journal of Economic Geography*, 23(4), 847-870.

Farronato, C. and Fradkin, A. (2018): "The welfare effects of peer entry in the accommodation market: The case of airbnb." *National Bureau of Economic Research.*

Ferreira, F., Gyourko, J., and Tracy, J. (2010): "Housing busts and household mobility." *Journal of Urban Economics*, 68(1), 34-45.

Fischel, W. A. (2001): *The Homevoter Hypothesis*. Harvard U. Press, Cambridge, MA.

Flemming, J. (2020): "Costly commuting and the job ladder." *Finance and Economics Discussion Series 2020-025*, Board of Governors of the Federal Reserve System (U.S.).

Folke, O. (2014): "Shades of brown and green: party effects in proportional election systems." *Journal of the European Economic Association*, 12(5), 1361-1395.

Fort, T. C., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013): "How firms respond to business cycles: The role of firm age and firm size." *IMF Economic Review*, 61(3), 520-559.

Franco, S. F. and Santos, C. D. (2021): “The impact of airbnb on residential property values and rents: Evidence from portugal.” *Regional Science and Urban Economics,* 88:103667.

Fujita, M. (1989): *Urban economic theory.* Cambridge Books.

Garcia-López, M. À., Jofre-Monseny, J., Martínez-Mazza, R., and Segú, M. (2020): “Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona.” *Journal of Urban Economics,* 119, 103278.

Gehlbach, S., Sonin, K., and Zhuravskaya, E. (2010): “Businessman candidates.” *American Journal of Political Science*, 54(3), 718-736.

Gerber, E.R., and Hopkins, D.J. (2011): When Mayors Matter: Estimating the Impact of Mayoral Partisanship on City Policy. *American Journal of Political Science*, 55(2), 326-339.

Glaeser, E. L. (2011): *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier,* Macmillan.

Glaeser, Edward L. and Joseph Gyourko. 2003: “The Impact of Building Restrictions on Housing Affordability.” *FRBNY Economic Policy Review* 9:21-39.

Glaeser, E. L., and Gyourko, J. (2005): “Urban decline and durable housing.” *Journal of Political Economy*, 113(2), 345-375.

Glaeser, E. L., Gyourko, J. and Saks, R.E. (2005): “Why have housing prices gone up?” *American Economic Review*, 952, 329–333.

Glaeser, E.L. and Gyourko, J. (2018): “The economic implications of housing supply.” *Journal of Economic Perspectives*, 32(1), 3-30.

Glaeser, E.L., Gyourko, J., and Saks, R.E. (2006): “Urban growth and housing supply.” *Journal of Economic Geography*, 6(1), 71-89.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018): “Bartik Instruments: What, When, Why, and How.” *American Economic Review* 110(8), 2586-2624.

Gollin, D., Jedwab, R., and Vollrath, D. (2016): “Urbanization with and without industrialization.” *Journal of Economic Growth* 21(1), 35-70.

González, L., and Surovtseva, T. (2020): “Do more tourists promote local employment?" Processed Graduate School of Economics.

Gössling, S., and Peeters, P. M. (2015): "Assessing tourism's global environmental impact 1900–2050." *Journal of Sustainable Tourism*, 23(5), 639-659.

Güell, M. and Petrongolo, B. (2007): "How binding are legal limits? Transitions from temporary to permanent work in Spain." *Labor Economics* 14(2), pp. 153–83.

Guren, A. M., McKay, A., Nakamura, E., and Steinsson, J. (2021): "Housing wealth effects: The long view." *The Review of Economic Studies*, 88, 669–707.

Gylfason, T. (2001a): "Natural resources, education, and economic development." *European Economic Review* 45(4-6), 847-859.

Gylfason, T. (2001b): "Nature, power and growth." *Scottish Journal of Political Economy* 48(5), 558-588.

Gyourko, J., and Molloy, R. (2015): "Regulation and housing supply." *Handbook of Regional and Urban Economics*, Vol. 5. Elsevier, 1289-1337.

Gyourko, J., Mayer, Ch., and Sinai, T. (2013): "Superstar cities." *American Economic Journal: Economic Policy*, 54, 167-99.

Hainmueller, J. (2012): "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20(1), 25–46.

Hankinson, M. (2018): "When do renters behave like homeowners? High Rent, Price Anxiety, and NIMBYsm." *American Political Science Review*, 66(5), 743-753.

Hankinson, M. and Magazinnik, A. (2023): "The supply-equity trade-off: the effect of spatial representation on the local housing supply." *The Journal of Politics*, Forthcoming.

Helble, M., Ok Lee, K., and Gia Arbo, M. A. (2021): "How (Un) affordable is housing in developing Asia?" *International Journal of Urban Sciences*, 25, 80–110.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018): "Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown." *Journal of Monetary Economics*, 93, 89–109.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014): "Growth and structural transformation." *Handbook of Economic Growth* 2, 855-941.

Hilber, Ch. and Robert-Nicoud, F. (2013): "On the origins of land use regulations: Theory and evidence from US metro areas." *Journal of Urban Economics*, 75, 29-43.

Hjort, J., and Poulsen, J. (2019): "The arrival of fast internet and employment in Africa." *American Economic Review* 109(3), 1032-79.

Holzner M. (2011): "Tourism and economic development: The beach disease?" *Tourism Management* 32 (4), 922-933.

Hornbeck, R., and Moretti, E. (2022): "Estimating who benefits from productivity growth: local and distant effects of city productivity growth on wages, rents, and inequality." *Review of Economics and Statistics*, Forthcoming.

Hoyt, J. and Schuetz, J. (2020): Making apartments more affordable starts with understanding the costs of building them. Report. The Brookings Institution.

Hsieh, C.-T. and Moretti, E. (2019): "Housing constraints and spatial misallocation." *American Economic Journal: Macroeconomics*, 11, 1–39.

Ihlanfeldt, Keith R. 2007: "The Effect of Land Use Regulation on Housing and Land Prices." *Journal of Urban Economics* 61: 420-435.

Imai, K. and Khanna, K. (2016): "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records." *Political Analysis*, 24, 263-272.

Inchausti-Sintes, F., Voltes-Dorta, A.,and Suau-Sánchez, P. (2021): "The income elasticity gap and its implications for economic growth and tourism development: the Balearic vs the Canary Islands." *Current Issues in Tourism* 24(1), 98-116.

Jackson, K.K. (2018): "Regulation, land constraints, and California's boom and bust." *Regional Science and Urban Economics*, 68, 130-147.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018): "Shift-share instruments and the impact of immigration." National Bureau of Economic Research 24285.

Johnes, G., and Hyclak, T. (1994): "House prices, migration, and regional labor markets." *Journal of Housing Economics*, 3(4), 312-329.

Johnes, G., and Hyclak, T. (1999): "House prices and regional labor markets." *The Annals of Regional Science*, 33, 33-49.

Kahn, M. (2011): "Do liberal cities limit new housing development? Evidence from California." *Journal of Urban Economics*, 692, 223-228.

Kim, D., Liu, F., and Yezer, A. "Do inter-city differences in intra-city wage differentials have any interesting implications?." *Journal of Urban Economics*, 66, 203–209.

Kirkland, P.A. (2021): "Business Owners and Executives as Politicians: The Effect on Public Policy." *The Journal of Politics*, 83(4), 1652-1668.

Koster, H. R., and Ozgen, C. (2021): "Cities and tasks." *Journal of Urban Economics*, 126, 103386.

Kuznets, S. (1973): "Modern economic growth: findings and reflections." *The American economic review,* 63(3), 247-258.

Lanzara, G., and Minerva, G. A. (2019). "Tourism, amenities, and welfare in an urban setting", *Journal of Regional Science*, 59(3), 452-479.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021): "Gender differences in job search: Trading off commute against wage." *The Quarterly Journal of Economics*, 136(1), 381-426.

Lee, D.S., and Lemieux, Th. (2010): "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2), 281-355.

Lee, D.S., Moretti, E., and Butler, M.J. (2004): "Do Voters Affect or Elect Policies? Evidence from the U.S. House." *The Quarterly Journal of Economics*, 119(3), 807-859.

Leffers, P. (2017): "Real Estate developers influence on land use legislation in the Toronto region." *Urban Studies*, 55/14, 3059-3075.

Makarski, K., Tyrowicz, J., and Lutynski, J. (2022): "Structural change and inequality in general equilibrium." GRAPE Working Papers 67.

Manning, A., and Petrongolo, B. (2017): "How local are labor markets? Evidence from a spatial job search model." *American Economic Review*, 107(10), 2877-2907.

Marble, W. and Nall, C. (2021): "Where interests trump ideology: the persistent influence of homeownership in local development politics." *Journal of Politics*, forthcoming.

Mason, K., Morrison, W. G., and Stockman, I. (2016): "Liberalization of air transport in Europe and the evolution of 'low-cost' airlines." *Liberalization in Aviation* 163-177.

Mast, E. (2020): "Warding off development: Local control, housing supply, and NIMBYs." *Review of Economics and Statistics*, forthcoming.

McMillan, M., Rodrik, D., Bacchetta, M., and Jensen, M. (2011): "Making globalization socially sustainable." International Labour Office.

Mian, A., and Sufi, A. (2011): "House prices, home equity–based borrowing, and the us household leverage crisis." *American Economic Review*, 101(5), 2132-2156.

Mian, A., Rao, K., and Sufi, A. (2013): "Household balance sheets, consumption, and the economic slump." *The Quarterly Journal of Economics*, 128(4).

Mion, G. and Naticchioni, P. (2009): "The spatial sorting and matching of skills and firms." *Canadian Journal of Economics*, 42(1):28-55.

Molotch, H. (1976): "The city as a growth machine: towards a political economy of place." *American Journal of Sociology*, 82(2), 309-332.

Molloy, R. (2020): "The effect of housing supply regulation on housing affordability: A review." *Regional science and urban economics*, 80(C), 1-5.

Moretti, E. (2004): "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data." *Journal of Econometrics*, 121, 175–212.

Moretti, E. (2013): "Real wage inequality." *American Economic Journal: Applied Economics*, 5, 65–103.

Mortensen, D. T., and Pissarides, C. A. (1999): "Chapter 39 New developments in models of search in the labor market." In *Handbook of Labor Economics* (Vol. 3, pp. 2567-2627). Elsevier.

Muñoz-Bullón, F. (2009): "The gap between male and female pay in the Spanish tourism industry." *Tourism Management* 30(5), 638-649.

NAHB (2017): Building a successful local advocacy program. National Association of Home Builders. [Online] Available at: https://www.nahb.org/en/advocate/-/media/8F3269649F134D44821CD579

NAR (2020): "How and why to get members elected." REALTOR AE Magazine. [Online] Available at: https://www.nar.realtor/realtor-ae-magazine/how-and-why-to-get-members-elected

Neumayer, E., 2004: "The impact of political violence on tourism: Dynamic cross-national estimation." *Journal of Conflict Resolution* 48(2), 259-281.

Nocito, S., Sartarelli, M., and Sobbrio, F. (2021): “A Beam of Light: Media, Tourism and Economic Development.” *Tourism and Economic Development.*

OECD Housing prices (indicator) (2023), 10.1787/63008438-en

Oi, W. (1976): “Residential location and labor supply.” *Journal of Political Economy*, 84(4-2), 221-238.

Oliver, X., and Villalonga, J. R. (2019): “Does the economic productive structure determine early school leaving rates? Evidence from Spanish regions, 2000-2013.” *Estudios sobre Educación,* 36, 181-205.

Ong, R., and Leishman, C. (2020): “Home truths: the role of housing in economic recovery.” CEDA Working paper.

Pardo, Á. (2004): “Un modelo espacial de renta per capita regional: evidencias provincial, comarcal y municipal.” *Investigaciones Regionales - Journal of Regional Research* 2004, 11761-11766.

Parellada, M. (1992): “Una aproximación a la estimación de la renta municipal en España.” *Economía española, cultura y sociedad. Homenaje a Juan Velarde Fuentes, tomo II, Eudema, Madrid.*

Parkhomenko, A. (2020): “Local causes and aggregate implications of land use regulation.” *Journal of Urban Economics,* 138, 103605.

Parrilla, J.C., Font, A.R. and Nadal, J.R. (2007): “Tourism and long-term growth a Spanish perspective.” *Annals of Tourism Research* 34(3), 709-726

Patacchini, E., and Zenou, Y. (2005): “Spatial mismatch, transport mode and search decisions in England.” *Journal of Urban Economics*, 58(1), 62–90.

Pei, Z., Card, D. and Weber, A. (2020): “Local Polynomial Order in Regression Discontinuity Designs.” *Journal of Business and Economic Statistics,* 40(3), 1259-1267.

Petrongolo, B., and Ronchi, M. (2020): *Gender gaps and the structure of local labor markets.* Labour Economics, 64, 101819.

Piketty, T. (2018): “Capital in the twenty-first century.” *Harvard University Press*.

Quigley, J. and Raphael, S. (2005): “Regulation and the High Cost of Housing in California.” *American Economic Review*, 95 (2), 323-328.

Quigley, John M. and Larry A. Rosenthal. 2005: "The Effects of Land Use Regulation on the Price of Housing: What Do We Know? What Can We Learn?" *Cityscape* 8(1): 69-137.

Quintero, L. (2020): "Fewer players, fewer homes: concentration and the new dynamics of housing supply." Processed, John Hopkins University.

Riley, M., Ladkin, A., and Szivas, E. (2002): "Tourism employment: Analysis and planning." *Channel View Publications* 6, 10.

Roca, J. D. L., and Puga, D. (2017): "Learning by working in big cities." *The Review of Economic Studies*, 84(1), 106-142.

Rotemberg, J.J. (1983): "Instrumental variable estimation of misspecified models." Working Paper 1508-83, MIT Sloan.

Sá, F. (2015): "Immigration and House Prices in the UK." *The Economic Journal* 125, 1393-1492.

Saiz, A. (2010): "The geographic determinants of housing supply." *The Quarterly Journal of Economics*, 125 (3), 1253-1296.

Saks, R. E. (2008): "Job creation and housing construction: Constraints on metropolitan area employment growth." *Journal of Urban Economics*, 64(1), 178-195.

Sánchez-Sánchez, E. M. (2001): "El auge del turismo europeo en la España de los años sesenta." *Arbor* 170(669), 201-224.

Santos, L. D., and Varejão, J. (2007): "Employment, pay and discrimination in the tourism industry." *Tourism Economics* 13(2), 225-240.

Schmutz, B., and Sidibé, M. (2019): "Frictional labour mobility." *The Review of Economic Studies*, 86(4), 1779-1826.

Schmutz, B., Sidibé, M., and Vidal-Naquet, E. (2021): "Why Are Low-Skilled Workers Less Mobile? The Role of Mobility Costs and Spatial Frictions." *Annals of Economics and Statistics*, 142, 283-304.

Sequeira, T.N. and Maçãs Nunes, P. (2008). "Does tourism influence economic growth? A dynamic panel data approach." *Applied Economics* 40(18), 2431.

Shetzer, A., Twinam, T. and Walsh, R. (2016): "Race, ethnicity, and discriminatory zoning." *American Economic Journal: Applied Economics*, 8(3), 217-246.

Solé-Ollé, A. and Viladecans-Marsal, E. (2012): "Lobbying, political competition, and local land supply: recent evidence from Spain." *Journal of Public Economics*, 961, 10-19.

Solé-Ollé, A. and Viladecans-Marsal, E. (2013): "Do political parties matter for local land use policies?" *Journal of Urban Economics*, 78, 42-56.

Song H., Dwyer L. and Zhengcao G.L. (2012): "Tourism economics research: a Review and assessment" *Annals of Tourism Research* 39 (3), 1653–1682.

Szakonyi, D. (2021): "Private sector policymaking: business background and politicians' behavior in office." *Journal of Politics*, 83(1).

Torvik, R. (2001): "Learning by doing and the Dutch disease." *European Economic Review* 45(2), 285-306.

Trounstine, J. (2018): "The Geography of Inequality: How Land Use Regulation Produces Segregation and Polarization." *American Political Science Review*, 114 (2), 443-455.

United Nations (2018): "World urbanization prospects: The 2018 revision."

Wetzstein, S. (2017): "The global urban housing affordability crisis." *Urban Studies*, 54, 3159–3177.

World Travel and Tourism Council (WTTC) (2021): "Economic Impacts Report." WTTC, Brussels.

Yu, R. (2022): "Returns to Political Contributions in Local Housing Markets." Processed, University of Pennsylvania.

Zabel, J. E. (2012): "Migration, housing market, and labor market responses to employment shocks." *Journal of Urban Economics*, 72(2–3), 267-284.

Zabel, Jeffrey and Maurice Dalton. 2011: "The Impact of Minimum Lot Size Regulations on House Prices in Eastern Massachusetts." *Regional Science and Urban Economics* 41: 571-583.

Zenou, Y. (2009): "Urban labor economics." Cambridge University Press.