NBER WORKING PAPER SERIES

# TOURISM AND ECONOMIC DEVELOPMENT: EVIDENCE FROM MEXICO'S COASTLINE

Benjamin Faber
Cecile Gaubert

Working Paper 22300
http://www.nber.org/papers/w22300

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2016

We are grateful to David Atkin, Arnaud Costinot, Dave Donaldson, Pablo Fajgelbaum, Gordon Hanson, Pete Klenow, Marc Muendler, Natalia Ramondo, Andrés Rodríguez-Clare and participants at multiple conferences and seminars for helpful comments. Nicholas Li and Jose Vasquez-Carvajal provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Tourism and Economic Development: Evidence from Mexico's Coastline
Benjamin Faber and Cecile Gaubert
NBER Working Paper No. 22300
June 2016
JEL No. F15,F63,O24

**ABSTRACT**

Tourism is one of the most visible and fastest growing facets of globalization in developing countries. This paper combines a rich collection of Mexican microdata with a quantitative spatial equilibrium model and a new empirical strategy to learn about the long-run economic consequences of tourism. We begin by estimating a number of reduced-form effects on local economic outcomes in today's cross-section of Mexican municipalities. To base these estimates on plausibly exogenous variation in long-term tourism exposure, we exploit geological and oceanographic variation in beach quality along the Mexican coastline to construct instrumental variables. To guide the estimation of the aggregate implications of tourism, we then write down a spatial equilibrium model of trade in goods and tourism services, and use the reduced-form moments to inform its calibration for counterfactual analysis. We find that tourism causes large and significant local economic gains relative to less touristic regions, and that these gains are in part driven by significant positive spillovers on manufacturing production. In the aggregate, however, we find that these local spillovers are largely offset by reductions in agglomeration economies among less touristic regions, so that the national gains from tourism are mainly driven by a classical market integration effect.

Benjamin Faber
Department of Economics
University of California, Berkeley
697A Evans Hall
Berkeley, CA 94720
and NBER
benfaber@econ.berkeley.edu

Cecile Gaubert
Department of Economics
University of California, Berkeley
687 Evans Hall
Berkeley, CA 94720
and NBER
cecile.gaubert@berkeley.edu

# 1 Introduction

Tourism is a particular form of market integration: instead of shipping goods across space, tourism involves the export of otherwise non-traded local services and amenities by temporarily moving consumers across space. Tourist expenditures on these local services are then reported as tourism exports in cross-country data on services trade flows. Over recent decades, these tourism exports have grown to become a quantitatively important channel of global integration, and this is particularly the case for developing countries.[1]

Unsurprisingly in this context, tourism has attracted widespread policy attention. Virtually every country in the world has one or several publicly funded tourism promotion agencies. Some governments and international organizations have also been advocating the promotion of tourism to foster local economic development in economically lagging regions within countries.[2] At the same time, much of the existing social sciences literature on tourism has been critical about its long-term economic consequences, especially in developing countries.[3] For example, Honey (1999) and Dieke (2000) have questioned the extent to which the gains from tourism accrue to the local population, rather than being captured by multinationals or domestic elites. In economics, the existing literature has argued that tourism may give rise to a particular form of the "Dutch disease" by reallocating factors of production towards stagnant services activities and away from traded sectors with higher potential for productivity growth (Copeland, 1991).

Despite the rapid growth of tourism and widespread policy attention, the existing literature on trade and development has so far paid relatively little attention to this channel of market integration. This paper seeks to fill this gap using the empirical context of Mexico, a country where tourism has grown to become an important economic force since the 1950s. Our analysis aims to contribute to our understanding of two central questions: i) what are the long-run economic consequences of tourism in a developing country, and ii) what are the channels underlying these effects. Since the development of tourism in Mexico is driven by both international and domestic tourism flows, we set out to study the consequences of both cross-border and inter-regional tourism integration, and decompose the gains from tourism into an international and domestic component.

In answering these questions, the paper also makes a contribution to the growing empirical literature that exploits within-country variation to estimate the effects of policy shocks across regions (e.g. Autor et al. (2013), Mian & Sufi (2009), Topalova (2010)). This literature has focused on credibly identifying the effect of a shock on relative regional economic outcomes within a country. While certainly of interest in its own right, this approach generally does not allow to shed light on the corresponding aggregate implications, as those are being soaked up by the constant term in

[1]World tourism exports were USD 1.25 trillion in 2014, making it the single largest sector of global trade in services (accounting for one quarter). Tourism exports of low and middle income countries have grown at an average annual rate of 11 percent over the period 1982-2012. For this group of countries, tourism exports over the past decade have been of the same magnitude as 75 percent of all food and agriculture exports combined. Figures are based on UNCTAD statistics (see http://unctad.org/en/pages/Statistics.aspx).

[2]See for example "*Passport to Development*" (WorldBank, 1979) or "*Tourism and Poverty Alleviation: Untapped Potential*" (DFID, 1999).

[3]See Hawkins & Mann (2007) for a review of this literature.

regressions on a cross-section of regions or by time fixed effects when using regional panel data. This shortcoming is particularly acute when the objective is to estimate the long-run effects of policy shocks, because over time workers are mobile to arbitrage away regional variation in real incomes.[4]

To make progress on this trade-off, this paper combines a reduced-form analysis that exploits within-country variation with a quantitative spatial equilibrium model. We first use a rich collection of Mexican microdata in combination with an instrumental variable (IV) strategy to estimate the effects of tourism on a number of *local* economic outcomes in today's cross-section of Mexican municipalities. We then use these moments to inform the calibration of a spatial equilibrium model that allows us to interpret the relative regional effects, decompose them into a number of underlying channels, and explore the *aggregate* implications of both domestic and international tourism integration.

The model explicitly allows for the possibility that tourism integration may lead to adverse long-run effects by introducing different sources of local production externalities. By altering the scale of production in different sectors both locally and in the aggregate, the development of tourism can affect the productivity of traded goods production in the economy. If agglomeration economies mainly operate within the manufacturing sector, the aggregate gains from tourism can be diminished or over-turned compared to the conventional neoclassical gains from falling frictions to tourism trade (e.g. travel costs). On the other hand, if agglomeration economies also operate at the cross-sector level, then the development of the services sector due to tourism can generate spillovers on manufacturing production that reinforce the classical gains from integration.

At the center of the analysis lies the construction of a rich collection of microdata. We assemble a database containing: i) municipality-level hotel revenues, employment, population, wages and output by sector from the Mexican Censos Economicos in 1998 and 2008 and the Mexican population censuses in 2000 and 2010; ii) a long time series of population census data for consistent spatial units going back to 1921; iii) remote sensing satellite data at a resolution of 30x30 meter pixels covering roughly 9,500 km of Mexican coastline across six different bands of wavelength during the 1980s and 90s; and iv) panel data on bilateral tourism exports and relative prices covering 115 countries over the period 1990-2011.

Armed with this database, the analysis proceeds in two parts. In the first part, we estimate reduced-form effects of tourism on current-day municipality-level population, employment, Mincerized wages and local GDP by sector of activity. Our empirical strategy takes inspiration from the tourism management literature arguing that variation in tourism activity is to a large extent determined by the quality of a very specific set of local natural characteristics (e.g. Weaver et al. (2000), Leatherman (1997)). We identify a set of beach quality criteria from that literature, such as the presence of a nearby offshore island or the fraction of onshore coastline covered by white sand beaches, that we can capture using the satellite data to construct instrumental variables for tourism attractiveness along the Mexican coastline. Our empirical strategy exploits cross-sectional varia-

[4] This limitation and the need for a more structured approach to get at general equilibrium effects has been highlighted by, for example, Kline & Moretti (2014) and Donaldson & Hornbeck (2015).

tion to capture the long-term effects of tourism on relative regional economic outcomes, leveraging the fact that beach tourism has had more than half a century in Mexico to materialize into today's observed distribution of regional outcomes.

The identifying assumption is that islands or the fraction of coastline covered by picturesque sand beaches do not affect local economic outcomes in today's cross-section of Mexican municipalities relative to other coastal municipalities, except through their effect on tourism activity. We assess the validity of this assumption in several ways. We report how point estimates are affected by the inclusion of pre-determined municipality controls, and assess whether six different instrumental variables yield similar estimates. To further assess the extent to which the IVs may impact the local economy by directly affecting the amenities of local residents or by being correlated with natural advantages in other sectors, we conduct a placebo falsification test. We construct a long time series of population census data and estimate the effect of the IVs on municipality populations in periods before beach tourism became a discernible force in Mexico. Finally, we verify the extent to which our instruments are correlated with model-based estimates of local amenities of residents in today's cross-section of municipalities.

Using this design, we find that variation in local tourism activity has strong and significant positive effects on municipality total employment, population, local GDP and wages relative to less touristic regions. According to our preferred specification, a 10 percent increase in local tourism revenues leads to a 2.8 percent increase in municipality total employment, and a 4.3 percent increase in nominal municipality GDP in today's cross-section of Mexican municipalities. These effects appear to be driven by sizable local multiplier effects on traded sector production. We find that a 10 percent increase in local tourism revenues leads to a 3.2 percent increase in local manufacturing GDP. This effect holds conditional on municipality differences in access to transport infrastructure, as well as among manufacturing sectors that are not intensively used as inputs in the production of tourism-related services.

The positive effect of tourism on local manufacturing production, while seemingly at odds with the notion that tourism acts as a Dutch disease, does not by itself provide prima facie evidence for positive spillovers from tourism development onto manufacturing. In a world with trade costs, labor mobility and input-output linkages, the net effect of tourism on local manufacturing is a priori ambiguous and could be positive through neoclassical channels alone. On the one hand, the development of tourism increases local labor demand and hence wages which, ceteris paribus, tends to hurt local manufacturing. On the other hand, however, tourism can have a positive impact on local manufacturing due to increased local market access that is caused by both in-migration and input demand from tourism-related activities.

To account for these channels in our assessment of the long-term economic impact of tourism, we next investigate them quantitatively through the lens of a spatial equilibrium model. This second part of the analysis allows us to interpret and decompose the reduced-form effects of tourism development, and to explore the aggregate implications that are consistent with the observed local effects. We build on the theoretical framework developed by Allen & Arkolakis (2014), Ahlfeldt et al. (2015) and Redding (2015), and extend it to capture the economic forces that are relevant in our context. In addition to trade in goods and migration across regions, the model features trade in tourism-related services across regions and countries, input-output linkages between tourism and manufacturing, as well as local production externalities.

We allow for manufacturing production to be subject to both within and cross-sector spillovers. The within-sector spillover is the standard source of agglomeration economies in economic geography models, and captures the extent to which a larger scale in local manufacturing production is beneficial for manufacturing productivity. In its presence, reducing the scale of manufacturing as the economy re-allocates factors towards services leads to adverse productivity effects in the aggregate. This adverse effect works in the opposite direction of the neoclassical gains from falling frictions to tourism trade. On the other hand, the cross-sector spillover captures the extent to which a larger scale of the local services sector affects traded sector productivity. By increasing local services production, the development of tourism may generate long-run positive spillovers on traded goods production by, for example, improving access to business services for local firms, such as finance, accounting or consulting, by loosening local credit constraints directly (through tourism revenues), or by facilitating contacts and business networks. In the presence of such cross-sectoral agglomeration economies, tourism can give rise to gains in traded-sector productivity that would not have occurred otherwise, thus reinforcing the neo-classical gains from tourism integration.

To quantify these forces, we estimate the model parameters, and calibrate the model to current-day Mexico as a reference equilibrium. In particular, we estimate the intensity of the within and cross-sector spillovers using an approach that combines model-based indirect inference with the exclusion restrictions of our instrumental variables. We find that both within and cross-sector agglomeration economies are necessary to rationalize the observed reduced-form effects of tourism on Mexican regions, while accounting for purely neoclassical linkages between tourism and local traded goods production. In addition to the conventional within-manufacturing agglomeration economies, we find that tourism –through its effect on the development of the local services sector– leads to positive spillovers on local traded goods production.

Armed with the model parameters and a number of observed moments in our data, we proceed to explore general equilibrium counterfactuals. To quantify the welfare gains from tourism, we solve for the welfare implications of moving from the current levels of domestic and international tourism to a prohibitive level of tourism travel frictions. We find that tourism causes significant long-run gains to the average Mexican household that are in the order of 4.4 percent of household consumption. Slightly more than one third of these gains are driven by international tourism, and the remainder by domestic tourism across Mexican regions.

Turning to the underlying channels, we find that slightly more than half of the observed effect on local GDP can be explained by neoclassical forces, including the direct mechanical effect due to local tourism expenditures. The remainder is driven by gains in local manufacturing activity due to both cross and within-sector agglomeration forces. In the aggregate, however, we find that these spillover effects contribute relatively little to the estimated welfare gains. That is, while the presence of within and cross-sector spillovers reinforce one another leading to the large observed re-allocations of economic activity towards touristic regions, we find that they largely offset one another at the aggregate level, so that the aggregate gains from tourism are mainly driven by a classical market integration effect.

Finally, an interesting difference to these findings emerges when we focus on the gains from international-only tourism. In this case, we find that the gains from tourism integration are dampened compared to what they would have been in the absence of agglomeration forces. In regions relatively more affected by international tourism, the reduction in the within-manufacturing scale effect outweighs the gains from the local expansion in services. We find that this result is driven by differences in the sectoral composition of the regions most affected by international tourism compared to domestic tourism. As a result, the gains from opening up to international tourism in the absence of agglomeration economies would be larger (2.4 percent) compared to the gains that we estimate (1.6 percent).

This paper relates and contributes to the recent literature on trade and development (e.g. Topalova (2010), Donaldson (in press), Atkin et al. (2015)). Relative to the existing literature, we focus on tourism, an important and fast-growing but so far understudied facet of globalization in developing countries. There is a small existing empirical literature that has analyzed cross-country data to shed light on the determinants and consequences of tourism.[5] In contrast, this paper leverages within-country variation to estimate the long-run effects of tourism on both regional and national economic outcomes. The paper also relates to the literature that studies possible "Dutch disease" effects associated with natural resource booms by comparing regional outcomes within countries (e.g. Caselli & Michaels (2009), Allcott & Keniston (2014)). Both the methodology we propose and the focus on tourism as a special kind of natural resource boom differ from the existing literature, but the economic questions are closely related.

Methodologically, the paper follows a recent but growing literature that uses quantitative spatial equilibrium models to analyze the welfare consequences of aggregate or local shocks, taking into account the frictions to trade and mobility between regions within countries (e.g. Redding (2015), Caliendo et al. (2014), Monte et al. (2015), Bryan & Morten (2015), Caliendo et al. (2015), Fajgelbaum et al. (2015) and Galle et al. (2014)).[6] We build on the framework developed by Allen & Arkolakis (2014) and Ahlfeldt et al. (2015) and extend the model and methodology in several ways as discussed above. In particular, we study the role of within and cross-sector agglomeration externalities, a novel dimension in this class of quantitative frameworks. In order to not confound the extent of these spillovers, we also introduce within and cross-sector input-output linkages a la Caliendo & Parro (2014). We then combine the structure of the model with observed empirical moments to identify the strength of the spillover effects, close to the approach in Ahlfeldt et

[5] Eilat & Einav (2004) use panel data on bilateral tourism flows over time to estimate the effect of factors such as political risk or exchange rates on bilateral tourism demand. Sequeira & Macas Nunes (2008) use country-level panel data to estimate the effect of tourism specialization on country growth. Arezki et al. (2009) regress average country-level growth rates over the period 1980-2002 on a measure of tourism specialization in a cross-section of 127 countries, and use the list of UN World Heritage sites as an instrumental variable for tourism specialization.

[6] Work by Ahlfeldt et al. (2015) and Allen & Arkolakis (2016) also follow closely related approaches, but focus on spatial equilibria within cities rather than within countries.

al. (2015). Finally, our approach combines a credible reduced-form estimation with a more structured approach to get at general equilibrium effects, following recent work by Kline & Moretti (2014) and Donaldson & Hornbeck (2015).

The remainder of the paper proceeds as follows. Section 2 describes the background of tourism in Mexico and the data. Section 3 presents the reduced-form evidence. Section 4 presents the theoretical framework that guides the welfare analysis. Section 5 presents the model calibration and parameter estimation. Section 6 presents the counterfactual analysis. Section 7 concludes.

# 2 Background and Data

## 2.1 Tourism in Mexico

According to Mexico's national accounts statistics, tourism activity in Mexico has grown over time to account for about 10 percent of total GDP. As depicted in Table 1, the bulk of this tourism activity is driven by beach tourism that is located among the 150 coastal municipalities. They account for two thirds of total hotel revenues in Mexico.

By 2014, Mexico received 29 million foreign visitors. As depicted in online appendix Figure A.1, this number was close to zero before the 1960s. Beach tourism started to emerge in Mexico during the 1950s and 60s, about three decades after a devastating civil war had ended in the 1920s. By that time, the first generation of Mexican tourist destinations, such as the colonial port city of Acapulco on the Pacific coast or the border city of Tijuana in the North, were starting to emerge and to become popular in Hollywood and among the international jet set. In terms of domestic beach tourism, one major hurdle before the 1960s were prohibitively high travel costs. For example, the first highway to connect Acapulco to Mexico City was completed in 1960. The next generation of Mexican destinations for beach tourism appeared during the 1970s and 80s, which witnessed the emergence of the Yucatan peninsula (e.g. Cancun) and other popular contemporary destinations such as Los Cabos, Ixtapa or Huatulco.

US Americans account for the largest share of foreign tourists in Mexico (57%), followed by Canadians (14%) and Britons (3%).[7] As is the case for the majority of countries in the world, the bulk of total tourism activity in Mexico is driven by domestic travel rather than cross-border tourism, with a share of roughly 80 to 20 percent in terms of revenues over recent decades according to the Mexican tourism satellite account (part of the Mexican national accounts statistics). Against this backdrop, our analysis sets out to quantify the gains from both domestic tourism integration across regions within Mexico and international tourism integration across borders. In particular, the counterfactual analysis aims to decompose the effect of tourism integration (e.g. falling travel costs) into its international and domestic components.

Finally, tourism revenues in Mexico can be divided into different types of expenditure. According to the tourism satellite account, 13 percent are spent on artisanals and other goods, and the rest of tourist expenditure goes to local services, with accommodation (hotels and other temporary accommodation), restaurants and transportation as the three main categories.

[7] Figures are from the Mexican Secretariat for Tourism (SECTUR).

## 2.2 Data

This subsection provides an overview of the main datasets used in the analysis. Table 2 provides descriptive statistics and Figure 1 depicts the satellite and GIS data.

**Censos Economicos for 1998 and 2008** Every five years the Mexican statistical institute INEGI undertakes a census of all economic establishments located in municipalities with more than 2500 inhabitants, and covers a representative sample of establishments in rural locations with less than 2500 inhabitants. The survey questionnaires of these Censos Economicos differ by sector of activity (e.g. construction, retail, manufacturing, etc). In our analysis, we use the municipality-level data of the Censos Economicos 1999 and 2009, which contain information about economic activity in 1998 and 2008 respectively. The timing of these two economic census rounds closely coincide with the two most recent national population censuses in Mexico in 2000 and 2010 that we describe below.

Our main explanatory variable of interest is municipality-level sales of hotels and other temporary accommodation (hostels, etc). In our specifications, we label this variable as hotel sales. Hotels and other temporary accommodation are covered as part of the Censos Economicos Comerciales y de Servicios, from which we obtain two cross-sections of municipality hotel revenues for 1998 and 2008. We combine this information on hotel sales with data from the Censos Economicos for the same years on total municipality GDP, total municipality wage bill, and GDP broken up into manufacturing, services and agriculture.

In the analysis, we will interpret differences in log hotel sales across municipalities as effectively capturing proportional differences in total local tourism expenditures. The reason for this is that the available data for other tourist expenditures, such as restaurants, do not distinguish between sales to local residents as opposed to visiting non-residents. The underlying assumption is that hotel sales are a constant share of total tourist expenditure. Using data from Mexico's tourism satellite account, online appendix Table A.1 documents that this assumption is supported by the available data: accommodation expenditures accounted for on average 13 percent of total tourist expenditure over the period 2003-2013, with very little variation across years.

**Population Census Data** We use IPUMS microdata from the Mexican Population Census in 2000 and 2010 to construct municipality-level total population and employment, as well as individual-level wages including information on gender, education, age and ethnicity. The IPUMS microdata provide us with 10 percent random census samples in addition to population weights that are linked to each observation.

To construct municipality population, we sum up the number of people surveyed and weight the summation by population weights. To construct total municipality-level employment, we make use of the fact that the Mexican population censuses in 2000 and 2010 asked people in which municipality they work, and sum up the number of people (again weighted by population weights) that work in a given municipality. To verify that the 10 percent random samples from IPUMS do not give rise to concerns about sparseness given our focus is at the municipality level, we also report a robustness check using municipality-level population data that is computed from 100 percent samples at INEGI.[8]

In order to construct wages, we first divide monthly incomes by four times the reported weekly hours worked in the census data. We then construct Mincerized wage residuals from a regression of log wages on dummies for gender and ethnicity in addition to the cubic polynomials of years of education and years of age as well as census year fixed effects. We weight these regressions by population weights. The final step is to take the population weighted average of the log wage residuals by year and municipality in the data.

In addition to the two most recent census rounds, we also use Mexican population census data for the years 1921, 1930, 1940 and 1950 in order to estimate a set of placebo falsification tests. To do so, we use INEGI's database Archivo Historico de Localidades to construct time-consistent municipality-level spatial units for the year 2010 that we can trace back to 1921. To do so, we extract the history of each census tract that existed in each of the 10 national population censuses conducted between 1921-2010. For example if municipality boundaries changed over time, or a census tract was split or merged, these instances are reported and traceable.

**GIS and Satellite Data** We use GIS and satellite data to build various measures of the attractiveness for beach tourism. As discussed in the next section, we use these measures to build a set of instrumental variables that influence local tourism demand. To this end, we use the earliest high-resolution satellite data that we could obtain.[9] The data source is the Global Land Survey (GLS) 1990 dataset that is based on the raw data from the LandSat 4-5 Thematic Mapper (TM). The GLS dataset provides a consolidation of the best quality LandSat imagery that were taken during the period of 1987-1997 over the coast of Mexico. We obtained these data at the original resolution of 30x30 meter pixels for six different wavelength bands: Band 1 covers 0.45-0.52, Band 2 covers 0.52-0.60, Band 3 covers 0.63-0.69, Band 4 covers 0.76-0.90, Band 5 covers 1.55-1.75, and Band 6 covers 2.08-2.35.[10]

When restricted to a 2 km buffer around the Mexican shoreline, these satellite data provide us with six raster data layers that each have approximately 52 million 30x30 meter pixels. Each pixel reports the wavelength value of the given bandwidth that the data layer corresponds to. Figure 1 provides an illustration of the satellite data when illustrated with all six bands for all the GLS data tiles that intersect with the Mexican coastline.

The satellite data also provide us with detailed information on coastal elevation and relief at the same level of spatial resolution (30 m pixels). For a robustness check discussed below, we use these data to construct the mean and the standard deviation of coastal elevation within 200 m of the shoreline for each coastal municipality.

[8]While the 100 percent sample data are available for total population, we do not have access to the microdata, which we would need to compute Mincerized wages as well as employment.

[9]We are interested in historical satellite coverage to limit the potential concern that some municipalities invest more to maintain high quality beaches (e.g. efforts against coastal erosion). As we discuss in the empirical section, we also present a number of additional robustness checks against such concerns (e.g. reporting results before and after including controls, and verifying to what extent the island instrument yields similar point estimates).

[10]We do not make use of a seventh band covering thermal infrared (10.40-12.50) that was only recorded at a resolution of 120 instead of 30 m pixels.

We combine these satellite data with a number of GIS data layers that we obtain from the Mexican statistical institute INEGI. These geo-coded data layers include the administrative shape file of municipality boundaries for the 2010 population census, the position of the Mexican coastline, the Mexican transport network for the year 2009 (airports, seaports, paved roads and railways), and the coordinates for each island feature within the Mexican maritime territory from the Mexican census of maritime land territory. The second panel in Figure 1 depicts the position of islands within 5 km of the Mexican coast.

Finally, we also obtain GIS data from two additional sources. The first is a measure of monthly temperature and precipitation at the level of 30 arc seconds (roughly 1km) for the period 1950-2000 from the WorldClim database. We take annual means of precipitation and temperature from the monthly data and collapse the grid cells to the municipality-level mean values of these two variables. The second is a measure of primary ocean productivity at the level of 0.1 degree cells from the Nasa Earth Observation (NEO) program. Primary productivity indicates the amount of biomass created from photosynthesis (measured by chlorophyll concentrations), which is an important determinant of the density of fish populations that can be sustained. We use these data to measure the mean primary ocean productivity within 50 km of the coastline among coastal municipalities for the year 2005.

**Data on Bilateral Tourism Exports 1990-2011** To estimate the tourism trade elasticity, we use data on bilateral tourism exports from the World Bank WITS database on trade in services. We link these data to information from the IMF on PPP rates for final consumption goods across countries in order to empirically capture the relative price of local consumption for origin-destination country pairs over time. The database spans the years 1990-2011 and includes 115 origin and destination countries.

# 3 Reduced-Form Evidence

This section uses the data described above in combination with the empirical strategy outlined below to estimate a number of reduced-form effects of tourism in today's cross-section of Mexican municipalities on municipality-level employment, population, wages and local GDP by sector of activity. As well as being of interest in their own right, these reduced-form moments inform the calibration of the model in Section 5 and the quantification of tourism's welfare implications in Section 6. The section proceeds in three parts. We first describe the empirical strategy, and proceed to discuss the estimation results. We then present additional results to further investigate the interpretation of the observed effects.

## 3.1 Empirical Strategy

Given beach tourism has had more than half a century to materialize into today's observed regional economic outcomes in Mexico, our aim is to exploit cross-sectional variation to capture tourism's long-term economic consequences across Mexican municipalities. To estimate the effect of differences in local tourism revenues on relative municipality-level economic outcomes in today's cross-section of Mexican municipalities, we estimate the following baseline specification:

$$log\left(y_{nct}\right) = \alpha_{ct} + \beta log\left(HotelSales_{nct}\right) + \alpha' X_{nct} + \epsilon_{nct} \tag{1}$$

where $n$ indexes municipalities, $c$ indexes coastal versus non-coastal municipalities and $t$ indexes census years. In our baseline specification, we regress the two most recent cross-sections of municipality-level outcomes, $y_{nct}$, in 2000 and 2010 for outcomes computed using the population censuses, and in 1998 and 2008 for outcomes computed using the Censos Economicos, on the log of municipality sales of temporary accommodation (hotels, hostels, vacation rentals) in 1998 and 2008, a vector of pre-determined municipality controls, $X_{nct}$, and coast-by-period fixed effects.[11] To address concerns about autocorrelated error terms for the same municipality over time, we cluster standard errors at the municipality level.[12] As noted in the previous section, we address the lack of data for total local tourism expenditure by making the assumption that variation in log hotel sales effectively captures proportional changes in total tourism expenditure across municipalities. As documented in online appendix Table A.1, the assumption of a constant share of accommodation in total tourist expenditure seems to be supported in the available data.

The main concern for causal identification in (1) is that municipalities with higher hotel sales are also subject to other local unobserved factors that affect both tourism activity as well as economic outcomes. For example, economically vibrant municipalities could report higher tourism sales because of business travel. Similarly, hotels could locate in municipalities with better transport links or skilled labor with foreign language skills. Conversely, tourism could locate in remote locations with cheaper land prices where hotel resorts can find large stretches of available space with little opportunity cost for land use. A third possibility is that given the bulk of Mexican tourism appears to be beach-oriented (see Table 1), tourist resorts could instead follow a quite specific set of natural amenities that are largely unrelated to economic outcomes except than through tourism.

To address such concerns and investigate which of these scenarios is the case in our empirical setting, we propose the following empirical strategy that proceeds in several steps. In the first step, we report how OLS estimates of $\beta$ are affected before and after including an additional set of pre-determined municipality controls. In the baseline specification, $X_{nct}$ includes the log distance to Mexico City, the log distance to the closest stretch of the US border and the log municipality area. These geographical controls are aimed to address concerns that larger municipalities that are located close to the main domestic or foreign economic centers have both higher tourism sales as well as more economic activity on the left hand side of specification 1. We then report how the estimate of $\beta$ is affected after additionally including dummies for state capitals, historical cities[13], colonial ports, as well as the logarithm of the average annual temperature and the average annual

[11] We use the inverse hyperbolic sine (IHS) transformation, $log\left(HotelSales_{nct} + \left(HotelSales_{nct}^2 + 1\right)^{1/2}\right)$, in order not to throw away variation from municipalities in places with zero hotel sales. In practice, this does not affect the estimates since the identifying variation in our IV estimation stems from coastal municipalities that except for three instances in the two cross-sections have no reported zeroes for hotel sales. As discussed below, we also report results without this transformation, or after assigning the log of 1 to values of zero.

[12] Clustering instead at the state-level or the state-by-year level leads to slightly smaller standard errors.

[13] Following INEGI's definition of cities with a population above 20k in 1930.

precipitation. Reporting point estimates before and after adding these controls helps us document the extent to which variation in local tourism activity within a given coast-by-year cell may be correlated with a number of observable pre-determined confounding factors that also matter for local economic outcomes.

In the second step, we then construct a number of instrumental variables for $log\,(HotelSales_{nct})$. We take inspiration from the tourism management literature (e.g. Weaver et al. (2000), Leatherman (1997)) arguing that tourism activity is to a large extent determined by the quality of a set of very specific local natural amenities. We identify two criteria for touristic beach quality that we can empirically capture along the roughly 9500 km of Mexican coastline using our GIS database: i) the presence of a nearby offshore island; and ii) the fraction of coastline covered by white sand beaches.

The first instrumental variable that we construct is whether or not a coastal municipality has access to an offshore island within 5 km of its coastline.[14] This measure is aimed at capturing both scenic beauty, as well as the availability of popular beach activities, such as snorkeling around the island or taking a boat trip to the offshore beaches. To measure offshore islands, we use the Mexican census of maritime land territory conducted by the INEGI. To assess the sensitivity of the 5 km cutoff, we alternatively report results using islands within 10 km of the shoreline.

The second set of instrumental variables is aimed at capturing the presence of picturesque white sand beaches along the Mexican coastline. Their construction using the satellite data is slightly more involved. Because an explicit quantifiable specification of what constitutes an attractive stretch of beach in Mexico has not been formulated in the remote sensing literature, we proceed by binding our hands to the best existing ranking of Mexican beaches that we could find. That ranking refers to the "Eight Best Beaches of Mexico" published by the ranking analytics company U.S. News and World Report.[15]

We take the top four of these eight beaches, Playa del Carmen, Tulum, Cozumel and Cancun, and construct 5 alternative municipality-level beach measures using the historical satellite data. For each of these beaches, we start by computing the wavelength ranges in the six different LandSat sensors computed across all 30 m pixels that cover the beach. Online Appendix Table A.2 presents these 6 x 4 ranges. We then use raster processing tools in ArcGIS to classify all 30 m pixels within 100 m of the Mexican shoreline into zeroes and ones depending on whether they fall within the wavelength ranges in each of the six original LandSat raster layers. By aggregating up which pixels are within the range of all six wavelength ranges, this yields four different measures of the fraction of coastline within 100 m of the shoreline that is covered by either definition of picturesque

[14] Our instrumental variables have no variation across non-coastal municipalities (we set them to zero for inland regions). Given specification (1) features coast-by-period fixed effects, it follows that the identifying variation is purely within the coastal municipality group. We include the full sample of Mexican municipalities to increase power when estimating additional municipality controls in $X_{nct}$. As a robustness check, we also allow the controls to have heterogeneous effects among coastal and non-coastal regions.

[15] In their description (http://travel.usnews.com/Rankings/Best_Mexico_Beaches/), they write: "*To help you find the ideal Mexican destination for sunbathing on the sand and splashing in the waves, U.S. News considered factors like scenery, water clarity, crowd congestion, and nearby amenities. Expert insight and user votes were also taken into account when creating this list of the country's best beaches.*"

white sand beaches and for each of the 150 coastal municipalities. In addition to these four instrumental variables, we also construct the fraction of 30 m pixels within 100 m of the shoreline that is covered by either of these four types of high quality beaches. Finally, to assess the sensitivity to the 100 m range, we also report results using a 200 m distance from the shoreline instead.

Having constructed these six instrumental variables (one for islands and five for beach quality), we proceed as follows. We use the island instrument and the beach quality instrument based on the top ranked beach (Playa del Carmen) as our baseline instrumental variable strategy. The identifying assumption is that the presence of an offshore island within close proximity of the shoreline or a higher fraction of coastline within 100 m of the shore covered by white sand beaches affect municipality-level economic outcomes relative to other coastal municipalities only through their effect on local tourism activity. To assess this assumption, we report the IV point estimates both before and after including the full set of municipality control variables, and test whether the island instrument and the beach instrument –which rely on very different variation across the Mexican coastline (Figure 1)– yield similar point estimates of the effect of tourism on local economic outcomes. We also report a number of additional robustness checks that we discuss in detail below.

## 3.2 Estimation Results

### Municipality Employment and Population

Using the empirical strategy outlined above, we begin by estimating the effect of differences in local tourism activity on municipality-level total employment and population. Viewed through the lens of a spatial equilibrium with labor mobility, these are two of the most informative long-term regional economic outcomes. To this end, we estimate specification (1) with log employment or log population on the left hand side that we construct from the Mexican census microdata for 2000 and 2010 as described in the data section. Table 3 presents the OLS and IV estimation results for our two baseline instrumental variables (the island IV and the first of the five beach instruments).

The OLS point estimate of the effect of tourism on municipality employment is statistically significant at the 1 percent level and changes little before and after including the full set of municipality controls. Given that the vast majority of Mexican tourism is beach-oriented and located along the coastline, and the fact that our baseline specification includes coast-by-period fixed effects, one interpretation of the OLS results is that tourism in Mexico is to a large extent determined by a specific set of natural amenities, such as beaches, that appear not be strongly correlated with some of the obvious observable and pre-determined control variables.

To further assess these results, columns 3-8 present the IV estimates. As for the OLS, the IV point estimates of the effect of tourism on municipality total employment change very little before and after including the full set of controls for both the island instrument and the beach instrument, as well as when using both instruments jointly. Both instruments lead to slightly higher point estimates compared to the OLS of the effect of tourism on municipality employment in the full specifications, and both instruments yield similar point estimates as reported by the p-value of the over-identification test in columns 7 and 8. The likeliest explanation for why the IV point estimates are higher than the OLS estimates is the concern of measurement error in our measure of local tourism activity, which relies on reported establishment revenues in the survey questionnaires of the Censos Economicos for 1998 and 2008.

The results suggest that local tourism activity has a strong and significant positive effect on total municipality employment relative to other municipalities. The elasticity is estimated to be 0.28 in the full specification with both instruments in column 8, suggesting that a 10 percent increase in local tourism activity in 1998 and 2008 on average leads to an increase in total municipality employment of on average 2.8 percent in 2000 and 2010 respectively. Given these estimates are based on cross-sectional variation, we interpret them as long-term effects of local exposure to tourism on municipality total employment relative to other regions.

Columns 9 and 10 of Table 3 report the estimation results of the IV specification using both instruments on log municipality total population instead of total employment. Interestingly, the point estimate is about 0.055 below the point estimate for employment, suggesting that a 10 percent increase in local tourism expenditure leads to an increase of 2.2 percent in total population compared to 2.8 percent in total employment. We interpret this as indicative evidence that some workers who are attracted to municipalities with more tourism activity do not end up residing in the same municipality. Having said this, 0.05 is a relatively small difference in the two point estimates. This is consistent with the raw moments in our census microdata, where the total share of workers commuting outside their residential municipality for work is 15 percent, and falls to below 10 percent once we exclude the Mexico City region.

**Robustness**

To further assess the robustness of these results to potentially remaining concerns about the IV identification strategy, we estimate and report a number of additional results in Tables 4-6, as well as online appendix Tables A.3-A.6.[16] In in columns 1-3 of Table 4, we address the concern that the island or beach instruments may be correlated with sea accessibility by picking out municipalities with relatively flat terrain relative to coastal cliffs and/or rugged terrain. A related concern could be that the island or beach instruments may be correlated with oceanographic conditions that affect the local fishery sector. To this end, we use additional high-resolution remote sensing data on elevation as well as ocean primary productivity. We report the previous IV point estimates

[16] Table A.3 confirms that the estimation results are not sensitive to using the IHS transformation for dealing with zero hotel revenues in the log specification. The reason is that the identifying variation of the IV strategy stems from differences across coastal municipalities, which except for three instances report positive amounts of hotel revenues. Table A.4 confirms that the 10 percent census samples do not give rise to sparseness concerns for our analysis at the municipality level. To this end, we report results on municipality population when measured from the 100 percent census samples, rather than our 10 percent samples for which we have the microdata. Table A.5 first confirms that the identifying variation is purely driven by coastal municipalities, and then reports close to identical point estimates after allowing all municipality controls to be interacted with the coastal region dummy. Finally, Table A.6 addresses the concern that the first stage F-statistic drops from 14.4 to 12.6 when including the full set of controls in the IV specification using both IVs in columns 8 and 10 in Table 3. To this end, we compare two stage least squares estimates using both IVs to limited information maximum likelihood (LIML) estimates. Given that the LIML estimator has been found to be more robust to weak instrument bias, the fact that the reported LIML point estimates are slightly higher provides reassurance against this concern.

of the effect on total employment after including the mean level of elevation within 200 m of the shoreline, its standard deviation, as well as mean ocean primary productivity withing 50 km of the municipality shoreline. As reported in Table 4, the point estimates remain virtually unchanged, providing some reassurance against the omitted variable concerns.

In columns 4 and 5 of Table 4, we test the sensitivity of the IV point estimates to excluding the origin municipality of the top ranked beach in Mexico (Playa del Carmen). This serves to address the concern that the ranking agency U.S. News partly based their ranking on the popularity of destinations by US tourists. Reassuringly, the IV point estimate remains unchanged. Columns 6-9 then report IV point estimates when using the five alternative beach instruments, while also each time excluding the respective origin municipality, and reporting over-identification tests relative to the island instrument. The point estimates are similar, and very slightly higher on average compared to our baseline estimates reported in Table 3. Reassuringly, all six instruments provide similar estimates of the effect of tourism activity on municipality total employment, as indicated by the point estimates in columns 4-9 and the reported p-values of the overidentification tests. The final two columns of Table 4 report results aimed at testing the sensitivity of the 5 km cutoff for the island instrument, and the 100 m cutoff for the beach instrument. The point estimates remain practically unchanged when doubling those cutoff values to 10 km and 200 m respectively.

In Tables 5 and 6 we address two potentially remaining concerns. First, maybe the previous robustness checks missed omitted variables that are both correlated with the instruments conditional on controls, and affect local economic outcomes. Second, islands or white beaches may affect the local economy not just through their effect on local demand for tourism-related services, but also by directly affecting the amenities of local residents. Even though we tried to be careful in constructing our IVs to capture a very particular set of features of the local environment that are arguably specific to beach tourism, it could be the case that these characteristics have significant direct effects on local employment and populations by altering the amenities of local residents in a significant way (relative to other coastal municipalities in Mexico).

We assess the extent to which this is the case in two ways. First, as reported in Table 5, we run a placebo falsification test on the identical sample of municipalities during a period before beach tourism had become a discernible force in Mexico. This involves the construction of a long time series of population census data for consistent spatial units for the years 1921, 1930, 1940 and 1950 in addition to the two most recent rounds of population census data 2000 and 2010 that we use in our baseline regressions. As discussed in Section 2, beach tourism in Mexico started to emerge in the 1950s and 1960s. To further assess the validity of our IV identification strategy, we regress log municipality population on each of the six instrumental variables for the same set of municipalities both before and after beach tourism could have significantly affected economic outcomes and thus local populations.

Table 5 reports the results of these specifications. We report the results across three panels, that deal in different ways with the potentially important feature of the data that not all municipalities reported non-zero populations for all census rounds between 1921-2010.[17] The first panel uses the

[17]We include the basic set of geographical controls used in the previous tables rather than the complete set, as some

same strategy that we use for the log hotel sales above, and uses the IHS transformation on the left-hand side in order to not ignore zero populations. The second panel replaces zero population values with the log of 1, instead. And the third panel reports results for log populations on the left hand side, while ignoring all municipalities that ever reported zero populations.

For the island instrument, we get a slightly negative but insignificant point estimate of the effect on municipality populations before 1960, and a significant positive effect afterward in all three panels. Importantly, the estimates on the geographical municipality controls are estimated with similar precision in both periods, providing some reassurance against the concern that the historical census population data could simply be more noisy than the more recent data. For the five remaining beach instruments, the reduced-form effect on population in the recent periods is slightly less precisely estimated than that for the island instrument. As for the island IV, the point estimate for the period before 1960 is negative and imprecisely estimated. The fact that the pre-tourism point estimates are consistently negative for all five beach instruments, and sometimes marginally statistically significant, points to the fact that an abundance of attractive beach characteristics may have been somewhat negatively correlated to municipality populations along the coastline before tourism emerged (pre-1960).

This pattern starts to make sense when taking a closer look at the U.S. News beach ranking: the highest ranked beaches are concentrated in the Caribbean part of Mexico along the Yucatan coast. These coastal municipalities were virtually empty fishing villages before tourism started growing in the region (e.g. Cancun in the 1970s and 80s). Both the fact that tourism is very plausibly the main reason for why these places switched from less to more populous,[18] and the fact that our empirical analysis is interested in cross-sectional differences in levels rather than growth rates –where mean reversion could be a concern– provide us with reassurance that our IV point estimates are unlikely to be biased upwards. The three panels of Table 5 confirm these findings across the different treatments of zero populations in our log specification on the left hand side.

The final robustness exercise is aimed at addressing the potentially remaining concern that while Mexicans may not have cared about white sand beaches or islands when deciding where to live and work along the coastline in the 1950s, their tastes may have evolved over time so that by 2000 these instrumental variables pick up significant direct amenity effects (again, relative to other coastal municipalities). To this end, we verify in today's cross-section of Mexican municipalities to what extent our model-based estimates of local amenities that we derive in Section 4 (essentially local population residuals left unexplained by spatial variation in real incomes) are significantly related to the presence of islands or a higher fraction of white sand coverage along the coastline.

Table 6 reports the estimation results. We construct these model-based measures of local amenities five times. Each time, we exclude one of the five instruments (1 island IV and the 4 different beach IVs)[19] in all steps of the model's parameter estimation in order to ensure that there

of the controls were arguably not pre-determined in the early census periods. Results remain unchanged, however.

[18]For some interesting background on this point, see the New York Times article "Why the Computer Chose Cancun" published on March 05 in 1972.

[19]Note that the sixth instrument (the fifth beach instrument) is a linear combination of the first four beach instruments.

is no mechanical orthogonality condition built into the estimation of the local amenities.[20] Consistent with the findings of the placebo falsification test above, we find that current-day estimates of local amenities are not significantly correlated with the instruments.

In summary, the additional results discussed above provide some reassurance that our measures of islands or the fraction of coastline covered by white sand beaches capture a specific set of shifters to local tourism demand that do not appear to have discernible direct effects on local populations, or to be correlated with other omitted variables affecting local economic outcomes.

**Municipality Wage Bill, GDP by Sector and Wages**

Table 7 reports the OLS and IV estimation results of the effect of differences in local tourism revenues on the municipality-level total wage bill (labor income), GDP and GDP by sector of economic activity. Tourism has a strong and significant positive effect on local aggregate labor income and GDP. According to the IV point estimates in the full specification with both instruments, a 10 percent increase in local tourism revenues leads to 4.8 percent increase in the local wage bill, and a 4.3 percent increase in local GDP.

Given tourism only accounts for on average roughly 10 percent of total GDP in Mexico (and about 20 percent among coastal municipalities),[21] these results suggest strong multiplier effects on the local economy. Interestingly, the strong effect of tourism on total local GDP appears to be driven by significant positive effects on local manufacturing GDP, while the point estimate on local agriculture is also positive, but not significant at conventional levels in the IV estimation. According to the point estimate reported in column, a 10 percent increase in local hotel revenues leads to an increase in local manufacturing GDP of 3.2 percent.

Finally, the last two columns of Table 7 report OLS and IV estimates of the effect of local tourism expenditure on average municipality wages. The dependent variable of interest here are Mincerized log wage residuals as described in Section 2 (flexibly controlling for age, education, gender and ethnicity using the microdata of the Mexican population censuses). We find that changes in local tourism exposure have a positive and significant effect on local wages with an elasticity of 0.033 in the full IV specification with both instruments. A 10 percent increase in local tourism revenue leads to an increase of 0.33 percent in local nominal wages.

## 3.3 Discussion

The previous subsection suggests that tourism activity has led to large and significant local increases in total employment, population, wages and GDP relative to less touristic regions in today's cross-section of Mexican municipalities. Underlying these results, we find sizable positive multiplier effects on traded sector production. Given tourism activity in Mexico has had several decades to materialize into today's observed outcomes across regions, these findings plausibly reflect long-term regional economic effects.

[20] See discussion in Section 5 where we refer back to Table 6.

[21] As discussed above, these figures are not directly observable and based on the assumption that hotel expenditures account for on average roughly 15 percent of total tourism expenditure (from the Mexican national tourism satellite account).

In this subsection we address two important remaining questions concerning the interpretation of these results. The first question is to what extent the effects are driven by tourism's effect on infrastructure investments, such as airports, seaports, roads or railways. The second question is to what extent the positive multiplier effect on traded sector production may be driven by a subset of manufacturing sectors that are used intensively as local inputs for tourism-related services.

To address the first of these questions, we present a number of additional results that we report in Table 8. The table presents IV estimates of the effect on total employment, population, GDP and manufacturing GDP both before and after additionally conditioning on differences in the log distance from the municipality centroid to the nearest airport, the log distance to the nearest sea port, the log of the paved road kilometers in the municipality and the log of the railway kilometers in the municipality. Interestingly, the point estimates conditional on differences in access to infrastructure are only slightly lower compared to our full baseline specification and confirmed in terms of their statistical significance. In particular, the point estimates of the effects on total local employment, population, GDP and manufacturing GDP change from 0.28 to 0.26, from 0.22 to 0.2, from 0.43 to 0.4 and from 0.32 to 0.29 respectively. Reassuringly, all four infrastructure variables enter significantly and in the expected sign in all specifications. These results suggest that while tourism's effect on local economic activity appears to be partly driven by increased investments in local infrastructure, the vast majority of the observed effects remain after conditioning on this channel. Reflecting these results, our quantitative analysis in 6 will also explore the sensitivity of the estimated gains from tourism after allowing the development of tourism to endogenously reduce transport costs for trade in goods.

Turning to the second question that we raise above, we report a set of additional estimation results in Table 9. In particular, we break up the 21 three-digit manufacturing sectors into above and below-median intensity of touristic input use among traded industries. In particular, we construct two different measures. The first is based on the three-digit level total requirement coefficients from the 2007 Mexican input output tables. We use the total (direct and indirect) input requirement coefficients for the hotel sector across the 21 manufacturing sectors, and divide these sectors into above and below the median. The two most intensively used input sectors are chemical products and petroleum/carbon-based products (both used in building hotels and resorts), and the two least used input sectors are leather products and the food industry.

Alternatively, to better capture sectors that tourists demand directly –rather than solely relying on what the hotel sector uses as inputs in the Mexican IO tables– we also construct a second measure of tourism's input intensity. In particular, the Mexican tourism satellite account splits up total tourist tradable consumption into five 3-digit sectors. These are (in decreasing order of importance): the food industry, artisanal products (part of other manufacturing), pharmaceuticals (part of chemical industry), clothing industry, and printed media (part of printing industry). In addition to the input-output based groupings of our first measure above, we use these five sectors as our second binary measure of traded sectors which may be used intensively by the tourism sector.

As reported in Table 9, we find that, as expected, sectors more intensively used in tourism are slightly more strongly affected by variation in local tourism activity. At the same time, the positive multiplier effects remain sizable and statistically significant in sectors with below median tourism input intensity. These results suggest that part of the positive effect of tourism on local traded goods production may be driven by better market access for local input suppliers to tourism. To reflect this finding in our quantification, we allow for input-output linkages between tourism related services and manufacturing in the theoretical framework that follows.

# 4 Theoretical Framework

In the previous section we have estimated a number of reduced-form effects of variation in tourism activity on local economic outcomes in today's cross-section of Mexican municipalities. While interesting in their own right, these reduced-form results also leave a number of important questions unanswered.

First, the estimates are by construction based on relative effects, since the empirical setting is based on comparing outcomes across regions with higher or lower levels of local tourism activity. This limitation is particularly acute since we are interested in estimating the long-term effects of tourism in Mexico: as we report above, local populations and employment strongly respond to differences in tourism activity, suggesting that the initial regional welfare differentials brought about by tourism activity have been smoothed by mobile labor over the long run (since the 1950s). We thus require a more structured approach to evaluate the aggregate national implications of tourism.[22]

Second, the previous section suggests that tourism has strong positive effects on local economic activity, both directly and indirectly, i.e. through its effect on manufacturing production. To what extent are these estimated multiplier effects a sign of possible productivity spillovers between the development of tourism and traded goods production? The answer is a priori unclear, as this result could be driven by cross-sector spillovers, but also by local demand effects: local population and input demand from the tourism sector increase, both of which improve the market access for local manufacturing once we allow for trade costs. Furthermore, to the extent that these multiplier effects do reflect productivity spillovers, it is also a priori unclear whether such localized effects on manufacturing may be offset by a decrease in agglomeration forces in other non-touristic regions of the country. These questions naturally feed back into the welfare evaluation of tourism: depending on the sign and magnitude of within and cross-sector agglomeration forces, the aggregate gains from tourism can either be magnified or diminished compared to the conventional neoclassical gains from market integration in tourism.

To make progress on these questions and guide the estimation of the welfare implications of tourism in the long run, we propose a spatial equilibrium model of trade in goods and tourism services. The following subsections outline the theoretical framework, Section 5 describes the model calibration, and Section 6 presents the counterfactual analysis.

[22] For a closely related discussion see for example Kline & Moretti (2014).

## 4.1 Model Setup

The theoretical framework builds on existing work by Allen & Arkolakis (2014), Ahlfeldt et al. (2015) and Redding (2015), and extends it in three dimensions. First, we allow for trade in tourism-related services in addition to trade in manufacturing goods across regions and countries. Second, we follow Caliendo et al. (2014) in allowing for input-output linkages between sectors. Third, we allow for cross-sector co-agglomeration economies in addition to within-sector agglomeration economies.

The model features regions within Mexico that differ in three dimensions: their level of productivity for manufacturing goods, their level of attractiveness for tourism, and their level of local amenities for residents. Furthermore, regions are linked economically through three ties. First, they trade goods with each other and the rest of the world. Second, they host international and domestic tourists that spend part of their income outside of their region of residence. Third, workers are mobile and choose their region of residence within countries.

The model is static and aims at capturing the long run equilibrium of the economy. The world is comprised of $N$ regions indexed by $n$. Labor is mobile between regions within countries but not between countries. The subset $\mathcal{M} \subset (1..N)$ corresponds to the regions of Mexico. The subset $\overline{\mathcal{M}}$ designates countries other than Mexico. For simplicity, we do not model intra-country heterogeneity for them. The total population of each country is taken as given: for countries other than Mexico, $L_n$ for $n \in \overline{\mathcal{M}}$ is exogenous; for Mexico, total population $L_{\mathcal{M}} \equiv \sum_{n \in \mathcal{M}} L_n$ is also given. In contrast, the share of workers in each Mexican region $\frac{L_n}{L_{\mathcal{M}}}$ for $n \in \mathcal{M}$ is an endogenous outcome, determined in spatial equilibrium.

**Households**

In each region $n \in 1..N$, there is a population of $L_n$ workers. Each worker supplies one unit of labor inelastically. Workers derive utility from the consumption of a bundle of goods and services as well as from the local amenities of the region where they live, subject to idiosyncratic preference shocks. The utility of a worker living in region $n$ of her country is:

$$U_n(\omega) = \varepsilon_n(\omega) C_n B_n L_n^{\epsilon}, \tag{2}$$

where $C_n$ is the consumption bundle of goods and services, $B_n$ is the exogenous amenity differences between regions, and term $L_n^{\epsilon}$ allows for that level of attractiveness of a region to respond endogenously to how populated a region is. This aims to capture, in a reduced-form way, the notion that more populated regions can be either more congested, leading to a decrease in the utility of local residents (if $\epsilon \leq 0$), or more attractive, as the concentration of population gives rise endogenously to better local amenities (e.g. more sources of entertainment, variety in consumption, etc). Finally, each worker has a set of idiosyncratic preferences for living in different regions in his country. We denote this vector of idiosyncratic preferences $\varepsilon_n(\omega)$ for worker $\omega$ and regions $n$ of his own country, and assume that they are drawn from a Frechet distribution with mean 1 and dispersion parameter $\kappa$. Workers within Mexico choose to live in the region that maximizes their utility, so that:[23]

$$U(\omega) = \max_{n \in \mathcal{M}} \varepsilon_n(\omega) C_n B_n L_n^{\epsilon}.$$

The goods and services workers consume are a bundle of local non-traded services ($C_s$), tourism-related services ($C_T$) and manufacturing goods ($C_M$), according to the following preferences:[24]

$$C_n = \left( \frac{\left[ C_{M,n}^{\frac{\rho-1}{\rho}} + C_{T,n}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}}{\alpha_{MT}} \right)^{\alpha_{MT}} \left( \frac{C_{S,n}}{\alpha_S} \right)^{\alpha_s}, \tag{3}$$

where the elasticity of substitution between tourism services and manufactured goods is $\rho > 1$, and $\alpha_{MT} + \alpha_S = 1$. Local services represent a constant share of spending.[25]

Given the demand function (3), the share of total spending in region $n$ spent on manufactured goods is:

$$\alpha_{MT} \chi_n \equiv \alpha_{MT} \frac{P_{M,n}^{1-\rho}}{P_{MT,n}^{1-\rho}},$$

where $P_{MT,n}$ is the composite price index for the bundle of manufactured and tourism goods, $P_{M,n}$ is the price of the composite manufacturing good, and $P_{T,n}$ is the price of the bundle of tourism-related services for a consumer located in region $n$:

$$P_{MT,n} = \left( P_{M,n}^{1-\rho} + P_{T,n}^{1-\rho} \right)^{\frac{1}{1-\rho}}.$$

**Manufacturing Production**

Production in manufacturing is as follows. Intermediate varieties are produced in each region using labor and a manufacturing composite good as an input to production. A competitive local sector aggregates intermediate varieties and sells this composite to i) local final consumers and ii) local intermediate producers in manufacturing and tourism, who use it as an input to their production. This is detailed below.

[23]The idiosyncratic preferences and local amenities play no role in the model for workers outside of Mexico as we do not model intra-country heterogeneity for these countries.

[24]More generally, the demand function can be parametrized as $\left( \frac{\left[ \beta_M C_{M,n}^{\frac{\rho-1}{\rho}} + \beta_T C_{T,n}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}}{\alpha_{MT}} \right)^{\alpha_T} \left( \frac{C_{S,n}}{\alpha_S} \right)^{\alpha_s}$, but the preference weights $\beta_M$ and $\beta_T$ that capture the relative strength of consumer tastes for each good cannot be separately identified from the difference in productivities between these two sectors, so we normalize these weights to 1. The calibrated productivities in each sector should be understood therefore as capturing both a productivity effect as well as demand weights.

[25]This is consistent in particular with the interpretation of this local spending as housing expenditure. For example, Davis & Ortalo-Magné (2011) show that housing expenditure constitutes a nearly constant fraction of household income.

**Intermediate Varieties in Manufacturing** A continuum of intermediate varieties $x \in [0,1]$ is produced in the manufacturing sector. Intermediate varieties are produced in each region combining labor and a a manufacturing composite good, according to the production function:

$$q_{M,n}(x) = M_n z_n(x) L_{M,n}(x)^{\nu_M} m_{M,n}(x)^{1-\nu_M},$$

where $q_{M,n}(x)$ is the quantity of the intermediate variety produced, $M_n$ is the local productivity in manufacturing, common to all varieties in region $n$, $z_n(x)$ is the variety $x$-specific efficiency in region $n$, $L_{M,n}(x)$ is labor and $m_{M,n}(x)$ is manufacturing composite good used for the production of variety $x$.

We allow for the possibility of local production externalities: the productivity of a region for manufacturing goods can respond to the level of local economic activity. This externality can stem from the level of economic activity in the manufacturing sector ($L_{M,n}$) and/or the level of economic activity in the services sector ($L_{ST,n} = L_{T,n} + L_{S,n}$). In each case, the externality increases with the size of economic activity with a constant sector-specific elasticity (denoted respectively $\gamma_M$ and $\gamma_S$), so that:

$$M_n = M_n^o L_{M,n}^{\gamma_M} L_{ST,n}^{\gamma_S}, \tag{4}$$

where $M_n^o$ is the exogenous component of local productivity. This expression captures in a reduced-form way the channels through which local tourism expenditures have positive or negative effects on manufacturing in the long run, beyond their neoclassical demand linkages. For example, it has been hypothesized that tourism could act as a special case of the "Dutch Disease" and attract activity away from more innovation-intensive traded industries, so that the long-term effect on productivity could be negative. Expression (4) allows for tourism to have such adverse long-run economic consequences if, for example, $\gamma_M > 0$ but $\gamma_S = 0$. In that case, the development of tourism attracts workers away from manufacturing, a sector in which scale matters for productivity, causing a decrease in manufacturing productivity. On the other hand, tourism could give rise to productivity spillovers that would not have materialized otherwise – if, for example, $\gamma_S > 0$ while $\gamma_M = 0$. There are a number of channels through which the development of tourism can a priori lead to positive spillovers on the manufacturing sector. For example, the development of tourism can improve the provision of local business services, such as finance, accounting or consulting. Moreover, tourism revenues can also directly loosen the credit constraints of local firms. Alternatively, tourism could lead to a better trained local workforce, or facilitate the fostering of domestic and international business networks through increased travel activity. All these effects are summarized by the parameter $\gamma_S$.

Beyond these region-specific productivity levels, regions draw idiosyncratic efficiency levels $z(x)$ for each variety from a Frechet distribution with shape parameter $\theta$ and mean 1 in all regions:

$$F(z) = e^{-Z^{-\theta}}.$$

Finally, we assume that firms behave competitively. Therefore, firms price at unit cost $\frac{c_{M,n}}{M_n z_n(x)}$

where we have defined the unit cost of the local input bundle in region $n$:

$$c_{M,n} = \Psi_M w_n^{\nu_M} P_{M,n}^{1-\nu_M},$$

where $\Psi_M$ is a constant.[26] Firms incur an iceberg trade cost $\tau_{ni}$ to ship the manufacturing good from region $i$ to region $n$.

**Composite Manufacturing Good** A perfectly competitive local sector aggregates these intermediate manufacturing varieties. Producers of the composite manufacturing good make zero profit and have no value added. They source across regions and countries and purchase intermediate varieties from the lowest cost supplier. The composite manufacturing good is a CES aggregate of individual varieties $x \in [0,1]$ with elasticity of substitution $\sigma_M$:

$$Q_{M,n} = \left[\int q_{M,n}(x)dx\right]^{\frac{\sigma_M}{\sigma_M-1}}.$$

The price index for the composite manufacturing goods is:

$$P_{M,n} = \left[\int p_{M,n}(x)^{1-\sigma_M}dx\right]^{\frac{1}{1-\sigma_M}},$$

where

$$p_{M,n}(x) = \min_{i\in 1\dots N}\left\{\frac{c_{M,i}\tau_{ni}}{M_i z_i(x)}\right\},$$

as local aggregators in region $n$ source from the lowest cost region. Given the properties of the Frechet distribution that governs local efficiency levels, the share of manufacturing spending that region $n$ spends on goods produced in region $i$ is:

$$\pi_{ni} = \frac{(\tau_{ni}c_{M,i})^{-\theta}M_i^{\theta}}{\sum_{k=1}^{N}(\tau_{nk}c_{M,k})^{-\theta}M_k^{\theta}}, \tag{5}$$

and the price index for the composite manufacturing good in region $n$ is:

$$P_{M,n} = \left[K_1\sum_{k=1}^{N}(\tau_{nk}c_{M,k})^{-\theta}M_k^{\theta}\right]^{-\frac{1}{\theta}}. \tag{6}$$

where $K_1 = \left(\Gamma(\frac{\theta-\sigma_M+1}{\theta})\right)^{\frac{1}{1-\sigma_M}}$ is a constant.

**Tourism-Related Services**

Workers living in region $n$ consume a bundle of tourism-related services $C_{T,n}$. They travel to various destination regions, including abroad, to consume these services. We assume that tourism-related services are differentiated by region of destination. The bundle of tourism-related services consumed by a worker living in region $n$ is a CES aggregate of the services consumed in each

[26]Specifically, $\Psi_M = (\nu_M)^{-\nu_M}(1-\nu_M)^{-1+\nu_M}$.

region $i$:[27]

$$Q_{T,n} = \left[\sum_{i\neq n} A_i^{\frac{1}{\sigma_T}} q_{T,i}^{\frac{\sigma_T - 1}{\sigma_T}}\right]^{\frac{\sigma_T}{\sigma_T - 1}},$$

where $\sigma_T$ is the elasticity of substitution between the various touristic destinations, $q_{T,i}$ is the amount of tourism-related services consumed in region $i$ and $A_i$ is a taste shifter for each destination region $i$. It summarizes the quality of the local site for tourism.[28] For example, a site with attractive beaches or a rich set of historical buildings is more attractive for tourists. Given the demand function, the price index $P_{T,n}$ of tourism-related services for the inhabitants of region $n$ is:

$$P_{T,n} = \left(\sum_{i\neq n} A_i {p_{T,ni}}^{1-\sigma_T}\right)^{\frac{1}{1-\sigma_T}},$$

where $p_{T,ni}$ is the price of tourism services in $i$ for tourists coming from region $n$.

Tourism services are produced under perfect competition by combining local labor and a composite manufacturing input, according to the production function:

$$q_{T,n} = {L_{T,n}}^{\nu_T} {m_{T,n}}^{1-\nu_T},$$

where $L_{T,n}$ is the local workforce working in the tourism industry, $m_{T,n}$ is the manufacturing good used as input in production of tourism services, and $q_{T,n}$ is the quantity of tourism services produced in region $n$. The unit cost of production of tourism-services provided in region $n$ is therefore:

$$c_{T,n} = \Psi_T w_n^{\nu_T} P_{M,n}^{1-\nu_T},$$

where $\Psi_T$ is a constant.[29]

Tourists from region $n$ incur an additional cost $t_{ni}$ when visiting region $i$, which captures travel costs from the region of residence to the region visited, as well as other barriers to tourism (cultural differences between regions, language barrier, duration of travel, etc).[30] Accounting for these frictions, the price of consuming a bundle of tourism-related services for a resident of region $n$ visiting region $i$ is:

[27] In reality, each tourist tends to visit very few regions. As shown in Anderson et al. (1992), the CES assumption made for a representative worker is isomorphic to the aggregation of a continuum of discrete choices made by individual consumers.

[28] Note that, in contrast to the case of manufacturing, we do not allow for the tourism shifter $A_i$ to be endogenous. This assumption is without loss of generality for the estimation of the spillover effects and the aggregate gains from tourism: because the welfare gains from tourism are computed by comparing the current equilibrium to a world with no tourism, it does not matter whether the tourism shifters $A_n$ are endogenous or exogenous. When estimating the gains from international-only tourism, this assumption is conservative as we do not account for the possibility that tourism is subject to scale effects.

[29] Specifically, $\Psi_T = (\nu_T)^{-\nu_T} (1-\nu_T)^{-1+\nu_T}$.

[30] As is common in the trade literature, we do not model explicitly the transportation industry. In this context, it is the industry that provides travel services to tourists between their region of residence and their destination. The impact of the corresponding costs is captured in a reduced-form way in the bilateral tourism frictions $t_{ni}$, that may systematically vary with the distance between $n$ and $i$ as we discuss in Section 5.

$$p_{T,ni} = c_{T,i} t_{ni}.$$

The price index for the bundle of tourism-related services for a resident of region $n$ is therefore:

$$P_{T,n} = \left( \sum_{i \neq n} A_i t_{ni}^{1-\sigma_T} c_{T,i}{}^{1-\sigma_T} \right)^{\frac{1}{1-\sigma_T}}. \tag{7}$$

If follows that the share of region $n$ spending on tourism services that is spent in region $i$ is:

$$\lambda_{ni} = \frac{A_i \left( t_{ni} c_{T,i} \right)^{1-\sigma_T}}{\sum_{k \neq n} A_k \left( t_{nk} c_{T,k} \right)^{1-\sigma_T}}. \tag{8}$$

**Local Non-Traded Services**

Finally, local services are produced and consumed by local residents. They are produced using local labor with constant returns to scale and productivity $R_n$, so that:[31]

$$q_{S,n} = R_n L_{S,n},$$

and

$$p_{S,n} = \frac{w_n}{R_n}.$$

Since $R_n$ is not identified independently from the level of local amenities $B_n$ in what follows, we choose to normalize $R_n = 1$ and interpret $B_n$ as indicating a combination of the level of local amenities and the productivity of the local non-traded services.

## 4.2 Equilibrium

A worker who lives in region $n$ consumes a bundle $C_n = \left( \frac{\left[ C_{M,n}^{\frac{\rho-1}{\rho}} + C_{T,n}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}}{\alpha_T} \right)^{\alpha_{MT}} \left( \frac{C_{S,n}}{\alpha_S} \right)^{\alpha_s}$, so that:

$$C_n = \frac{w_n}{P_{MT,n}^{\alpha_{MT}} w_n^{\alpha_S}}.$$

Mexican workers choose in which region to live within Mexico. Given the properties of the Frechet distribution and the workers' utility maximization problem in (2), the share of Mexican workers who choose to live in region $n \in \mathcal{M}$ is:

$$\frac{L_n}{L_{\mathcal{M}}} = \frac{\left( B_n L_n^{\epsilon} C_n \right)^{\kappa}}{\sum_{k \in \mathcal{M}} \left( B_k L_k^{\epsilon} C_n \right)^{\kappa}}, \text{ for } n \in \mathcal{M}. \tag{9}$$

[31] These services can be interpreted as housing. Formally, modeling housing as in Redding (2015) leads to isomorphic expressions.

This expression leads to another formulation that will prove useful in estimation:

$$\frac{L_n}{L_{\mathcal{M}}} = \frac{\left(B_n \left(\frac{w_n}{P_{MT,n}}\right)^{\alpha_{MT}}\right)^{\tilde{\kappa}}}{\sum_{k \in \mathcal{M}} \left(B_k \left(\frac{w_k}{P_{MT,k}}\right)^{\alpha_{MT}}\right)^{\tilde{\kappa}}}, \text{ for } n \in \mathcal{M}, \tag{10}$$

where we define:

$$\tilde{\kappa} = \frac{\kappa}{1 - \kappa\epsilon}.$$

Outside of Mexico, other countries' populations $L_n$ for $n \in \overline{\mathcal{M}}$ are given exogenously.

Finally, the three market clearing conditions for the manufacturing goods market, the tourism services market and the market for local services lead to the following system of $3 \times N$ wage equations that closes the model.[32] For all regions $i \in (1, ..., N)$:

$$w_i L_{M,i} = \nu_M \sum_{n=1}^{N} \left(\alpha_{MT} w_n L_n \chi_n + \frac{1-\nu_T}{\nu_T} w_n L_{T,n} + \frac{1-\nu_M}{\nu_M} w_n L_{M,n}\right) \pi_{ni} \tag{11}$$

$$w_i L_{T,i} = \nu_T \sum_{n=1}^{N} \alpha_{MT} w_n L_n (1\text{-}\chi_n) \lambda_{ni} \tag{12}$$

$$w_i L_{S,i} = \alpha_S w_i L_i \tag{13}$$

Equations (5)- (8) and (10)-(13) define the equilibrium of the economy.

### 4.3 Welfare Impact of Tourism Integration

The model lends itself naturally to welfare analysis. We use as a measure of welfare in a region the average utility level enjoyed by workers who live in this region. In a given spatial equilibrium, because of the free mobility of workers and the properties of the Frechet distribution, this level of welfare is equalized across all Mexican regions. To quantify how this representative level of welfare is impacted by the development of tourism in Mexico, we need to solve for a counterfactual equilibrium without tourism integration, and compare welfare between the current equilibrium and the counterfactual one.

To this end, we first derive the expression for welfare per capita in Mexico in any given spatial equilibrium. We then derive an expression for how welfare changes between two equilibria with different levels of frictions to tourism. This expression depends on how several endogenous variables change between the two equilibria. Finally, we describe how we solve for these changes in endogenous variables between the current equilibrium and a counterfactual one with prohibitive frictions to tourism travel.

[32] We assume for simplicity that aggregate trade is balanced in Mexico. In the data, Mexico runs a very small trade deficit. The model can be readily adapted to account for this aggregate deficit in the spirit of Dekle et al. (2007) and Caliendo & Parro (2014). We have experimented with this specification, allocating the aggregate deficit to regions in proportion to local GDP, and found that results are very similar when accounting for this deficit.

### Expression for Aggregate Welfare

Given the properties of the Frechet distribution and the workers' utility maximization problem in (2), the common level of welfare in the Mexican economy is:

$$U_{\mathcal{M}} = K_2 \left[ \sum_{k \in \mathcal{M}} \left( B_k L_k^{\epsilon} \left( \frac{w_k}{P_{MT,k}} \right)^{\alpha_{MT}} \right)^{\kappa} \right]^{\frac{1}{\kappa}}. \tag{14}$$

where $K_2 = \Gamma(\frac{\kappa-1}{\kappa})$ is a constant. Welfare is a power mean, across all Mexican regions, of a measure of local utility that includes local real income and local amenities. The parameter $\kappa$ measures how weak are idiosyncratic preferences of workers for different regions. When this parameter is high, regional labor supply responds very elastically to regional differences in amenities-adjusted real income. In turn, aggregate welfare responds strongly to the level of real income in the most attractive region. At the limit when $\kappa \to \infty$, welfare is simply proportional to $\max_{k \in \mathcal{M}} \; B_k L_k^{\epsilon} \left( \frac{w_n}{P_{MT,n}} \right)^{\alpha_T}$.

Combining equations (9) and (14) leads to the following alternative expression for per-capita welfare in Mexico, which holds for any region $n$ in Mexico:

$$U_{\mathcal{M}} = K_3 B_n \left( \frac{w_n}{P_{MT,n}} \right)^{\alpha_{MT}} L_n^{-\frac{1}{\kappa}}, \; \forall n \in \mathcal{M}. \tag{15}$$

where the constant $K_3$ is equal to $\Gamma(\frac{\kappa-1}{\kappa})^{-1} L_{\mathcal{M}}^{\frac{1}{\kappa}}$.

### Expression for Change in Welfare

Using expression (15), it follows that difference in welfare between two equilibria is identical across all regions, irrespective of their level of exposure to tourism, because of free mobility: in the new equilibrium, workers relocate and arbitrage away differences in welfare across regions. The difference in welfare for any region $n$ takes the following expression:

$$\widehat{U_{\mathcal{M}}} = \left( \widehat{\frac{w_n}{P_{MT,n}}} \right)^{\alpha_{MT}} \widehat{L_n}^{-\frac{1}{\kappa}} \; \forall n \in \mathcal{M}.$$

The hat notation indicates percentage changes. For a variable $x$, we write $\widehat{x} = \frac{x'}{x}$, where $x$ the value in the current equilibrium (observed in the data) and $x'$ is the value of the same variable in the counterfactual equilibrium. To evaluate these welfare gains/losses, we need to quantify how $(L_n, w_n, P_{MT,n})$ change when Mexico's level of openness to tourism changes.

To this end, we follow the methodology introduced by Dekle et al. (2007) and generalized to spatial equilibria by Caliendo et al. (2014) and Redding (2015). We first define more precisely the counterfactual equilibrium we use as a reference to estimate the welfare gains from the development of tourism in Mexico, then write the system of equations that allows to solve for the change in welfare in Mexico between the baseline equilibrium and this counterfactual.

**Definition of the Counterfactual Equilibria**

We consider two counterfactual worlds in turn. In the first one, we assume that frictions to international tourism become prohibitive but there is still inter-regional tourism within the borders of Mexico. Compared to the current equilibrium, this corresponds to the following changes in frictions to tourism:

$$\begin{aligned}
\widehat{t_{ni}} &= 1 \text{ if } (n,i) \in \mathcal{M} \times \mathcal{M}, \\
\widehat{t_{ni}} &= 1 \text{ if } (n,i) \in \overline{\mathcal{M}} \times \overline{\mathcal{M}}, \\
\widehat{t_{ni}}^{-1} &= 0 \text{ if } (n,i) \in \overline{\mathcal{M}} \times \mathcal{M}, \\
\widehat{t_{ni}}^{-1} &= 0 \text{ if } (n,i) \in \mathcal{M} \times \overline{\mathcal{M}}.
\end{aligned}$$

All other exogenous parameters are held constant. In the second counterfactual, we investigate the case in which both international tourism and inter-regional travel for Mexicans are shut down. This corresponds to the following changes in frictions to tourism:

$$\begin{aligned}
\widehat{t_{ni}} &= 1 \text{ if } n = i, \\
\widehat{t_{ni}}^{-1} &= 0 \text{ if } n \neq i.
\end{aligned}$$

In both cases, we assume that all other exogenous fundamentals of the economy stay unchanged between equilibria.

**Solving for Counterfactual Changes**

Given the expression for the prices indexes (6) and (7), the expression for trade and tourism shares, (5) and (8), and the expression of manufacturing productivity (4), changes in trade shares and prices between two equilibria are simple functions of changes in wages and local populations working in the services or the manufacturing sector:

$$\widehat{\pi_{ni}} = \frac{\widehat{c_{M,i}}^{-\theta}\widehat{M_i}^{\theta}}{\widehat{P_{M,n}}^{-\theta}} \tag{16}$$

$$\widehat{P_{M,n}}^{-\theta} = \sum_j \pi_{nj}\widehat{c_{M,j}}^{-\theta}\widehat{M_j}^{\theta} \tag{17}$$

$$\widehat{c_{M,i}} = \widehat{w_i}^{\nu_M}\widehat{P_{M,i}}^{1-\nu_M} \tag{18}$$

$$\widehat{M_i} = \widehat{L_{M,i}}^{\gamma_M}\widehat{L_{ST,i}}^{\gamma_S} \tag{19}$$

$$\widehat{\lambda_{ni}} = \frac{\widehat{c_{T,i}}^{1-\sigma_T}\widehat{t_{ni}}^{1-\sigma_T}}{\widehat{P_{T,n}}^{1-\sigma_T}} \tag{20}$$

$$\widehat{P_{T,n}}^{1-\sigma_T} = \sum_j \lambda_{nj}\widehat{t_{nj}}^{1-\sigma_T}\widehat{c_{T,j}}^{1-\sigma_T} \tag{21}$$

$$\widehat{c_{T,i}} = \widehat{w_i}^{\nu_T}\widehat{P_{M,i}}^{1-\nu_T} \tag{22}$$

$$\widehat{\chi_n} = \frac{\widehat{P_{M,n}}^{1-\rho}}{\widehat{P_{MT,n}}^{1-\rho}} \tag{23}$$

$$\widehat{P_{MT,n}}^{1-\rho} = \left((1-\chi_n)\widehat{P_{T,n}}^{1-\rho} + \chi_n\widehat{P_{M,n}}^{1-\rho}\right), \tag{24}$$

Change in local population levels within Mexico stems from the location choice equation (9) together with the maintained assumption that total population is unchanged in the counterfactual equilibrium, i.e. $\sum_{n\in\mathcal{M}} L'_n = L_{\mathcal{M}}$:

$$\widehat{L_n} = \frac{\left(\widehat{w_n}^{\alpha_{MT}}\widehat{PT_n}^{-\alpha_{MT}}\right)^{\tilde{\kappa}}}{\sum \frac{L_i}{L_{\mathcal{M}}}\left(\widehat{w_i}^{\alpha_{MT}}\widehat{PT_i}^{-\alpha_{MT}}\right)^{\tilde{\kappa}}} \ \forall n \in \mathcal{M}. \tag{25}$$

Finally, the system is closed by the market clearing conditions in each sector, that is, equations (11)-(13) expressed in the counterfactual equilibrium, together with:

$$L'_i = L'_{M,i} + L'_{T,i} + L'_{S,i} \tag{26}$$

$$L'_{ST,i} = L'_{T,i} + L'_{S,i} \tag{27}$$

Knowing the values of $(\pi_{nj}, \lambda_{nj}, \chi_n, w_n, L_{M,n}, L_{T,n}, L_{S,n})$ in the baseline equilibrium, the parameters of the model $(\nu_M, \nu_T, \alpha_T, \sigma_T, \theta, \rho, \widetilde{\kappa})$ and the change in frictions $\widehat{t_{ni}}$, the system of equations (16)-(27) defines a counterfactual equilibrium of the economy.[33]

**Role of Local Spillovers**

One of the questions we are after is whether the positive reduced-form effect of tourism on manufacturing reported in Section 3 is indicative of cross-sector productivity spillovers, and whether

[33]In the presence of within and cross-sector spillovers, the uniqueness of the equilibrium is not a priori guaranteed. We come back to this issue when discussing our quantification strategy.

the interplay of within and cross-sector agglomeration economies gives rise to a positive or negative effect of tourism in the aggregate. To gain a better intuition on these questions, let us consider theoretically how local manufacturing productivity changes when the economy moves from a given equilibrium to one with less frictions to trade in tourism (both across regions and borders).

When frictions to tourism decrease, two things happen. First, the spatial distribution of population $\{L_n\}$ changes: labor reallocates towards regions with comparative advantage in tourism. Second, the tourism-related services grow and the share of workers working in manufacturing decreases: on average, $s_{M,n}$ decreases, where $s_{M,n}$ denotes the share of workers of region $n$ working in manufacturing.[34] To see how these two channels impact manufacturing productivity, consider the expression for local productivity in manufacturing. It is given by equation (4), which we can re-write as: $M_n = M_n^o s_{M,n}^{\gamma_M} (1 - s_{M,n})^{\gamma_S} L_n^{\gamma_S+\gamma_M}$. Manufacturing productivity changes locally as a function of (i) the change in local scale of economic activity, captured by the term $L_n^{\gamma_S+\gamma_M}$ (the "agglomeration effect"), and (ii) the change in the composition of economic activity, captured by the term $s_{M,n}^{\gamma_M} (1 - s_{M,n})^{\gamma_S}$ (the "reallocation effect").

The first effect leads to a local boost in manufacturing productivity in regions that attract population, and to muted effects in the aggregate because the total population is fixed.[35] The sectoral reallocation effect driven by $\{s_{M,n}\}_{n \in \mathcal{M}}$ is novel in our approach and arises in the presence of both within and cross-sectoral spillovers. On average, the share $1 - s_{M,n}$ increases when frictions to tourism are lifted. The sign of the derivative of manufacturing productivity $M_n$ with respect to $1 - s_{M,n}$ is given by:

$$\frac{dM_n}{d(1 - s_{M,n})} \propto \frac{\gamma_S}{1 - s_{M,n}} - \frac{\gamma_M}{s_{M,n}} \tag{28}$$

We first examine the case where one of the spillovers is zero, a simple case that is useful to illustrate the forces at play. If $\gamma_S = 0$ but $\gamma_M > 0$, then $\frac{dM_n}{d(1-s_{M,n})}$ is unambiguously negative. Moving towards lower frictions to tourism leads to an increase in $1 - s_M$ on average, therefore a decrease in $M_n$ on average. This would be an amplified Dutch disease effect of tourism: the development of tourism leads to an adverse effect on manufacturing productivity, both locally and in the aggregate. In contrast, when $\gamma_M = 0$ but $\gamma_S > 0$, lowering frictions to tourism leads to an unambiguously beneficial effect on manufacturing productivity as the development of tourism generates spillovers that would not have occurred otherwise.

For non-zero values of $(\gamma_M, \gamma_S)$, the impact of the presence of spillovers on the welfare gains of lowering frictions to tourism is a priori ambiguous. It depends not only on the relative values of $\gamma_S$ and $\gamma_M$, but also on the initial distribution of activity across sectors $s_{M,n}$ in each region. For a given value of spillovers $(\gamma_M, \gamma_S)$, we see that $\frac{dM_n}{d(1-s_{M,n})}$ can be negative for regions that have a low share of the manufacturing sector, and positive for those with high shares. Therefore, if the regions affected by the increase in tourism tend to have relatively low (resp. high) manufacturing

[34]This holds true in all regions in the limiting case where all regions are identical. It holds only on average in the more general case.

[35]Kline & Moretti (2014) show that in a spatial setting with the production of a homogenous good and no trade costs, agglomeration externalities have no aggregate effect.

activity, the spillover effect of tourism, ceteris paribus, is negative (resp. positive). The intuition for this result is that the spillovers display decreasing returns with respect to local concentration: given $(\gamma_M, \gamma_S)$, there can be positive net spillovers from reallocating resources towards services in regions already highly concentrated in manufacturing.

# 5 Calibration

## 5.1 Methodology

To quantify the long-run economic impact of tourism in Mexico, the analysis proceeds in three consecutive steps. In the first step, we calibrate the model to the reference equilibrium corresponding to the current level of trade and tourism frictions, using wages, production, employment and trade data on the one hand, and estimates for the trade elasticity $\theta$ and the tourism elasticity $\sigma_T$ on the other hand. We detail below how we use the data at hand to calibrate the model. Importantly, this step does not require estimating the spatial labor supply elasticity, nor the spillover function. This first step allows us to recover a vector of –possibly endogenous– model-based manufacturing productivities, $M_n$, and a set of local demand shifters for tourism $A_n$.

In the second step, we estimate the reduced-form expression of the spatial labor supply elasticity $\tilde{\kappa} = \frac{\kappa}{1-\kappa\epsilon}$. We derive an estimating equation from the model, and exploit the instrumental variables for local beach quality that we discuss in Section 3 as a set of exogenous shifters to local real wages to identify $\tilde{\kappa}$. This procedure requires an estimate of the local price indices for which sufficiently rich data do not exist in our empirical setting.[36] To circumvent this issue, we use the model-consistent price indices that we calibrate in the first step. Importantly, this step does not require knowledge of the spillover function in traded goods production.

In the third step, armed with these moments and parameters, we investigate the existence of productivity spillovers both within manufacturing and between the local services sector and the manufacturing sector using indirect inference. The approach we follow makes use of the exclusion restrictions that we introduced in the reduced-form analysis. We simulate a counterfactual equilibrium of the Mexican economy without tourism, and then back out the extent of local production externalities that ensure that, absent tourism, the distribution of population and manufacturing productivities are uncorrelated with our instrumental variables. This imposes a strong discipline on how manufacturing productivities, $M_n$, can be decomposed between an exogenous part, independent of the level of economic activity, and an endogenous part, coming from both within and cross-sector production externalities.

## 5.2 Calibration of the Baseline Equilibrium

To calibrate the baseline equilibrium, we follow a methodology close to the one laid out by Redding (2015). Using direct data on $L_{M,n}$, $L_{T,n}$, $L_{S,n}$ and $w_n$, aggregate exports, a parametric specification for the trade and tourism frictions $\tau_{ij}$ and $t_{ij}$, and estimates for the elasticities $\nu_M$, $\nu_T$,

[36] Note that consumption microdata such as the AC Nielsen Mexican consumer panel focus on relatively large cities and would only cover a fraction of coastal municipalities (less than 20).

$\alpha_S$, $\theta$, $\sigma_T$ and $\rho$, we can recover the unique vectors of local manufacturing productivities, $M_n$, and local demand shifters for tourism, $A_n$, for $n \in \mathcal{M}$, that are consistent with the data.[37]

**Data and Calibration**

The model is calibrated to the mean of inflation-adjusted outcomes for 2000 and 2010 as the baseline period.[38] In order to limit the computational requirement, we aggregate the data coming from each of the 2455 Mexican municipalities described in the reduced-form empirical exercise into a set of 300 regions. In particular, we keep the 150 coastal municipalities unchanged, but aggregate the interior municipalities to 150 economic centers that are located at the centroids of the largest 150 interior municipalities (in terms of mean inflation-adjusted GDP in 2000 and 2010). The remaining interior municipalities are assigned to their closest centroid so that all Mexican municipalities are aggregated to 300 regions in the model. Since the empirical moments that we use to inform the calibration of the model and estimate the key parameters are based on variation among coastal municipalities (similar to the reduced-form analysis above), this aggregation is greatly convenient for computational power, but largely inconsequential for our welfare quantification, as the robustness tables A.8 and A.9 show in the online appendix.

We use nominal wage and and local employment data from the Mexican population censuses as our measure of $w_n$ and $L_n$.[39] We calibrate the Cobb-Douglas share of local non-traded services such that aggregate non-traded services GDP matches the Mexican national accounts. This leads to $\alpha_S = .35$, from which we derive our measure of $L_{S,n}$ for all regions. To measure the relative share of tourism and manufacturing employment and production in each region, we measure the relative GDP of these sectors in each region using the economic census data. The manufacturing GDP is observable in all regions. To measure local tourism GDP, we use hotel sales as a basis. We scale these sales with a constant factor of proportion across all regions so that in the aggregate tourism GDP represents 10 percent of total GDP, as observed in the Mexican national accounts in 2000 and 2010.[40] Given our measure of $w_n$, this delivers our measures for $L_{M,n}$ and $L_{T,n}$.

For simplicity, we aggregate all countries but Mexico into a "Rest of the World" ("RoW") aggregate.[41] We calibrate the wage in RoW as the trade-weighted average wage of Mexico's trading partners (measured as GDP per capita), and adjust population of RoW so that the ratio of GDP of Mexico to the GDP of RoW equals the actual ratio between Mexico and the rest of the world. The shares of workers in the manufacturing and tourism industries for RoW are calibrated to the share of world GDP in each sector. Aggregate trade flows for manufacturing and tourism between Mexico and RoW are taken from the the World Bank's WITS database for trade in goods and services.

Turning to trade costs, there is no available data on bilateral trade flows between municipality pairs within Mexico. To calibrate these shares using the available data, we use a parametric

[37] These two vectors are unique up to a multiplicative constant.

[38] Note that the population census data is for 2000 and 2010, while the economic census data is for 1998 and 2008.

[39] To aggregate interior regions, we take the sum of population and the population-weighted mean of wages.

[40] For the few regions for which this procedure predicts an employment in the tourism sector that is higher than the total employment in services reported in this region, we cap tourism employment at the level reported for the services sector as a whole.

[41] Consumers from RoW consume tourism services both in their own country and in Mexico.

specification of the frictions to trade and tourism within Mexico. There is a long literature finding robust evidence that trade flows decay with distance with a constant elasticity close to 1 ((Mayer, 2014)). We therefore parametrize trade fictions within Mexico such that

$$\tau_{nj}^{-\theta} = d_{nj}^{-D_M}, \text{ for } (n, j) \in \mathcal{M} \times \mathcal{M},$$

where $D_M = 1$, where $d_{nj}$ is the distance between the centroid of the two municipalities $n$ and $j$. We follow a similar parametrization for tourism flows. To estimate the distance decay of tourism flows, we run a model-based gravity equation on country-level bilateral tourism exports, taking the log of equation (8) aggregated at the country level:

$$logE_{nkt} = \delta_{nt} + \zeta_{kt} - D_T \log d_{nk} + X_{nk} + \xi_{nkt}, \tag{29}$$

where $E_{nkt}$ is the spending of country $n$ on tourism in country $k$ in year $t$, $\delta_{nt}$ is an country of origin-by-year fixed effect, $\zeta_{kt}$ is a destination country-by-year fixed effect, $X_{nk}$ are three dummy variables capturing other sources of frictions to international tourism besides distance, such as whether countries share a common border, language or colonial ties. Finally, $d_{nk}$ is our regressor of interest, the distance between country $n$ and $k$. Specification (29) leads to a tightly estimated distance decay elasticity of $D_T = 1.46$ (standard error 0.0364).[42] As can be seen in Figure 2 subject to a much less parametric specification than in (29), tourism flows strongly follow gravity, and the log-linear specification fits the data extremely well. We proceed to parametrize tourism fictions within Mexico as varying with distance, such that:

$$t_{nj}^{1-\sigma} = d_{nj}^{-D_T}, \text{ for } (n, j) \in \mathcal{M} \times \mathcal{M}.$$

Following the literature, we assume that international frictions are subject to an additional sector-specific border effect, such that the frictions between region $n$ in Mexico and RoW is $d_{nj}^{-D_M} \tau_{Border}$ for manufacturing, and $d_{nj}^{-D_T} t_{Border}$ for tourism.[43] The two border frictions are calibrated such that the model matches exactly the aggregate trade data in manufactured goods and tourism between Mexico and the rest of the world.

**Elasticity Estimates**

We present here how we estimate or calibrate the model elasticities $\nu_M$, $\nu_T$, $\theta$, $\sigma_T$ and $\rho$, which are required to back out the vector of productivities and tourism demand shifters conditional on data. Note that, as a robustness check, we also revisit our counterfactual analysis in Section 6 using a range of alternative parameter combinations for the trade and tourism elasticities $\theta$, $\sigma_T$ and $\rho$.

We calibrate the traded input shares of manufacturing and tourism services to match the total requirement coefficients of the 2003 Mexican input-output table. These correspond to the labor shares $\nu_M = 0.55$ and $\nu_T = 0.8$. For the baseline value of the trade elasticity for flows of goods, we

[42] Standard errors are clustered at the level of origin-destination pairs.

[43] Since the US and Canada account for the vast majority of both trade in goods and tourism in Mexico, we compute the distance between Mexican regions and the RoW region based on bilateral distances to a central point in the US: Denver City.

use an existing estimate from the literature of $\theta = 6.1$ (Adao et al., 2015).

To estimate the tourism trade elasticity $\sigma_T$, we use the panel data on country-level bilateral tourism exports. Equation (8) leads to the following estimation equation:

$$log E_{nkt} = \delta_{nt} + \zeta_{nk} + (1 - \sigma_T)\, log\, w_{kt} + \xi_{nkt}, \tag{30}$$

where $E_{nkt}$ is the spending of country $n$ on tourism in country $k$ at period $t$, $\delta_{nt}$ is an origin-by-time fixed effect (e.g. capturing productivity shocks), $\zeta_{nk}$ is an origin-by-destination fixed effect (e.g. capturing distances or cultural proximity), $ln\, w_{kt}$ is the relative consumption price of tourism services across destinations, and $\xi_{nkt}$ is a mean zero error term. To empirically measure $log\, w_{kt}$, we use country-level PPP rates for final consumption goods that the International Price Comparison (ICP) program computes for all 115 countries over the period 1990-2011 in our database. The ICP constructs this measure, $PPP_{kt}$ as the number of units of a country $k$'s currency required to buy the same basket of goods and services in $k$'s domestic market as one US Dollar would buy in the United States. To measure $log\, w_{kt}$, we take the log of $(1/PPP_{kt})$. Given the inclusion of origin-by-period and origin-by-destination fixed effects, this measure effectively captures (with some error) relative consumption price changes across different destination countries from the point of view of a given origin-by-time cell.

The main concern for the identification of $(1 - \sigma_T)$ is that changes in consumption prices across destinations are correlated with other factors that may increase or decrease bilateral tourism flows in the error term. For example, if prices in a destination increase at the same time that travelers at the origin become more likely to travel to the destination for other reasons (e.g. due to business travel or attractiveness), this would lead to an upward biased estimate of $(1 - \sigma_T)$ (towards zero).

The first step we take to address this concern is to condition on a basic control for time changing economic conditions in the destination countries, by including the log of country GDP as a control. Second, to address remaining concerns, in addition to the very likely concern of measurement error in our measure of $log\, w_{kt}$, we use nominal exchange rate changes across destination countries with respect to the US Dollar, $log\, e_{kt}$, as an instrumental variable for $log\,(1/PPP_{kt})$. The exchange rates are used as part of the PPP rate construction by the ICP, so that we can expect a strong first stage. The exclusion restriction is that differential exchange rate changes across different destination markets to not affect bilateral tourism expenditure except through relative price changes of tourism services, conditional on the included fixed effects and destination-specific changes in log GDP.

To further assess the validity of this assumption, we also estimate specification (30) after restricting attention to what we label touristic destinations: i.e. destinations for which more than 80 percent of total travel inflows are due to leisure rather than business travel. Finally, to allow for tourist flows to respond to relative price information across destination markets with some time lag, we also estimate specifications in which we lag the independent variable by 1-5 years.

Table 10 presents the estimation results. We find a negative and statistically significant tourism trade elasticity that reaches $\sigma_T = 1.7$ when we lag the relative destination price changes by 3-4 year (1.5 with lesser lags). These results are confirmed with a slightly lower point estimate of 1.6 once we restrict attention to destinations with more than 80 of travel inflows driven by leisure rather than business purposes. These results indicate that the tourism trade elasticity appears to be significantly lower than common estimates of the trade elasticity for flows of goods. Finally, given the structure of the model the upper-nest elasticity that governs the substitution patterns between manufacturing and tourism has to be smaller than the lower nest elasticity of $\sigma_T = 1.7$. To be conservative in our quantification of the gains from tourism, we pick the highest estimate of the tourism trade elasticity $\sigma_T = 1.7$ reported in Table 10 as our baseline estimate, and in a similar vein we choose the value of $\rho = \sigma_T = 1.7$. As discussed above, we also report results of the counterfactual analysis across a range of different parametrizations in Section 6.

**First Step: Regional Productivity Levels and Tourism Shifters**

Armed with the calibrated values for $w_n$, $L_{M,n}$, $L_{T,n}$, $L_{S,n}$, $d_{nm}^{-D_T}$ and $d_{nm}^{-D_M}$, values for aggregate trade and tourism flows between Mexico and RoW, and estimates for the elasticities $\nu_M$, $\nu_T$,$\alpha_S$, $\sigma_T$, $\theta$ and $\rho$ discussed above, we use numerical methods to recover the unique (up to a normalization factor) vectors of local manufacturing productivities ($M_n$) and local demand shifters for tourism ($A_n$), for $n \in \mathcal{M}$, as well as border effects $\tau^{Border}$ and $t^{Border}$ that are consistent with the data. Specifically, given the observed distribution of economic activity $(w_n,\ L_{M,n},\ L_{T,n},\ L_{S,n})$ and aggregate trade flows that are fed into the model, these parameters ensure that equations (5), (6), (8), (7), (11), (12) and (13) hold simultaneously –i.e. that the labor and the goods markets clear.

Two remarks are in order here. First, note that knowledge of the labor supply elasticity $\widetilde{\kappa}$ is not needed for this step, because wages and the distribution of population are taken directly from the data, so that equation (9) is not used at this stage of the quantification. It will be required, however, to compute counterfactual equilibria. Second, we mentioned above (and will discuss further below) that in the presence of spillovers there is a potential for multiple equilibria in the model. Conditional on the data we observe, though, the mapping to unobserved productivities and tourism shifters is unique.

**Second Step: Spatial Labor Supply Elasticity**

The estimation equation for the long-run spatial labor supply elasticity is derived from Equation (9) of the model. It leads to the following estimating equation:

$$log\, L_n = K_o + \widetilde{\kappa}\, log \left( \left( \frac{w_n}{P_{MT,n}} \right)^{\alpha_{MT}} \right) + \xi_n \text{ for } n \in \mathcal{M}\ . \tag{31}$$

While wages are part of our data, price indices are not. However, $P_{MT,n}$ is pinned down in the calibrated model as described in step 1. We use these to infer the vector of local real incomes, $\left(\frac{w_n}{P_{MT,n}}\right)^{\alpha_{MT}}$.[44] Note that the spatial labor supply elasticity $\kappa$ is not identified separately from the amenity externality $\epsilon$. For our purposes, and in particular for welfare computations, we do not require a separate estimate of these two elasticities, but only an estimate of the observable reduced-

[44] Recall from the model that non-traded services do not show up in the local price indices. As noted above, since $R_n$ (productivity in non-traded services) is not identified independently from the level of local amenities $B_n$, we normalize $R_n = 1$ and interpret $B_n$ as indicating a combination of the level of local amenities and the productivity of the local non-traded services.

form expression $\widetilde{\kappa} = \frac{\kappa}{1-\kappa\epsilon}$.

The main concern to causally identify $\widetilde{\kappa}$ in specification (31) is that our measure of real wages is correlated with unobserved local amenities $B_n$ in the error term (that due to our normalization of $R_n = 1$ include both differences in amenities as well as in productivities of the local non-traded services sector). The second major concern is measurement error in our construction of local real wages across municipalities, which would again lead to a downward bias in our estimate of $\widetilde{\kappa}$. To address these concerns, we exploit the identifying assumption that our measures of local beach quality only affect local employment through either their effect on real wages (nominal wages and local price indices), and use them as instruments for $log\left(\frac{w_n}{P_{MT,n}}\right)^{\alpha_{MT}}$ in equation (31). Following the empirical strategy in Section 3, we also include the full set of pre-determined regional controls.[45]

Notice that this empirical strategy in principle allows for tourism to affect local amenities endogenously: through the resulting changes in local populations. This is captured as part of our reduced form labor supply elasticity, $\frac{\kappa}{1-\kappa\epsilon}$, where $\epsilon$ captures the change in local amenities due to changes in the local population density. The key remaining concern is that our instrumental variables may be correlated with local amenities, $B_n$, which is a point that connects to our discussion of the exogeneity of the instrumental variables in Section 3. In particular, Tables 5 and 6 provide some reassurance against this concern.

Table 11 presents the estimation results. We estimate significant positive point estimates for $\widetilde{\kappa}$ in both OLS and the IV specifications. The point estimates for $\widetilde{\kappa}$ is significantly larger in the IV specifications relative to the OLS specification. This is consistent with both downward bias due to unobserved local amenities (that in spatial equilibrium should be negatively correlated to local real wages) as well as significant measurement error in our real wage measures. The IV point estimates suggest that the long-run reduced-form elasticity $\widetilde{\kappa}$ is about 7.8 for employment.[46] This elasticity captures how much, in the long-run, the population reallocates in response to a long-run change in relative real wages brought about by tourism. Our estimate indicates that even in this long-term perspective, there are significant frictions to mobility –the elasticity is far from infinite. On the other hand, it is markedly higher than estimates of the short-run spatial labor supply elasticity that are often estimated to be quite low.[47] In our counterfactual analysis in Section 6, we use $\widetilde{\kappa} = 7.8$ as our baseline parameter, and then explore the sensitivity of our findings across a range of alternative parameter combinations.

Finally, as discussed in Section 3, we use the above empirical strategy to construct the regional amenity measures used as the outcome variables in the model-based robustness regressions in Table 6. In particular, we construct five different vectors of regional amenities. Each of them is computed as the residual variation in local population that is left unexplained by variation in real

[45]In model-based regressions, we take the mean of municipality-level control variables to aggregate the interior regions.

[46]We use the second-stage IV estimate when using all five instruments due to the relatively weak first stage when only using the first two IVs, as reported in Table 11. As noted in Section 3, the sixth IV (5th beach IV) is a linear combination of the first four IVs and would thus drop out of the estimation.

[47]See for example Fajgelbaum et al. (2015) for a discussion of the estimates of the labor supply elasticity in the literature.

wages (i.e. the residual in specification (31)). We construct this variable five different times using specification (31) in order to exclude each of the five instruments separately when estimating $\widetilde{\kappa}$. As discussed in Section 3, this ensures that we do not build in a mechanical orthogonality condition between local amenities and our instruments when testing whether or not our instruments are correlated with the model-based measures of the local amenities of residents.

**Third Step: Agglomeration Forces**

At this stage, we have a fully calibrated model except for one important remaining question. Local goods production is potentially subject to within and cross-sector production externalities. The development of tourism may have indirect long-run effects on manufacturing production by affecting the scale of local production across different sectors. We have modeled this possibility in a reduced-form way by assuming that productivity in manufacturing may depend on the scale of local economic activity in manufacturing and services, through the productivity function (4). In the calibration of the baseline equilibrium in step 1 above, we have recovered the vector of (endogenous) local manufacturing productivities $M_n$. To fully characterize the effect of tourism on long-run economic outcomes, we require estimates of the within and cross-sector spillovers on traded goods production $(\gamma_M, \gamma_S)$. To do so, we propose a strategy based on indirect inference that makes use of our instrumental variable strategy. Taking logs of equation (4) provides us with the following equation:

$$\log M_n = \gamma_M \log L_{M,n} + \gamma_S \log L_{ST,n} + \log M_n^o.$$

Note that the distribution of populations $(L_{M,n}, L_{ST,n})$ is endogenous to the level of local manufacturing productivity $M_n^o$. Fortunately, the instrumental variable strategy developed in Section 3 provides us with a series of instruments that impact local economic outcomes only through their impact on local tourism. In particular, the instruments are uncorrelated with the exogenous part of manufacturing productivity $M_n^o$. We formulate these exclusion restrictions in the form of moment conditions:

$$E\left[z_n^{(j)} \log M_n^o\right] = 0, \tag{32}$$

where $z_n^{(j)}$ for $j = 1...5$ denotes our five instrumental variables detailed in Section 3. Furthermore, under the identifying assumption discussed in Section 3, the instruments do not affect the distribution of economic activity directly beyond the effect they have through tourism. In particular, they would not affect the distribution of population in the absence of tourism. We can therefore exploit an additional set of moment conditions for the estimation of $\gamma_M$ and $\gamma_S$ (one for each instrument) that are based on the counterfactual distribution of population under prohibitive frictions to tourism trade in Mexico, denoted by $\left\{\widetilde{L_n}\right\}_{n\in\mathcal{M}}$:

$$E\left[z_n^{(j)}\widetilde{L_n}\right] = 0, \text{ for } j = 1, ..., 5. \tag{33}$$

We make use of these two sets of moments to estimate $(\gamma_M, \gamma_S)$ using model-based indirect inference. That is, we identify the spillover parameter values such that the corresponding moment conditions hold as close as possible within the model. Intuitively, this approach is based on identifying the parameter combination $(\gamma_M, \gamma_S)$ such that the otherwise fully calibrated spatial equilibrium model is capable to match the change from the zero correlations between the beach instruments and local productivity or population in the counterfactual equilibrium without tourism to the observed reduced-form correlations in the equilibrium we observe today. Using only the first set of moment conditions in (31) to estimate $(\gamma_M, \gamma_S)$ would be equivalent to directly estimating $(\gamma_M, \gamma_S)$ using today's cross-section of regional outcomes in an IV regression of $log\,(M_n)$ on $log\,(L_{M,n})$ and $log\,(L_{ST,n})$, with the identifying assumption that $M_n^o$ in the error term of that regression is unrelated to the IVs. The additional set of moments (33) exploits the fact that the distribution of population is a non-linear function of the distribution of exogenous productivity and other primitives of the model. It provides the estimation with additional information to pin down $(\gamma_M, \gamma_S)$.

We implement this approach as follows. For any candidate value of the agglomeration parameters $(\gamma_M, \gamma_S)$, we use the system of equations (16)-(27) to compute numerically the corresponding vector of exogenous manufacturing productivities, $M_n^o(\gamma_M, \gamma_S)$, as well as the distribution of population in Mexico in a counterfactual equilibrium with prohibitive tourism frictions, $\widetilde{L_n}(\gamma_M, \gamma_S)$, while holding all other exogenous parameters constant.[48] We then compute the correlation of these vectors with our set of instrumental variables, conditional on the full set of controls used in Section 3. Specifically, we estimate the following regressions within the model across alternative parameter combinations of $(\gamma_M, \gamma_S)$:

$$log\,(y_{nc}) = \alpha_c^{(j)} + \beta_y^{(j)} z_{nc}^{(j)} + \alpha^{(j)\prime} X_{nc} + u_{nc}^{(j)}, \tag{34}$$

for each of the instruments $j \in 1...5$, and for $y = \widetilde{L_n}(\gamma_M, \gamma_S)$ and $y = M_n^o(\gamma_M, \gamma_S)$ respectively. The vector $X_{nc}$ is the full vector of pre-determined controls described in (1). The indirect inference procedure finds the combination of spillover parameters such that:

$$(\widehat{\gamma_M}, \widehat{\gamma_S}) = argmin\ \boldsymbol{\beta}(\gamma_M, \gamma_S)'\mathbf{W}\boldsymbol{\beta}(\gamma_M, \gamma_S),$$

where $\boldsymbol{\beta}(\gamma_M, \gamma_S)$ is the vector of $\beta_y^{(j)}(\gamma_M, \gamma_S)$ for j=1...5 and $y = \{\widetilde{L_n}, M_n^o\}$, and $\mathbf{W}$ is a weighting matrix for which we use the inverse of the variance of the point estimates of each $\beta_y^{(j)}$ in equation (34).

The results of this procedure are reported graphically in Figure 3, where the loss function is given by $\mathcal{L} = \boldsymbol{\beta}(\gamma_M, \gamma_S)'\mathbf{W}\boldsymbol{\beta}(\gamma_M, \gamma_S)$. The best fitting combination of parameters to match our moment conditions is $\widehat{\gamma_M} = .084$ (with a standard error of 0.034) and $\widehat{\gamma_S} = .088$ (with a standard

[48] As discussed above, there is the potential for multiple equilibria in the model in the presence of spillovers. We choose to implement the following equilibrium selection rule. We solve for the closest counterfactual equilibrium compared to the baseline. That is, we use the values of the endogenous variables from the current equilibrium as a starting point for the counterfactual equilibrium. The procedure then updates the candidate value of endogenous variables in the counterfactual equilibrium based on a weighted average of this initial guess and the new values that come out of solving the model. The procedure is iterated until new values and initial values converge.

error of 0.011).[49]

Figure 3 presents a bowl-shaped graph suggesting that our indirect inference approach identifies a unique parameter combination that minimizes the loss function.[50] Note, however, that the loss function is flatter in the direction of $\gamma_M$ than it is in the direction of $\gamma_S$: the cross-sectoral spillover parameter is more precisely identified by the procedure, whereas the within-sector parameter has somewhat wider confidence intervals. Reassuringly, the value of this within-sector spillover is well within the range of existing measures of agglomeration externalities found in the literature (see e.g. (Rosenthal & Strange, 2004) for a review). Our estimated cross-sectoral spillover is of the same order of magnitude, but there are no existing references in the literature to compare this that we are aware of. As we do for the elasticities $\sigma_T$, $\rho$ and $\widetilde{\kappa}$ discussed above, we also test the the sensitivity of our welfare results to alternative parameter values for $\gamma_M$ and $\gamma_S$ as a robustness check at the end of the following section.

To summarize, we find evidence of both within and cross-sector agglomeration forces. This evidence is model-based in the following sense: to rationalize the strong positive causal impact of tourism on regional employment and production that we document in Section 3, the calibrated model demands a combination of both within-manufacturing as well well as services-to-manufacturing co-agglomeration economies. Importantly, this finding is net of potentially confounding neoclassical general equilibrium price effects on the local economy and, in particular, on traded goods production, which our estimation accounts for. The important remaining question concerns the aggregate implications of these findings, to which we turn in the following section.

# 6 Quantification

This section proceeds to the quantification of the long-run economic consequences of tourism in Mexico. We first present our baseline results for the aggregate welfare gains from tourism, and explore the underlying channels. Second, we investigate what the local welfare effects of tourism would have been in the absence of migration across Mexican regions. Finally, we report a number of additional quantification results to document the sensitivity across alternative modeling assumptions and parameter combinations.

## 6.1 Gains from Tourism

To evaluate the welfare gains from tourism integration in Mexico, we compare the level of welfare in two counterfactual equilibria to the one in the baseline current-day equilibrium. The counterfactual equilibria we consider are ones in which the extent of the development of tourism is limited compared to today's equilibrium. We examine two cases. In the first one, Mexico is closed to international tourism but within-country tourism is still present. In the second one, tourism frictions are prohibitive both across regions within Mexico and internationally. The difference in

[49] To get the standard errors we bootstrap the procedure accounting for sampling error as discussed below in Section 6.

[50] To minimize computing power requirements, we look for the parameter combination that delivers the best fit over a grid of possible values for $\gamma_M$ and $\gamma_S$ ranging from 0 to .25. After inspection of Figure 3, it is clear that the best-fitting parameter combination is insensitive to extending the grid space to larger (less realistic) values.

aggregate per capita welfare between the counterfactual equilibria and the baseline equilibrium captures the welfare loss from going back to a world without tourism (resp. without international tourism). For clarity of presentation, we refer to the inverse of these numbers as the gains from tourism.

Table 12 presents the quantification results. In the first column, we report the gains from tourism integration. In the second row, we focus on the gains from international tourism. The reported confidence intervals of these estimates account for sampling error in the parameter estimates that enter the first step of the model calibration (see Section 5). To obtain these confidence intervals, we bootstrap the whole quantification exercise (steps 1 to 3 above) 200 times. In each bootstrap, we draw the parameters for both the trade and tourism trade elasticities $\theta$ and $\sigma_T$ from a normal distribution with a mean equal to the point estimate and a standard deviation equal to the standard error of the estimate.[51] And as discussed above, for each draw of $\sigma_T$, we also adjust $\rho = \sigma_T$.

The estimated gains from the development of tourism amount to about 4.4 percent welfare gains per capita. The development of international tourism contributes about 36 percent of these gains (1.6 percent), with the remainder stemming from the gains of inter-regional tourism within Mexico. Table 12 also decomposes these welfare results into the neoclassical gains from lowering trade frictions in tourism, and those due to agglomeration economies. Interestingly, while the spillovers lead to large regional re-allocations of production in Mexico (rationalizing the observed reduced-form effects), their aggregate effect on Mexican welfare is relatively muted. In the absence of spillovers, the welfare gains from tourism development would have amounted to 4.2 percent, close to the 4.4 percent welfare gains that we estimate in the presence of spillovers.

This is reminiscent of the finding in Kline & Moretti (2014) that local agglomeration effects may be offset in the aggregate. They show that this is the case in the US, using a single-sector model, when agglomeration effects increase with constant elasticity with respect to local employment. Note, however, that in our setting this result is not purely driven by the log-linear functional form. Instead, it derives from the combination of within and cross-sector agglomeration forces that best fit the data. Despite the log-linear functional forms, both the local and the aggregate implications of tourism can a priori differ substantially in our framework, depending on the combination of $(\gamma_M, \gamma_S)$.

To illustrate this point, we can explore the counterfactual analysis subject to different agglomeration and co-agglomeration forces in the model. Table 13 reports what would have been the welfare gains from tourism integration for alternative spillover scenarios, and illustrates the impact that tourism would have had on relative local outcomes in these cases in the reduced-form regression analysis. In particular, the table reports the estimated gains from tourism alongside the

[51] This is a parametric bootstrap (e.g. Horowitz, 2001) that implicitly assumes errors are uncorrelated across datasets. The standard error associated to the point estimate 6.1 of the trade-in-goods elasticity is 1.046 as reported in Table 1 of Adao et al. (2015). The standard error for the 1.7 point estimate of the tourism trade elasticity is 0.281 as reported in Table 10. The bootstrap confidence intervals do not take into account cases of degenerate equilibria where in excess of half of Mexico's total population concentrates in just one region in the no-tourism equilibrium. For reference, the largest regional share of population in today's observed equilibrium with 300 regions is less than 5 percent.

estimates of the following regressions:

$$\Delta \widetilde{logGDP}_n^j = \alpha_{coast}^j + \beta_1^j logGDPTourism_n + \beta^{j'} X_n + \epsilon_n^j, \tag{35}$$

where the left-hand side measures model-based long-run regional changes in GDP when moving from a no-tourism counterfactual equilibrium to today's spatial equilibrium. Each different parametrization of the agglomeration economies $(\gamma_M, \gamma_S)$ that we index by $j$ here yields a different cross-section of regional changes in local GDP on the left-hand side. On the right-hand side, we replicate the regression specification from (1), and instrument for local tourism GDP in today's equilibrium (which is equal to the counterfactual change in local tourism GDP in each of the $j$ counterfactuals) with the island IV and the beach IV as in the reduced-form analysis. These regressions thus allow us to explore how the effect of cross-sectional variation in tourism activity on local GDP would have differed across alternative spillover scenarios.

In these model-based regressions, the IV approach still addresses the same types of concerns as in the reduced-form analysis: in the model, local variation in tourism activity is partly driven by regional differences in tourism attractiveness (captured by $A_n$). These local tourism shifters may be correlated with other local advantages, such as the $M_n$ and $B_n$, and in addition tourists incur a travel cost so that variation in tourism is also correlated with local market access. To address these confounding factors in (35), we use the island and beach IVs under the same identifying assumption as before: that these form part of the variation in $A_n$ that is unrelated to other determinants of local economic activity.

A first scenario that we explore in column 1 is without any form of agglomeration economies. In that case, tourism has an effect on local GDP that is about half the size of the effect we observe in the reduced-form analysis (0.24 vs 0.44). The reason for this is that in this scenario the effect of tourism on local manufacturing is actually negative (-0.12 with standard error of 0.0293), which underlies the significantly smaller positive effect on local total GDP. What these results suggest is that, in absence of agglomeration economies, increases in local market access to consumers and input demand from tourism through neoclassical channels are insufficient to overturn the adverse effect on local manufacturing through higher factor prices.

A second interesting focal case is the situation in which we shut down spillovers going from tourism to manufacturing ($\gamma_S = 0$), but allow for strong agglomeration economies within manufacturing ($\gamma_M = 0.15$). The corresponding results are reported in column 2. In this case, the development of tourism barely leads to an increase in local GDP. The adverse local effect of tourism on manufacturing is now reinforced by the presence of within-sector agglomeration externalities in manufacturing. The overall welfare gains from tourism are reduced from 4.5 to 1.1 percent compared to what the classical gains from integration would have been in the absence of production externalities in column 1. This adverse effect of tourism illustrates the case when tourism acts as a Dutch disease that is amplified by agglomeration forces. Resources are reallocated away from traded goods production to tourism. Due to localization economies within manufacturing, this has negative implications for productivity in traded goods production as tourism reduces resources allocated to the manufacturing sector.

Column 3 reports the polar opposite case where only cross-sector spillovers are at play. In this case, the development of tourism has a strong positive effect on local manufacturing GDP. In turn, this leads to a strong positive effect on total local GDP, significantly overshooting the effect in the reduced-form analysis and our preferred parametrization in column 4. This larger positive impact stems from a productivity effect: manufacturing now becomes more productive where tourism develops. In the aggregate, this leads to additional welfare benefits of the development of tourism, as it leads to a growth in manufacturing productivity that would not have otherwise occurred. Quantitatively, the welfare gains from tourism in this case are about 1.6 times higher than the classical gains from market integration without spillover effects in column 1.

Column 4 reports the impact of tourism on local GDP and aggregate welfare for the best fitting parameter values ($\widehat{\gamma_M} = .084$ and $\widehat{\gamma_S} = .088$) that we estimated in the previous section. This leads to an effect of tourism on local total GDP that is close to identical to what we observe in the reduced-form analysis (0.41 vs 0.44). Comparing this result to the first column (no spillovers) provides an interesting decomposition of the observed local impact of tourism in the long run: slightly more than half of this effect (0.24) is driven by purely neoclassical channels, including the mechanical increase in local GDP due to tourism revenues in addition to general equilibrium local price effects on other sectors. The remainder of tourism's observed effect on local GDP is driven by gains in local manufacturing activity through the combination of both within and cross-sector agglomeration economies.

These results are informative about the underlying channels of the gains from tourism that we estimate for the Mexican case. Spillovers from the services sector to manufacturing are strong enough to generate co-agglomeration, as attested by the positive impact of tourism on local manufacturing production. On the other hand, manufacturing is also estimated to be subject to within-sector agglomeration economies. While these two forces work in parallel to rationalize the large observed re-allocations of GDP and employment towards touristic regions, their combination also implies that the aggregate welfare effect of tourism on manufacturing productivity is muted: the aggregate gains from tourism in Mexico are only slightly higher than what would have prevailed without spillovers due to reductions in agglomeration economies among less touristic regions in Mexico.

Finally, as shown in Table 12, an interesting contrast to these findings emerges when we focus on the gains from international-only tourism. We find that the welfare gains brought about by international tourism are dampened compared to what they would have been in the absence of spillovers. This asymmetry in the role of the agglomeration forces between the gains from tourism as a whole and the gains from international tourism relate to our discussion above in Section 4.3. In the case of international tourism, the regions most impacted have on average a smaller share of manufacturing than the average regions impacted by domestic tourism across Mexican regions. Because of this, the reallocation of resources towards the services sector, and away from the manufacturing sector, generates a disproportional reduction in the scale effect for manufacturing compared to the gain in spillovers brought about by the services expansion. The result is that the estimated gains from international tourism are lower than the gains that would have occurred in the absence of agglomeration economies (1.6 vs 2.4 percent).

### 6.2 Local Gains from Tourism Without Migration

A second question is what the local welfare effect of tourism would have been in absence of a long-term spatial equilibrium with labor mobility. To address this question, we first solve for the counterfactual spatial equilibrium with prohibitive frictions to both domestic and international tourism in Mexico. Starting from this initial equilibrium, we then simulate the new equilibrium that arises when lowering the tourism travel frictions to today's level for either both domestic and international or international-only tourism, but now under the assumption that labor is immobile across regions within Mexico. We thus effectively shut down the economic geography dimension of the model, and evaluate the local welfare implications of tourism in a world with trade in goods and tourism-related services. All other forces in the model, such as input-output linkages and agglomeration economies, are held constant at their baseline parameters, but the model no longer allows workers to choose their region of residence, so that expected real incomes are no longer equalized across regions.

Table 14 reports the counterfactual effect of regional variation in tourism activity on local worker welfare. In particular, the table replaces the left-hand side in specification (35) by the log change in worker utility when moving from the no-tourism equilibrium to the current level of tourism trade frictions. We regress this variable on the counterfactual change in local tourism GDP (which in specification (35) was equal to today's level of tourism GDP, but this is no longer the case in absence of mobility) in addition to the full set of controls. As in the previous subsection, we instrument for the change in local tourism activity with our island and beach IVs.

We find that a 10 percent increase in local tourism activity causes a 1.7 percent increase in local worker welfare in the absence of immigration. When focusing on international-only tourism, we find almost the same point estimate (0.17 with standard error of 0.0494). In principle, there are several factors that could lead to differences in the local welfare elasticity with respect to international and domestic tourism activity. As we have discussed above, international and domestic tourism are concentrated in different regions of Mexico. For example, the initial sectoral composition of the local economies could differ when hit by the tourism shock, and this could lead to heterogeneous local welfare effects due to the presence of both within and cross-sector agglomeration forces. In practice, however, the local welfare effects of the two counterfactuals turn out to be very similar in terms of proportional changes.

In summary, we estimate large and significant local welfare gains of tourism that would have occurred in the absence of regional migration. These local welfare gains are the model-based counterpart of the strong migration responses to local variation in tourism that we have documented in the reduced-form analysis in Section 1.

### 6.3 Robustness

In the final section we address two remaining questions. We first explore the sensitivity of the estimated co-agglomeration forces and gains from tourism after allowing for the possibility that tourism development causes an endogenous reduction in transport costs for trading goods due to infrastructure investments among touristic regions. Second, we report our estimates of the gains from tourism and international tourism across a range of alternative parameter values for the tourism trade elasticity, the spatial labor supply elasticity, and the extent of within and cross-sector agglomeration forces.

**Endogenous Reduction in Transport Costs**

One of the channels through which the development of tourism can have a positive effect on the local manufacturing sector is through reductions in the cost of trading goods due to transportation infrastructure. As discussed at the end of Section 3, while the reduced-form effects of tourism on local economic outcomes appear to be remarkably robust to the inclusion of infrastructure controls, we did find suggestive evidence that the positive effects are somewhat reduced compared to the baseline results. In the context of our quantitative analysis, this could give rise to concern that part of the positive effect on traded goods production is driven by an endogenous reduction in transport costs. In turn, this could lead to over-stated welfare gains since the estimation of the cross-sector externality ($\gamma_S$) could be upward biased due to this omitted increase in local market access.

To get a sense of the sensitivity of our results to this concern, we re-evaluate the counterfactual analysis after allowing for the possibility that tourism leads to a substantial reduction in the costs of trading goods among touristic regions. In particular, we introduce the assumption that the development of tourism brings about a 50 percent reduction in the trade costs among the top 20 percent of touristic municipalities along the coastline for all their bilateral trading partners (domestic regions and RoW). That is, the tourism centers that we observe in Mexico today experience an endogenous increase in their transport costs as we move from today's spatial equilibrium to the counterfactual equilibrium in the absence of tourism. We then implement the exact same counterfactual quantification as above (taking as given steps 1 and 2) and, re-estimate the spillover parameters (step 3) in the same way that we do in our main specification. The only difference is that the counterfactual equilibrium outcomes used in the estimation of $\gamma_S$ and $\gamma_M$ and in the welfare quantification now also account for the hypothesized difference in transport costs brought about by tourism.

As reported in Table 15, we find estimates of $\gamma_S = 0.080$ and $\gamma_M = 0.090$ compared to $\gamma_S = 0.088$ and $\gamma_M = 0.084$ in our baseline estimation. As expected, the cross-sector co-agglomeration force ($\gamma_S$) is somewhat weaker and the within-sector spillover ($\gamma_M$) somewhat higher than in our baseline specification. However, the magnitude of these changes is minor. As a result, Table 15 reports that the welfare gains from tourism are virtually unchanged compared to our main specification (4.51 vs 4.4 percent, and 1.64 vs 1.6 percent for international tourism). In line with our reduced-form evidence on the robustness of the local effects to the inclusion of various infrastructure controls, the lack of sensitivity of our model-based spillover and welfare results provides some re-assurance that our findings are unlikely to be biased upwards due the omission of potentially endogenous reductions in trade costs due to tourism development.

**Alternative Parameter Values**

The final question we turn to is to what extent our findings are sensitive to different assumptions about some of the key parameters determining the size of the estimated gains from tourism. In particular, Table 16 reports the estimated gains from tourism as well as from international-only tourism across different parameter combinations for the trade elasticity of tourism ($\sigma_T$), the spatial labor supply elasticity ($\widetilde{\kappa}$) and the cross-sector co-agglomeration force $\gamma_S$. All other parameters are held constant at their values of our baseline calibration discussed above.

First, the tourism trade elasticity ($\sigma_T$) directly affects the magnitude of the estimated neoclassical gains from lower frictions to tourism trade. In particular, a larger tourism trade elasticity implies a lower gain from trade in tourism for a given set of empirical moments. This is analogous to the role of the trade-in-goods elasticity in the recent quantitative literature on the gains from trade (Arkolakis et al., 2012). Intuitively, moving from the observed level of tourism consumption to tourism autarky implies a larger loss in welfare if the demand elasticity of tourism consumption is lower (less elastic). As reported in Table 16, the gains from tourism are 20 percent lower if the tourism trade elasticity were to double relative to the point estimate that we estimate in the data.[52]

Second, we explore the sensitivity of the gains from tourism with respect to different assumptions about the spatial labor supply elasticity. As noted in Section 5 above, our preferred estimate of $\widetilde{\kappa} = 7.8$ is significantly larger than many of the estimates in the existing literature that have exploited shorter-term variation over time, rather than cross-sectional estimates. Table 16 thus reports the gains from tourism across three alternative parametrizations for $\widetilde{\kappa} = 7.8$, $\widetilde{\kappa} = 4.8$ and $\widetilde{\kappa} = 1.8$, with the first one equal to our empirical estimate from the data. Interestingly, our welfare quantification appears to be quite robust to different assumptions about the spatial labor supply elasticity. Holding all other parameters constant, the gains from tourism range between 4.4 and 4.8 percent across the different rows, and the gains from international-only tourism range between 1.6 and 2.8 percent.

Finally, we explore to what extent lower values of the estimated cross-sector spillover parameter affect the welfare results. As we have discussed in the previous section, the services-to-manufacturing externality matters directly for the extent of net gains or losses in traded goods production in the aggregate due to the development of tourism. This is also apparent in Table 16: holding other parameter values at their baseline, the estimated gains from tourism range between 4.4 and 2.5 percent as we move from the baseline calibration of $\gamma_S = 0.088$ to 0.058, 0.028 and finally 0. In turn, the gains from international tourism range between 1.6 percent in the baseline calibration to 0.75 percent in the absence of co-agglomeration forces.

For completeness, the table also reports the full cross of these parameter ranges. In particular, moving towards the lower left of each panel tends to increase the estimated gains from tourism (reducing $\sigma_T$ and $\widetilde{\kappa}$, and increasing $\gamma_S$). Conversely, moving toward the upper right of each panel

[52] Note that a very similar logic applies to the elasticity of substitution between tourism and manufacturing consumption ($\rho$). A lower value magnifies the gains from tourism because the less substitutable tourism becomes relative to other consumption, the more will an increase in the frictions to tourism trade deprive consumers from the benefits of tourism consumption. In the limit of $\rho = 1$ (Cobb-Douglas) this leads to infinite gains from tourism.

tends to lower the estimated gains from tourism (increasing $\sigma_T$ and $\widetilde{\kappa}$, and reducing $\gamma_S$). Online appendix Table A.7 also reports the identical exercises depicted in Table 16, but instead of varying the parameter of the cross-sector agglomeration force, we instead vary the value of the within-manufacturing spillover, $\gamma_M$, in the same way, while holding $\gamma_S$ constant. In line with the discussion of the role of the spillover parameters at the end of Section 4, we find that, for every given parameter combination of $\sigma_T$ and $\widetilde{\kappa}$, the estimated gains from both domestic and international tourism increase as we reduce the strength of the within-manufacturing agglomeration externality.

# 7 Conclusion

Much of the existing literature on tourism and economic development has been critical about tourism's long-term implications, especially in developing countries. At the same time, governments around the world and international organizations are showing widespread interest in tourism and have committed substantial amounts of public funds for national and regional tourism promotion policies. Somewhat surprisingly in this context, both the existing literature and current policy proposals have been based on limited empirical evidence on the long-term economic consequences of tourism.

This paper combines a rich collection of Mexican microdata with a spatial equilibrium model of trade in goods and tourism services and a new empirical strategy in order to contribute to our understanding of the long-term economic consequences of tourism in a developing country. To estimate the reduced-form effects of differences in tourism exposure on local economic outcomes in today's cross-section of Mexican municipalities, we exploit oceanographic and geological variation in beach quality along the Mexican coastline to construct instrumental variables. To guide the estimation of tourism's aggregate welfare implications, we then write down a spatial equilibrium model, inform its calibration using the reduced-form moments, and explore a number of model-based general equilibrium counterfactuals.

The analysis presents several findings. We find that tourism causes large and significant long-run local economic gains. Given that tourism has had more than five decades to shape relative regional economic outcomes in Mexico in a setting with labor mobility, the raw empirical moment speaking most directly to this effect is the fact that a 10 percent increase in local tourism revenues leads to a 2.8 percent increase in relative total local employment and a 2.2 percent increase in the local population. Contrary to much of the existing literature, we find that these local effects are in part driven by sizable positive multiplier effects on manufacturing production. Through the lens of the model, which takes into account other general equilibrium forces such as the gain in market access for local manufacturing producers brought about by tourism, we find that these multiplier effects provide evidence of positive spillovers from the local services sector on traded goods production. In particular, we estimate significant cross-sector spillovers in addition to within-sector localization economies within manufacturing. Interestingly, while these two sources of agglomeration economies reinforce one another leading to the large observed re-allocations of manufacturing and total GDP towards tourism centers in the data, we find that they largely offset one another for the aggregate implications of tourism. That is, while tourism leads to sizable gains in agglomeration economies at the local level, these gains are largely offset by reductions in agglomeration economies at the national level, so that the aggregate welfare gains from tourism are mainly driven by a classical market integration effect.

The analysis serves to inform currently ongoing policy debates in two central ways. First, we provide credible empirical evidence on the long-term effects of tourism activity on economic outcomes. Given that most of the current tourism promotion policies are targeted at either lowering travel frictions to particular regions (the $t_{nk}$ in our framework), or at increasing the local attractiveness for tourism (the $A_n$ in our framework), our results on both the local and aggregate implications of tourism integration are directly related to these policies. Second, the methodology that we propose in this paper provides a useful empirical tool to study a number of additional unanswered questions about the regional and aggregate effects of tourism, or similar economic shocks that affect regions within a country differently.

## References


Adao, R., Costinot, A., & Donaldson, D. (2015). *Nonparametric counterfactual predictions in neoclassical models of international trade* (Tech. Rep.). National Bureau of Economic Research.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the Berlin wall. *forthcoming, Econometrica*.

Allcott, H., & Keniston, D. (2014). *Dutch disease or agglomeration? the local economic effects of natural resource booms in modern america* (Tech. Rep.). National Bureau of Economic Research.

Allen, T., & Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, *1085*, 1139.

Allen, T., & Arkolakis, C. (2016). *Optimal city structure* (Tech. Rep.).

Anderson, S. P., De Palma, A., & Thisse, J. F. (1992). *Discrete choice theory of product differentiation*. MIT press.

Arezki, R., Cherif, R., & Piotrowski, J. (2009). *Tourism specialization and economic development: Evidence from the unesco world heritage list* (No. 9-176). International Monetary Fund.

Arkolakis, C., Costinot, A., & Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, *102*(1), 94–130.

Atkin, D., Faber, B., & Gonzalez-Navarro, M. (2015). Retail globalization and household welfare: Evidence from mexico. *UC Berkeley mimeo*.

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the United States. *The American Economic Review*, *103*(6), 2121–2168.

Bryan, G., & Morten, M. (2015). Economic development and the spatial allocation of labor: Evidence from Indonesia. *Manuscript, London School of Economics and Stanford University*.

Caliendo, L., Dvorkin, M., & Parro, F. (2015). *The impact of trade on labor market dynamics* (Tech. Rep.).

Caliendo, L., & Parro, F. (2014). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, rdu035.

Caliendo, L., Parro, F., Rossi-Hansberg, E., & Sarte, P.-D. (2014). *The impact of regional and sectoral productivity changes on the us economy* (Tech. Rep.). National Bureau of Economic Research.

Caselli, F., & Michaels, G. (2009). *Do oil windfalls improve living standards? evidence from brazil* (Tech. Rep.). National Bureau of Economic Research.

Copeland, B. R. (1991). Tourism, welfare and de-industrialization in a small open economy. *Economica*, 515–529.

Davis, M. A., & Ortalo-Magné, F. (2011). Household expenditures, wages, rents. *Review of Economic Dynamics*, *14*(2), 248–261.

Dekle, R., Eaton, J., & Kortum, S. (2007). Unbalanced trade. *American Economic Review*, *97*(2), 351–355.

DFID. (1999). *Tourism and poverty alleviation: Untapped potential*. DFID, London.

Dieke, P. (2000). *The political economy of tourism development in africa.* Cognizant Communication Corporation.

Donaldson, D. (in press). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*.

Donaldson, D., & Hornbeck, R. (2015). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*.

Eilat, Y., & Einav, L. (2004). Determinants of international tourism: a three-dimensional panel data analysis. *Applied Economics*, *36*(12), 1315–1327.

Fajgelbaum, P. D., Morales, E., Serrato, J. C. S., & Zidar, O. M. (2015). *State taxes and spatial misallocation* (Tech. Rep.). National Bureau of Economic Research.

Galle, S., Rodrıguez-Clare, A., & Yi, M. (2014). Slicing the pie: Quantifying the aggregate and distributional effects of trade. *Unpublished manuscript, UC Berkeley*.

Hawkins, D. E., & Mann, S. (2007). The world banks role in tourism development. *Annals of Tourism Research*, *34*(2), 348–363.

Honey, M. (1999). *Who owns paradise?* Island Press.

Horowitz, J. L. (2001). The bootstrap. In J. J. Heckman & E. Leamer (Eds.), *Handbook of econometrics* (Vol. 5, p. 3159 - 3228). Elsevier.

Kline, P., & Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority*. *Quarterly Journal of Economics*, *129*(1).

Leatherman, S. P. (1997). Beach rating: a methodological approach. *Journal of coastal research*, 253–258.

Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. *Handbook of International Economics*, *4*.

Mian, A., & Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics*, *124*(4), 1449–1496.

Monte, F., Redding, S., & Rossi-Hansberg, E. (2015). *Commuting, migration and local employment elasticities* (Tech. Rep.).

Redding, S. J. (2015). *Goods trade, factor mobility and welfare* (Tech. Rep.). National Bureau of Economic Research.

Rosenthal, S., & Strange, W. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, *4*, 2119–2171.

Sequeira, T. N., & Macas Nunes, P. (2008). Does tourism influence economic growth? a dynamic panel data approach. *Applied Economics*, *40*(18), 2431–2441.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India. *American Economic Journal: Applied Economics*, *2*(4), 1–41.

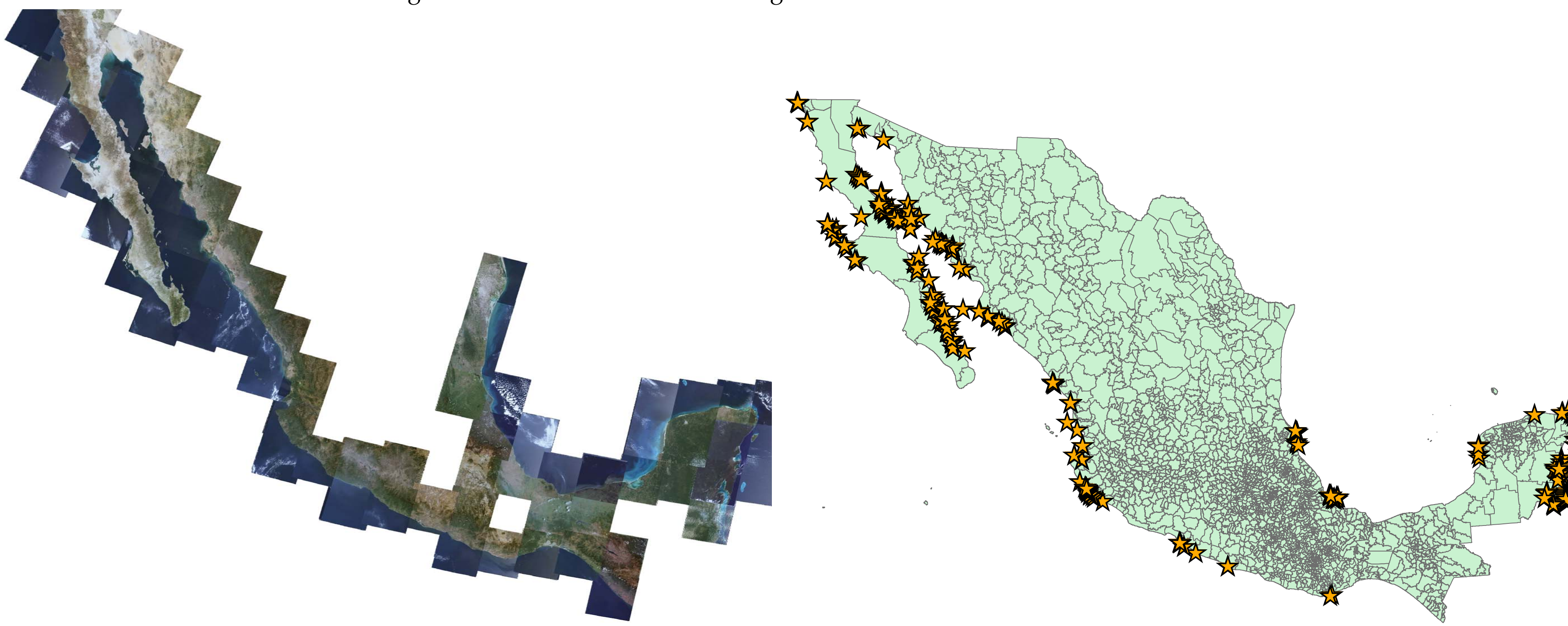
Weaver, D., Oppermann, M., et al. (2000). *Tourism management*. John Wiley and Sons.

WorldBank. (1979). *Tourism: passport to development?* World Bank/UNESCO Conference.

# 8 Figures and Tables

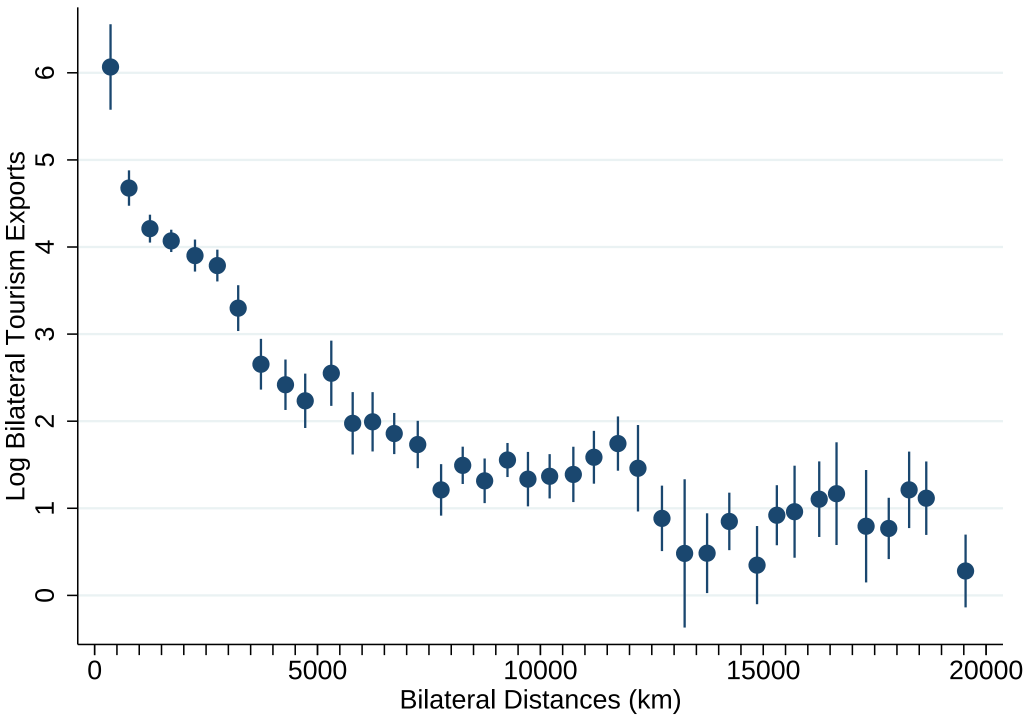
## Figures

Figure 1: Beach Characteristics along the Mexican Coastline



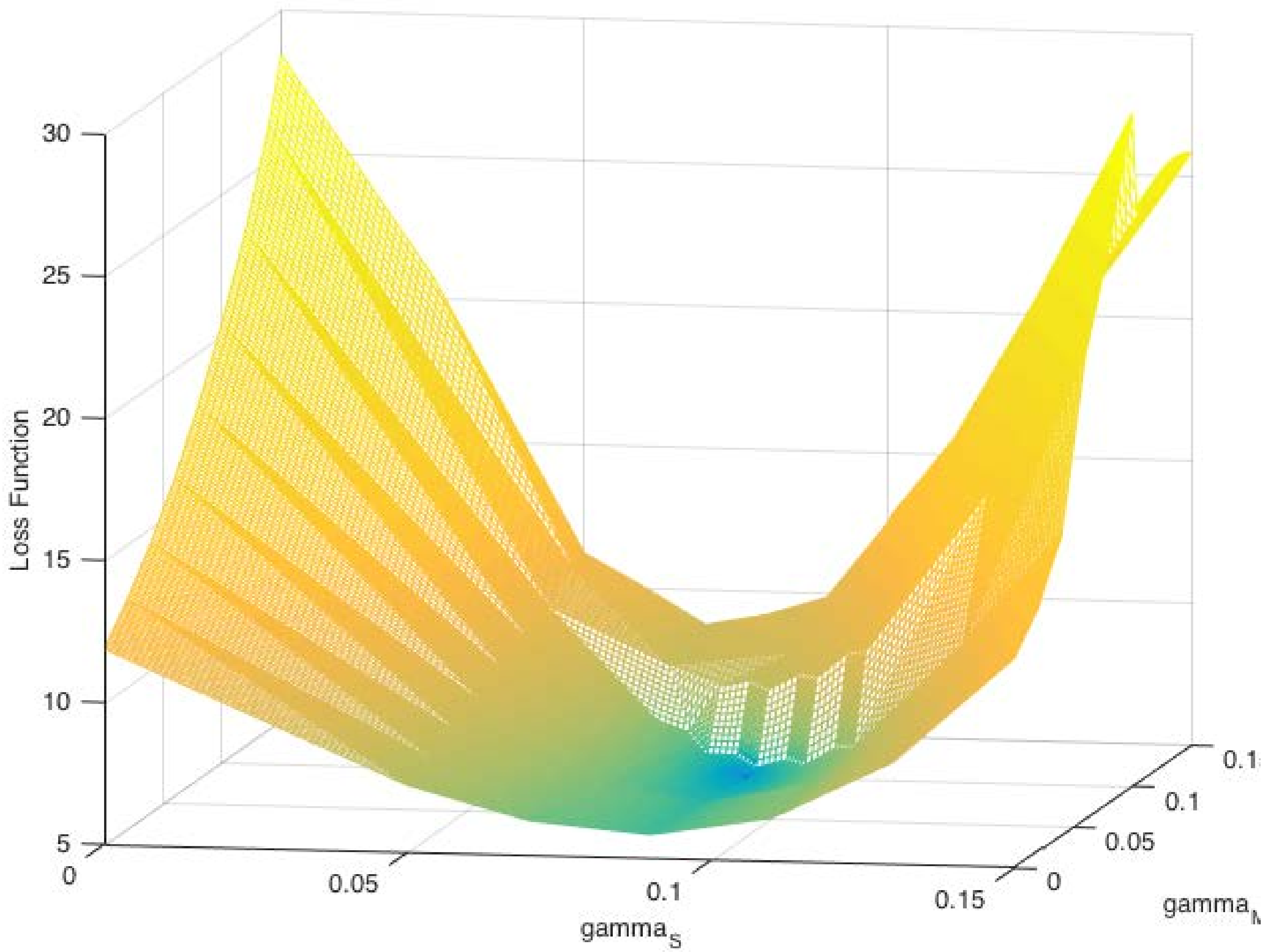
*Notes:* The left panel displays the remote sensing satellite data covering the Mexican coastline at a resolution of 30x30 meter. The right panel displays the location of islands within 5 kilometers from the Mexican coastline.

Figure 2: Tourism's Distance Decay




*Notes:* Point estimates from a regression of log tourism exports on 500 km bilateral distance bins in addition to origin-by-year fixed effects, destination-by-year fixed effects, and dummies for common border, language and colonial ties. The figure depicts 95% confidence intervals based on standard errors that are clustered at the level of origin-destination pairs.

Figure 3: Indirect Inference for Best-Fitting Combination of Agglomeration Forces




*Notes:* See Section 5 for discussion.

## Tables

Table 1: Beach Tourism in Mexico

| | Number of Municipalities | Sum of Hotel Revenues in 1998 and 2008 (Thousands of Pesos) | Share of National Hotel Revenues 1998 and 2008 |
|---|---|---|---|
| Inland Municipalities | 2305 | 46,070,000 | 0.365 |
| Coastal Municipalities | 150 | 80,130,000 | 0.635 |

*Notes:* Source: Censos Economicos for 1998 and 2008.

Table 2: Descriptive Statistics

| Data Source | Variable | 1998 Censos Economicos or 2000 Population Census | | | | | 2008 Censos Economicos or 2010 Population Census | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | mean | sd | min | max | N | mean | sd | min | max |
| Censos Economicos | state id | 2,434 | 19.30 | 7.32 | 1.00 | 32.00 | 2,455 | 19.26 | 7.34 | 1.00 | 32.00 |
| | gdp | 2,434 | 1,528,000.00 | 9,613,000.00 | 6.00 | 251,800,000.00 | 2,455 | 4,480,000.00 | 27,220,000.00 | 21.00 | 704,200,000.00 |
| | log gdp | 2,434 | 9.92 | 2.97 | 1.79 | 19.34 | 2,455 | 10.92 | 2.96 | 3.05 | 20.37 |
| | hotel sales | 2,434 | 12,847.00 | 138,994.00 | 0.00 | 5,230,000.00 | 2,455 | 38,668.00 | 433,757.00 | 0.00 | 13,730,000.00 |
| | log hotel sales | 2,434 | 3.28 | 3.92 | 0.00 | 16.16 | 2,455 | 4.53 | 4.26 | 0.00 | 17.13 |
| | number of hotels | 2,434 | 4.42 | 18.53 | 0.00 | 431.00 | 2,455 | 7.51 | 26.72 | 0.00 | 457.00 |
| Population Census | population | 2,434 | 39,832.00 | 119,060.00 | 105.00 | 1,763,000.00 | 2,455 | 45,603.00 | 132,175.00 | 90.00 | 1,794,000.00 |
| | log population | 2,434 | 9.34 | 1.50 | 4.65 | 14.38 | 2,455 | 9.42 | 1.56 | 4.50 | 14.40 |
| | employment | 2,434 | 14,542.00 | 48,042.00 | 34.00 | 825,945.00 | 2,455 | 17,999.00 | 60,391.00 | 37.00 | 874,120.00 |
| | log employment | 2,434 | 8.17 | 1.56 | 3.53 | 13.62 | 2,455 | 8.27 | 1.64 | 3.61 | 13.68 |
| Satellite Data and GIS | coast id | 2,434 | 0.06 | 0.24 | 0.00 | 1.00 | 2,455 | 0.06 | 0.24 | 0.00 | 1.00 |
| | island dummy | 2,434 | 0.02 | 0.14 | 0.00 | 1.00 | 2,455 | 0.02 | 0.14 | 0.00 | 1.00 |
| | beach type 1 share within 100 m of coast | 2,434 | 0.00 | 0.00 | 0.00 | 0.08 | 2,455 | 0.00 | 0.00 | 0.00 | 0.08 |
| | beach type 2 share within 100 m of coast | 2,434 | 0.00 | 0.00 | 0.00 | 0.01 | 2,455 | 0.00 | 0.00 | 0.00 | 0.01 |
| | beach type 3 share within 100 m of coast | 2,434 | 0.00 | 0.00 | 0.00 | 0.06 | 2,455 | 0.00 | 0.00 | 0.00 | 0.06 |
| | beach type 4 share within 100 m of coast | 2,434 | 0.00 | 0.00 | 0.00 | 0.05 | 2,455 | 0.00 | 0.00 | 0.00 | 0.05 |
| | beach types 1-4 share within 100 m of coast | 2,434 | 0.00 | 0.00 | 0.00 | 0.09 | 2,455 | 0.00 | 0.00 | 0.00 | 0.09 |
| | distance to northern border (km) | 2,434 | 753.40 | 265.80 | 6.59 | 1,348.00 | 2,455 | 755.10 | 266.00 | 6.59 | 1,348.00 |
| | distance to Mex City (km) | 2,434 | 453.70 | 372.50 | 2.30 | 2,271.00 | 2,455 | 454.20 | 372.10 | 2.30 | 2,271.00 |
| Mexican Statistical Institute (INEGI) | state capital dummy | 2,434 | 0.02 | 0.14 | 0.00 | 1.00 | 2,455 | 0.02 | 0.14 | 0.00 | 1.00 |
| | old city dummy | 2,434 | 0.02 | 0.13 | 0.00 | 1.00 | 2,455 | 0.02 | 0.13 | 0.00 | 1.00 |
| | colonial port dummy | 2,434 | 0.00 | 0.03 | 0.00 | 1.00 | 2,455 | 0.00 | 0.03 | 0.00 | 1.00 |
| | average monthly temperature (Celsius x 10) | 2,434 | 197.30 | 40.30 | 104.50 | 290.30 | 2,455 | 197.40 | 40.36 | 104.50 | 290.30 |
| | average monthly percipitation (mm) | 2,434 | 88.79 | 50.57 | 5.99 | 336.50 | 2,455 | 89.15 | 50.77 | 5.99 | 336.50 |

See Section 2 for a description of the datasets.

Table 3: Tourism's Effect on Municipality Employment and Population

| Dependent variables: | Log Municipality Employment 2000, 2010 | | | | | | | | Log Municipality Population 2000, 2010 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) Island IV | (4) Island IV | (5) Beach IV | (6) Beach IV | (7) Both IVs | (8) Both IVs | (9) Both IVs | (10) Both IVs |
| Log Hotel Sales | 0.236*** | 0.218*** | 0.295*** | 0.323*** | 0.228** | 0.243*** | 0.263*** | 0.275*** | 0.212*** | 0.221*** |
| | (0.00605) | (0.00568) | (0.0890) | (0.122) | (0.0917) | (0.0888) | (0.0573) | (0.0643) | (0.0613) | (0.0686) |
| Log Distance to US Border | 0.0790** | -0.0290 | 0.105** | 0.0206 | 0.0754 | -0.0171 | 0.0910** | -0.00217 | 0.138*** | 0.0444 |
| | (0.0386) | (0.0416) | (0.0513) | (0.0676) | (0.0568) | (0.0588) | (0.0436) | (0.0486) | (0.0465) | (0.0514) |
| Log Distance to Mexico City | -0.587*** | -0.578*** | -0.508*** | -0.463*** | -0.598*** | -0.550*** | -0.551*** | -0.516*** | -0.595*** | -0.568*** |
| | (0.0258) | (0.0284) | (0.122) | (0.137) | (0.125) | (0.101) | (0.0810) | (0.0761) | (0.0862) | (0.0809) |
| Log Municipality Area | 0.340*** | 0.351*** | 0.263** | 0.223 | 0.350*** | 0.320*** | 0.305*** | 0.282*** | 0.364*** | 0.343*** |
| | (0.0172) | (0.0169) | (0.118) | (0.150) | (0.121) | (0.110) | (0.0774) | (0.0810) | (0.0826) | (0.0863) |
| State Capital Dummy | | 0.796*** | | 0.378 | | 0.696* | | 0.570* | | 0.540* |
| | | (0.191) | | (0.506) | | (0.398) | | (0.304) | | (0.328) |
| Old City Dummy | | 1.028*** | | 0.624 | | 0.931** | | 0.809** | | 0.836** |
| | | (0.229) | | (0.513) | | (0.404) | | (0.323) | | (0.349) |
| Colonial Port Dummy | | 0.699*** | | 0.300 | | 0.603* | | 0.483* | | 0.589* |
| | | (0.141) | | (0.509) | | (0.364) | | (0.291) | | (0.308) |
| Log Average Percipitation | | 0.263*** | | 0.244*** | | 0.258*** | | 0.253*** | | 0.241*** |
| | | (0.0402) | | (0.0483) | | (0.0431) | | (0.0425) | | (0.0428) |
| Log Average Temperature | | 0.233** | | 0.194 | | 0.224** | | 0.212* | | 0.273** |
| | | (0.106) | | (0.123) | | (0.111) | | (0.111) | | (0.108) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 |
| R-squared | 0.659 | 0.682 | 0.642 | 0.635 | 0.658 | 0.679 | 0.655 | 0.668 | 0.642 | 0.660 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | | | 9.549 | 5.748 | 11.71 | 11.81 | 14.36 | 11.59 | 14.36 | 11.59 |
| Over-ID Test P-Value | | | | | | | 0.625 | 0.617 | 0.538 | 0.533 |

*Notes:* Island IV is a dummy indicating whether an offshore island is within 5 km of the municipalities' coastline. Beach IV is the fraction of municipality area within 100 m of the coastline covered by white sand pixels that lie within the wavelength ranges of the top-ranked Mexican beach. Log hotel sales are measured with the hyperbolic inverse sine transformation as described in Section 3. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

## Table 4: Robustness

| Dependent variable: | Log Municipality Employment 2000, 2010 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Omitted Variables | | | Different Beach IVs and Excluding Origin Municipalities | | | | | | Sensitivity to Cutoffs | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| | Coastal Elevation | Fishery Potential | Both | Beach Type 1 (Baseline) | Beach Type 1 Exclude Origin Municipality | Beach Type 2 Exclude Origin Municipality | Beach Type 3 Exclude Origin Municipality | Beach Type 4 Exclude Origin Municipality | All Beach Types Exclude Origin Municipalities | Beach Type 1 Island Within 10 km | Beach Type 1 Within 200 m of Shoreline |
| | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs | Both IVs |
| Log Hotel Sales | 0.258*** | 0.276*** | 0.256*** | 0.275*** | 0.279*** | 0.301*** | 0.324*** | 0.262*** | 0.334*** | 0.266*** | 0.282*** |
| | (0.0735) | (0.0642) | (0.0746) | (0.0643) | (0.0656) | (0.0660) | (0.0913) | (0.0576) | (0.127) | (0.0625) | (0.0601) |
| Log Mean Coastal Elevation | -0.384 | | -0.404 | | | | | | | | |
| | (0.300) | | (0.250) | | | | | | | | |
| Log Stand Dev of Coastal Elevation | 0.0643 | | 0.0800 | | | | | | | | |
| | (0.177) | | (0.175) | | | | | | | | |
| Log Mean Ocean Primary Productivity | | 0.0338 | 0.0327 | | | | | | | | |
| | | (0.0586) | (0.0604) | | | | | | | | |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 4,889 | 4,889 | 4,887 | 4,887 | 4,887 | 4,887 | 4,881 | 4,889 | 4,889 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 | 2454 | 2454 | 2454 | 2454 | 2451 | 2455 | 2455 |
| First Stage F-Stat | 10.55 | 11.58 | 10.47 | 11.59 | 11.22 | 7.399 | 4.580 | 22.38 | 2.957 | 12.81 | 16.11 |
| Over-ID Test P-Value | 0.363 | 0.596 | 0.359 | 0.617 | 0.560 | 0.841 | 0.897 | 0.277 | 0.720 | 0.709 | 0.660 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

## Table 5: Placebo Falsification Tests

| Dependent variable: | Log Municipality Census Population | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Census Years: | 1921, 1930, 1940, 1950 | 2000, 2010 | 1921, 1930, 1940, 1950 | 2000, 2010 | 1921, 1930, 1940, 1950 | 2000, 2010 | 1921, 1930, 1940, 1950 | 2000, 2010 | 1921, 1930, 1940, 1950 | 2000, 2010 | 1921, 1930, 1940, 1950 | 2000, 2010 |
| | Island IV | Island IV | Beach IV Type 1 | Beach IV Type 1 | Beach IV Type 2 | Beach IV Type 2 | Beach IV Type 3 | Beach IV Type 3 | Beach IV Type 4 | Beach IV Type 4 | Beach IV All Types | Beach IV All Types |
| *Panel A: Left Hand Side with Inverse Hyperbolic Sine Transformation for Log Population* | | | | | | | | | | | | |
| Instrumental Variable | -0.151 | 0.510** | -34.12* | 12.38 | -236.3 | 138.2* | -34.60 | 16.16 | -74.91 | 47.08*** | -15.20 | 6.757* |
| | (0.350) | (0.233) | (19.26) | (9.327) | (241.9) | (81.59) | (29.52) | (13.19) | (66.70) | (16.49) | (9.522) | (3.979) |
| Log Distance to US Border | 0.121* | 0.0415 | 0.127** | 0.0385 | 0.125** | 0.0386 | 0.124* | 0.0390 | 0.127** | 0.0370 | 0.126** | 0.0382 |
| | (0.0636) | (0.0574) | (0.0634) | (0.0578) | (0.0635) | (0.0578) | (0.0635) | (0.0578) | (0.0635) | (0.0578) | (0.0634) | (0.0578) |
| Log Distance to Mexico City | -0.419*** | -0.878*** | -0.412*** | -0.880*** | -0.415*** | -0.880*** | -0.414*** | -0.880*** | -0.415*** | -0.880*** | -0.413*** | -0.880*** |
| | (0.0574) | (0.0321) | (0.0574) | (0.0322) | (0.0574) | (0.0322) | (0.0575) | (0.0322) | (0.0573) | (0.0321) | (0.0574) | (0.0322) |
| Log Municipality Area | 0.497*** | 0.633*** | 0.494*** | 0.637*** | 0.495*** | 0.637*** | 0.494*** | 0.637*** | 0.497*** | 0.636*** | 0.494*** | 0.637*** |
| | (0.0215) | (0.0205) | (0.0215) | (0.0205) | (0.0215) | (0.0204) | (0.0215) | (0.0205) | (0.0214) | (0.0204) | (0.0215) | (0.0204) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 |
| R-Squared | 0.231 | 0.400 | 0.235 | 0.399 | 0.232 | 0.399 | 0.233 | 0.399 | 0.234 | 0.400 | 0.234 | 0.399 |
| Number of Municipalities | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 |
| *Panel B: Left Hand Side with Log of One for Zero Population* | | | | | | | | | | | | |
| Instrumental Variable | -0.144 | 0.510** | -33.79* | 12.38 | -234.1 | 138.2* | -34.73 | 16.16 | -74.31 | 47.08*** | -15.12 | 6.757* |
| | (0.337) | (0.233) | (19.08) | (9.327) | (240.2) | (81.59) | (29.50) | (13.19) | (66.16) | (16.49) | (9.478) | (3.979) |
| Log Distance to US Border | 0.116* | 0.0415 | 0.121** | 0.0385 | 0.119** | 0.0386 | 0.119** | 0.0390 | 0.121** | 0.0370 | 0.121** | 0.0382 |
| | (0.0607) | (0.0574) | (0.0605) | (0.0578) | (0.0605) | (0.0578) | (0.0606) | (0.0578) | (0.0605) | (0.0578) | (0.0605) | (0.0578) |
| Log Distance to Mexico City | -0.427*** | -0.878*** | -0.420*** | -0.880*** | -0.423*** | -0.880*** | -0.423*** | -0.880*** | -0.424*** | -0.880*** | -0.421*** | -0.880*** |
| | (0.0542) | (0.0321) | (0.0542) | (0.0322) | (0.0542) | (0.0322) | (0.0543) | (0.0322) | (0.0541) | (0.0321) | (0.0542) | (0.0322) |
| Log Municipality Area | 0.499*** | 0.633*** | 0.495*** | 0.637*** | 0.496*** | 0.637*** | 0.496*** | 0.637*** | 0.498*** | 0.636*** | 0.496*** | 0.637*** |
| | (0.0205) | (0.0205) | (0.0206) | (0.0205) | (0.0206) | (0.0204) | (0.0206) | (0.0205) | (0.0204) | (0.0204) | (0.0205) | (0.0204) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 | 9,736 | 4,868 |
| R-Squared | 0.246 | 0.400 | 0.251 | 0.399 | 0.248 | 0.399 | 0.249 | 0.399 | 0.250 | 0.400 | 0.250 | 0.399 |
| Number of Municipalities | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 | 2434 |
| *Panel C: Left Hand Side with Simple Logs (Dropping Municipalities with Zero Populations in the Past)* | | | | | | | | | | | | |
| Instrumental Variable | -0.0483 | 0.555** | -26.81 | 11.45 | -178.5 | 129.8 | -34.96 | 15.62 | -60.95 | 45.29*** | -12.85 | 6.372 |
| | (0.250) | (0.242) | (19.00) | (9.835) | (224.9) | (83.58) | (29.50) | (13.25) | (59.31) | (16.38) | (9.494) | (4.130) |
| Log Distance to US Border | 0.0404 | 0.0572 | 0.0439 | 0.0557 | 0.0424 | 0.0557 | 0.0433 | 0.0559 | 0.0445 | 0.0542 | 0.0441 | 0.0553 |
| | (0.0353) | (0.0600) | (0.0350) | (0.0604) | (0.0351) | (0.0604) | (0.0350) | (0.0604) | (0.0350) | (0.0604) | (0.0350) | (0.0604) |
| Log Distance to Mexico City | -0.541*** | -0.864*** | -0.536*** | -0.865*** | -0.538*** | -0.865*** | -0.536*** | -0.865*** | -0.539*** | -0.865*** | -0.536*** | -0.865*** |
| | (0.0258) | (0.0331) | (0.0257) | (0.0332) | (0.0257) | (0.0332) | (0.0257) | (0.0332) | (0.0256) | (0.0331) | (0.0257) | (0.0332) |
| Log Municipality Area | 0.523*** | 0.636*** | 0.521*** | 0.640*** | 0.522*** | 0.640*** | 0.521*** | 0.640*** | 0.523*** | 0.639*** | 0.521*** | 0.640*** |
| | (0.0130) | (0.0210) | (0.0132) | (0.0210) | (0.0132) | (0.0210) | (0.0132) | (0.0210) | (0.0131) | (0.0209) | (0.0132) | (0.0210) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 9,548 | 4,774 | 9,548 | 4,774 | 9,548 | 4,774 | 9,548 | 4,774 | 9,548 | 4,774 | 9,548 | 4,774 |
| R-Squared | 0.432 | 0.397 | 0.437 | 0.396 | 0.434 | 0.396 | 0.436 | 0.396 | 0.436 | 0.396 | 0.437 | 0.396 |
| Number of Municipalities | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 | 2387 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table 6: Model-Based Test of Direct Effect on Local Residential Amenities

| Dependent variable: | Log Municipality Residential Amenities | | | | |
|---|---|---|---|---|---|
| | (1) Not Using Island IV | (2) Not Using Beach IV 1 | (3) Not Using Beach IV 2 | (4) Not Using Beach IV 3 | (5) Not Using Beach IV 4 |
| Left -Out IV | 0.179 | 2.078 | 48.07 | 5.882 | -9.071 |
| | (0.287) | (4.929) | (56.76) | (7.078) | (7.903) |
| Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 300 | 300 | 300 | 300 | 300 |
| R-Squared | 0.416 | 0.390 | 0.390 | 0.386 | 0.373 |
| Number of Clusters | 32 | 32 | 32 | 32 | 32 |

*Notes:* See Sections 3 and 5 for discussion. Standard errors are clustered at the level of Mexican states. * 10%, ** 5%, *** 1% significance levels.

Table 7: Tourism's Effect on Municipality Wage Bill, GDP and Wages

| | Censos Economicos 1998, 2008 | | | | | | | | | | | | Population Census 2000, 2010 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Dependent Variables: | Log Labor Income | Log Labor Income | Log GDP | Log GDP | Log GDP (w/o Hotel) | Log GDP (w/o Hotel) | Log GDP (Manu+ Mining) | Log GDP (Manu+ Mining) | Log GDP (Manu) | Log GDP (Manu) | Log GDP (Agri) | Log GDP (Agri) | Log Wage Residual | Log Wage Residual |
| | OLS | Both IVs | OLS | Both IVs | OLS | Both IVs | OLS | Both IVs | OLS | Both IVs | OLS | Both IVs | OLS | Both IVs |
| Log Hotel Sales | 0.480*** | 0.477*** | 0.464*** | 0.425*** | 0.458*** | 0.392*** | 0.516*** | 0.273* | 0.530*** | 0.317** | 0.291*** | 0.195 | 0.0126*** | 0.0333*** |
| | (0.0104) | (0.0966) | (0.0104) | (0.0932) | (0.0106) | (0.0979) | (0.0144) | (0.147) | (0.0146) | (0.124) | (0.0164) | (0.156) | (0.00114) | (0.0108) |
| Log Distance to US Border | -0.364*** | -0.366*** | -0.299*** | -0.317*** | -0.304*** | -0.336*** | -0.290*** | -0.405*** | -0.181* | -0.282** | 0.267** | 0.222* | -0.0970*** | -0.0872*** |
| | (0.0713) | (0.0808) | (0.0691) | (0.0814) | (0.0696) | (0.0837) | (0.105) | (0.132) | (0.105) | (0.127) | (0.107) | (0.132) | (0.00661) | (0.00893) |
| Log Distance to Mexico City | -0.641*** | -0.645*** | -0.705*** | -0.747*** | -0.711*** | -0.783*** | -0.871*** | -1.137*** | -0.889*** | -1.123*** | -0.231*** | -0.336* | 0.00235 | 0.0251* |
| | (0.0510) | (0.111) | (0.0489) | (0.112) | (0.0494) | (0.117) | (0.0678) | (0.176) | (0.0690) | (0.152) | (0.0753) | (0.185) | (0.00564) | (0.0130) |
| Log Municipality Area | 0.183*** | 0.188 | 0.217*** | 0.264** | 0.221*** | 0.302** | 0.181*** | 0.478*** | 0.112*** | 0.373** | 0.451*** | 0.569*** | 0.00955*** | -0.0159 |
| | (0.0323) | (0.121) | (0.0310) | (0.118) | (0.0313) | (0.124) | (0.0430) | (0.186) | (0.0428) | (0.157) | (0.0435) | (0.193) | (0.00368) | (0.0137) |
| State Capital Dummy | 1.224*** | 1.240*** | 1.164*** | 1.317*** | 1.197*** | 1.461*** | 0.689** | 1.659** | 0.736** | 1.589** | 0.287 | 0.671 | 0.0519** | -0.0312 |
| | (0.207) | (0.434) | (0.210) | (0.431) | (0.214) | (0.456) | (0.319) | (0.711) | (0.348) | (0.641) | (0.661) | (0.955) | (0.0221) | (0.0534) |
| Old City Dummy | 1.310*** | 1.325*** | 1.307*** | 1.454*** | 1.324*** | 1.579*** | 1.242*** | 2.179*** | 1.241*** | 2.064*** | 0.733 | 1.104 | 0.0435 | -0.0367 |
| | (0.240) | (0.442) | (0.242) | (0.447) | (0.246) | (0.474) | (0.363) | (0.751) | (0.394) | (0.690) | (0.809) | (1.041) | (0.0271) | (0.0562) |
| Colonial Port Dummy | 0.829** | 0.843* | 0.548 | 0.693 | 0.551 | 0.803 | 0.400 | 1.326 | 0.462 | 1.275 | -0.873 | -0.507 | -0.0977*** | -0.177** |
| | (0.325) | (0.477) | (0.446) | (0.512) | (0.486) | (0.530) | (0.959) | (0.832) | (0.962) | (0.817) | (0.739) | (1.072) | (0.0320) | (0.0707) |
| Log Average Percipitation | -0.629*** | -0.628*** | -0.578*** | -0.571*** | -0.577*** | -0.566*** | -0.960*** | -0.917*** | -0.937*** | -0.900*** | -0.182 | -0.165 | -0.0640*** | -0.0677*** |
| | (0.0807) | (0.0839) | (0.0760) | (0.0787) | (0.0765) | (0.0802) | (0.106) | (0.118) | (0.106) | (0.114) | (0.111) | (0.114) | (0.00985) | (0.0105) |
| Log Average Temperature | 0.577*** | 0.578*** | 1.069*** | 1.083*** | 1.077*** | 1.102*** | 1.394*** | 1.486*** | 1.437*** | 1.518*** | 2.367*** | 2.403*** | 0.0160 | 0.00815 |
| | (0.197) | (0.197) | (0.184) | (0.187) | (0.186) | (0.191) | (0.270) | (0.291) | (0.276) | (0.291) | (0.305) | (0.314) | (0.0266) | (0.0282) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,596 | 4,596 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 |
| R-squared | 0.636 | 0.636 | 0.643 | 0.641 | 0.636 | 0.631 | 0.506 | 0.457 | 0.507 | 0.469 | 0.429 | 0.422 | 0.189 | 0.133 |
| Number of Municipalities | 2385 | 2385 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | | 11.27 | | 11.59 | | 11.59 | | 11.59 | | 11.59 | | 11.59 | | 11.59 |
| Over-ID Test P-Value | | 0.380 | | 0.107 | | 0.140 | | 0.137 | | 0.308 | | 0.214 | | 0.305 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table 8: Assessing the Role of Infrastructure

| | Population Census 2000, 2010 | | | | Censos Economicos 1998, 2008 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Dependent Variables: | Log Employment Both IVs Baseline | Log Employment Both IVs Infrastructure Controls | Log Population Both IVs Baseline | Log Population Both IVs Infrastructure Controls | Log GDP Both IVs Baseline | Log GDP Both IVs Infrastructure Controls | Log GDP (Manu+Mining) Both IVs Baseline | Log GDP (Manu+Mining) Both IVs Infrastructure Controls | Log Manu GDP Both IVs Baseline | Log Manu GDP Both IVs Infrastructure Controls |
| Log Hotel Sales | 0.275*** | 0.257*** | 0.221*** | 0.201*** | 0.425*** | 0.400*** | 0.273* | 0.245** | 0.317** | 0.287** |
| | (0.0643) | (0.0674) | (0.0686) | (0.0707) | (0.0932) | (0.0883) | (0.147) | (0.122) | (0.124) | (0.120) |
| Log Distance to US Border | -0.00217 | -0.0113 | 0.0444 | 0.0373 | -0.317*** | -0.293*** | -0.405*** | -0.350*** | -0.282** | -0.234** |
| | (0.0486) | (0.0433) | (0.0514) | (0.0448) | (0.0814) | (0.0704) | (0.132) | (0.111) | (0.127) | (0.109) |
| Log Distance to Mexico City | -0.516*** | -0.417*** | -0.568*** | -0.450*** | -0.747*** | -0.430*** | -1.137*** | -0.645*** | -1.123*** | -0.654*** |
| | (0.0761) | (0.0472) | (0.0809) | (0.0489) | (0.112) | (0.0666) | (0.176) | (0.0934) | (0.152) | (0.0932) |
| Log Municipality Area | 0.282*** | 0.238*** | 0.343*** | 0.285*** | 0.264** | 0.0188 | 0.478*** | 0.0343 | 0.373** | -0.0339 |
| | (0.0810) | (0.0490) | (0.0863) | (0.0510) | (0.118) | (0.0676) | (0.186) | (0.0974) | (0.157) | (0.0962) |
| State Capital Dummy | 0.570* | 0.285 | 0.540* | 0.223 | 1.317*** | 0.577* | 1.659** | 0.551 | 1.589** | 0.514 |
| | (0.304) | (0.241) | (0.328) | (0.255) | (0.431) | (0.307) | (0.711) | (0.463) | (0.641) | (0.466) |
| Old City Dummy | 0.809** | 0.605** | 0.836** | 0.620** | 1.454*** | 0.885** | 2.179*** | 1.251** | 2.064*** | 1.164** |
| | (0.323) | (0.278) | (0.349) | (0.295) | (0.447) | (0.347) | (0.751) | (0.534) | (0.690) | (0.539) |
| Colonial Port Dummy | 0.483* | 0.0936 | 0.589* | 0.127 | 0.693 | -0.438 | 1.326 | -0.340 | 1.275 | -0.327 |
| | (0.291) | (0.331) | (0.308) | (0.382) | (0.512) | (0.400) | (0.832) | (0.564) | (0.817) | (0.531) |
| Log Average Percipitation | 0.253*** | 0.442*** | 0.241*** | 0.462*** | -0.571*** | -0.0189 | -0.917*** | -0.0714 | -0.900*** | -0.0841 |
| | (0.0425) | (0.0793) | (0.0428) | (0.0817) | (0.0787) | (0.113) | (0.118) | (0.163) | (0.114) | (0.157) |
| Log Average Temperature | 0.212* | -0.0339 | 0.273** | 0.0179 | 1.083*** | 0.580*** | 1.486*** | 0.721*** | 1.518*** | 0.762*** |
| | (0.111) | (0.104) | (0.108) | (0.100) | (0.187) | (0.169) | (0.291) | (0.262) | (0.291) | (0.265) |
| Log Distance to Nearest Airport | | -0.373*** | | -0.414*** | | -0.807*** | | -1.113*** | | -1.104*** |
| | | (0.0678) | | (0.0702) | | (0.0982) | | (0.142) | | (0.140) |
| Log Distance to Nearest Seaport | | -0.0616* | | -0.0413 | | 0.0128 | | 0.0139 | | -0.00515 |
| | | (0.0341) | | (0.0341) | | (0.0580) | | (0.0904) | | (0.0907) |
| Log Paved Road Kilometers in Municipality | | 0.158** | | 0.188*** | | 0.505*** | | 0.825*** | | 0.774*** |
| | | (0.0667) | | (0.0700) | | (0.0930) | | (0.134) | | (0.129) |
| Log Railway Kilometers in Municipality | | 0.0975*** | | 0.105*** | | 0.269*** | | 0.441*** | | 0.429*** |
| | | (0.0235) | | (0.0244) | | (0.0327) | | (0.0452) | | (0.0457) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 | 4,889 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | 11.59 | 11.87 | 11.59 | 11.87 | 11.59 | 11.87 | 11.59 | 11.87 | 11.59 | 11.87 |
| Over-ID Test P-Value | 0.617 | 0.603 | 0.533 | 0.529 | 0.107 | 0.0594 | 0.137 | 0.0665 | 0.308 | 0.208 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table 9: Tourism's Effect on Traded Sector Production By Degree of Input Intensity

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: Log Manufacturing GDP | Below Median Input Intensity (10 Sectors) | Above Median Input Intensity (11 Sectors) | Sectors Not in Tourism Satellite Use Table (16 Sectors) | Sectors in Tourism Satellite Use Table (5 Sectors) |
| | Both IVs | Both IVs | Both IVs | Both IVs |
| *Panel A: Left Hand Side with HIS Transformation* | | | | |
| Log Hotel Sales | 0.586*** | 0.625*** | 0.530*** | 0.790*** |
| | (0.164) | (0.164) | (0.154) | (0.180) |
| Year-By-Coast-By-Sector FX | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ |
| Observations | 53,779 | 48,890 | 73,335 | 29,334 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | 11.61 | 11.61 | 11.61 | 11.61 |
| Over-ID Test P-Value | 0.222 | 0.722 | 0.363 | 0.526 |
| *Panel B: Left Hand Side with Log(Zero+1)* | | | | |
| Log Hotel Sales | 0.552*** | 0.586*** | 0.498*** | 0.742*** |
| | (0.156) | (0.154) | (0.145) | (0.170) |
| Year-By-Coast-By-Sector FX | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ |
| Observations | 53,779 | 48,890 | 73,335 | 29,334 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | 11.61 | 11.61 | 11.61 | 11.61 |
| Over-ID Test P-Value | 0.219 | 0.728 | 0.362 | 0.524 |
| *Panel C: Left Hand Side with Simple Logs (Dropping Zeroes)* | | | | |
| Log Hotel Sales | 0.326*** | 0.440*** | 0.377*** | 0.385*** |
| | (0.0778) | (0.0702) | (0.0581) | (0.112) |
| Year-By-Coast-By-Sector FX | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ |
| Observations | 19,637 | 13,516 | 21,184 | 11,969 |
| Number of Municipalities | 2224 | 2057 | 2161 | 2203 |
| First Stage F-Stat | 17.97 | 21.77 | 24.42 | 14.09 |
| Over-ID Test P-Value | 0.424 | 0.828 | 0.516 | 0.837 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table 10: Tourism's Trade Elasticity

| Dependent Variables: | Log Tourism Exports from Origin to Destination | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | Same Year | Same Year | 1-Year Lag | 2-Year Lag | 3-Year Lag | 4-Year Lag | 5-Year Lag |
| | OLS | IV | IV | IV | IV | IV | IV |
| *Panel A: All Destinations* | | | | | | | |
| Log Inverse Consumption PPP | -0.140*** | -0.201 | -0.419* | -0.550** | -0.715** | -0.710** | -0.351 |
| | (0.0402) | (0.205) | (0.227) | (0.222) | (0.281) | (0.301) | (0.227) |
| Log Destination GDP | 0.438*** | 0.410*** | 0.238** | 0.0699 | -0.104 | -0.102 | 0.0216 |
| | (0.0492) | (0.103) | (0.121) | (0.121) | (0.152) | (0.165) | (0.129) |
| Origin-by-Destination FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Origin-by-Period FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 25,089 | 25,089 | 20,935 | 18,328 | 16,084 | 14,361 | 12,497 |
| Number of Orig-Dest Pairs | 2899 | 2899 | 2596 | 2513 | 2265 | 2169 | 2098 |
| First Stage F-Stat | | 171.5 | 159.9 | 136.4 | 72.74 | 76.19 | 102.5 |
| *Panel B: Touristic Destinations Only* | | | | | | | |
| Log Inverse Consumption PPP | -0.114*** | -0.298 | -0.488** | -0.571** | -0.656** | -0.616* | -0.361 |
| | (0.0442) | (0.204) | (0.249) | (0.251) | (0.311) | (0.339) | (0.293) |
| Log Destination GDP | 0.402*** | 0.312*** | 0.132 | -0.00375 | -0.141 | -0.159 | -0.109 |
| | (0.0631) | (0.110) | (0.138) | (0.137) | (0.162) | (0.182) | (0.162) |
| Origin-by-Destination FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Origin-by-Period FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 17,165 | 17,165 | 14,294 | 12,535 | 11,052 | 9,874 | 8,603 |
| Number of Orig-Dest Pairs | 1981 | 1981 | 1771 | 1710 | 1511 | 1474 | 1428 |
| First Stage F-Stat | | 138.0 | 119.4 | 125.4 | 62.48 | 65.19 | 69.67 |

*Notes:* See Section 5 for discussion. Standard errors are clustered at the level of origin-by-destination pairs. * 10%, ** 5%, *** 1% significance levels.

Table 11: Spatial Labor Supply Elasticity

| Dependent Variable: | Log Municipality Employment 2000, 2010 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | Island IV & Beach IV 1 | All Five IVs | OLS | Island IV & Beach IV 1 | All Five IVs |
| Log Nominal Wage | 1.163*** | 5.425** | 4.235*** | | | |
| | (0.262) | (2.707) | (1.138) | | | |
| Log Real Wage | | | | 2.179*** | 9.446** | 7.811*** |
| | | | | (0.447) | (4.582) | (2.018) |
| Log Distance to US Border | 0.0325 | 0.469 | 0.347* | 0.0410 | 0.466 | 0.371* |
| | (0.0986) | (0.336) | (0.186) | (0.0957) | (0.325) | (0.191) |
| Log Distance to Mexico City | -0.0291 | 0.0704 | 0.0426 | 0.0532 | 0.418 | 0.336* |
| | (0.142) | (0.153) | (0.114) | (0.138) | (0.306) | (0.175) |
| Log Municipality Area | 0.297*** | 0.195* | 0.224*** | 0.306*** | 0.241** | 0.255*** |
| | (0.105) | (0.100) | (0.0788) | (0.102) | (0.0938) | (0.0780) |
| State Capital Dummy | 0.916*** | -0.393 | -0.0275 | 0.833*** | -0.636 | -0.306 |
| | (0.312) | (0.875) | (0.452) | (0.302) | (0.961) | (0.496) |
| Old City Dummy | -0.321 | -1.091 | -0.876 | -0.341 | -1.108 | -0.936* |
| | (0.479) | (0.697) | (0.533) | (0.468) | (0.681) | (0.518) |
| Colonial Port Dummy | 2.906*** | 2.451*** | 2.578*** | 2.820*** | 2.120*** | 2.278*** |
| | (0.298) | (0.444) | (0.323) | (0.285) | (0.564) | (0.363) |
| Log Average Percipitation | 0.375* | 0.702* | 0.610** | 0.349 | 0.560 | 0.513 |
| | (0.211) | (0.423) | (0.298) | (0.212) | (0.397) | (0.315) |
| Log Average Temperature | -0.313 | 0.354 | 0.168 | -0.282 | 0.431 | 0.271 |
| | (0.897) | (1.011) | (0.962) | (0.891) | (1.030) | (0.985) |
| Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 300 | 300 | 300 | 300 | 300 | 300 |
| Number of Clusters | 32 | 32 | 32 | 32 | 32 | 32 |
| First Stage F-Stat | | 3.014 | 43.03 | | 2.770 | 24.94 |
| Over-ID Test P-Value | | 0.966 | 0.588 | | 0.897 | 0.661 |

*Notes:* See Section 5 for discussion. Standard errors are clustered at the level of Mexican states. * 10%, ** 5%, *** 1% significance levels.

Table 12: The Gains from Tourism

| | Estimated | No Spillovers |
|---|---|---|
| Parameters | $\gamma_S = 0.088$<br>$\gamma_M = 0.084$ | $\gamma_S = 0$<br>$\gamma_M = 0$ |
| Gains from All Tourism | 4.42% | 4.16% |
| | (1.09, 8.12) | (2.57, 7.82) |
| | [2.52, 7.56] | [2.68, 6.57] |
| Gains from International Tourism | 1.60% | 2.43% |
| | (-0.69, 3.09) | (2.02, 3.09) |
| | [0.50, 2.86] | [2.05, 2.86] |

*Notes:* See Section 6 for discussion. 95% confidence intervals below point estimates in round brackets, and 90% confidence intervals in square brackets.

Table 13: The Role of Agglomeration Forces

| Dependent variable: | Counterfactual Change in Log Total GDP | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Parameters | $\gamma_S = 0$<br>$\gamma_M = 0$ | $\gamma_S = 0$<br>$\gamma_M = 0.15$ | $\gamma_S = 0.15$<br>$\gamma_M = 0$ | $\gamma_S = 0.088$<br>$\gamma_M = 0.084$ |
| Log Tourism GDP | 0.236*** | 0.0501*** | 0.637*** | 0.409*** |
| | (0.0575) | (0.0148) | (0.144) | (0.0928) |
| Coast FX | ✓ | ✓ | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ | ✓ | ✓ |
| Observations | 300 | 300 | 300 | 300 |
| **Gains from Tourism** | **0.0416** | **0.0111** | **0.0673** | **0.0442** |
| Number of Clusters | 32 | 32 | 32 | 32 |

*Notes:* See Section 6 for discussion. The point estimates are from an IV regression using the island instrument and the first beach instrument. Standard errors are clustered at the level of Mexican states. * 10%, ** 5%, *** 1% significance levels.

Table 14: The Local Gains from Tourism Without Labor Mobility

| Dependent variable: | Counterfactual Change in Log Local Worker Utility | |
|---|---|---|
| | (1) | (2) |
| Counterfactual | All Tourism | International Tourism |
| Parameters | $\kappa = 0$ | $\kappa = 0$ |
| Counterfactual Change in Log Tourism GDP | 0.172*** | 0.171*** |
| | (0.0496) | (0.0494) |
| Full Set of Controls | ✓ | ✓ |
| Coast FX | ✓ | ✓ |
| Observations | 300 | 300 |
| Number of Clusters | 32 | 32 |

*Notes:* See Section 6 for discussion. The point estimates are from an IV regression using the island instrument and the first beach instrument. Standard errors are clustered at the level of Mexican states. * 10%, ** 5%, *** 1% significance levels.

Table 15: The Gains from Tourism Before and After Allowing for Endogenous Transport Cost Reductions

| | Baseline Counterfactual | Allowing for 50 Percent Reduction in Transport Costs |
|---|---|---|
| Gains from Tourism | 4.42 | 4.51 |
| Gains from International Tourism | 1.60 | 1.64 |
| $\gamma_S$ Estimate | 0.088 | 0.080 |
| $\gamma_M$ Estimate | 0.084 | 0.090 |

*Notes:* See Section 6 for discussion.

Table 16: The Gains from Tourism Across Alternative Parameter Combinations

| | | Gains from Tourism | | | | | | | | Gains from International Tourism | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\sigma=\rho=1.1$ | $\sigma=\rho=1.3$ | $\sigma=\rho=1.5$ | $\sigma=\rho=1.7$ | $\sigma=\rho=2.2$ | $\sigma=\rho=2.7$ | $\sigma=\rho=3.2$ | $\sigma=\rho=3.7$ | $\sigma=\rho=1.1$ | $\sigma=\rho=1.3$ | $\sigma=\rho=1.5$ | $\sigma=\rho=1.7$ | $\sigma=\rho=2.2$ | $\sigma=\rho=2.7$ | $\sigma=\rho=3.2$ | $\sigma=\rho=3.7$ |
| $\kappa/(1-\kappa\varepsilon)=7.8$ | $\gamma_S=0$ | 8.34 | 4.04 | 2.95 | 2.46 | 1.95 | 1.75 | 1.66 | 1.63 | 1.47 | 0.93 | 0.83 | 0.78 | 0.73 | 0.69 | 0.67 | 0.65 |
| | $\gamma_S=0.028$ | 8.83 | 4.77 | 3.68 | 3.18 | 2.65 | 2.44 | 2.34 | 2.29 | 1.98 | 1.43 | 1.32 | 1.27 | 1.19 | 1.14 | 1.09 | 1.05 |
| | $\gamma_S=0.058$ | 9.11 | 5.41 | 4.37 | 3.86 | 3.3 | 3.07 | 2.97 | 2.91 | 2.32 | 1.79 | 1.67 | 1.61 | 1.51 | 1.44 | 1.37 | 1.3 |
| | $\gamma_S=0.088$ | 9.06 | 5.79 | 4.89 | $\boxed{4.42}$ | 3.85 | 3.62 | 3.49 | 3.42 | 2.29 | 1.79 | 1.67 | $\boxed{1.61}$ | 1.49 | 1.4 | 1.32 | 1.24 |
| $\kappa/(1-\kappa\varepsilon)=4.8$ | $\gamma_S=0$ | 8.77 | 4.01 | 2.92 | 2.44 | 1.95 | 1.77 | 1.69 | 1.65 | 1.58 | 1.03 | 0.92 | 0.87 | 0.82 | 0.79 | 0.77 | 0.75 |
| | $\gamma_S=0.028$ | 9.46 | 4.78 | 3.66 | 3.17 | 2.66 | 2.47 | 2.39 | 2.35 | 2.18 | 1.62 | 1.5 | 1.44 | 1.37 | 1.32 | 1.27 | 1.22 |
| | $\gamma_S=0.058$ | 10.06 | 5.5 | 4.37 | 3.87 | 3.35 | 3.15 | 3.05 | 3.02 | 2.7 | 2.15 | 2.02 | 1.96 | 1.86 | 1.78 | 1.71 | 1.63 |
| | $\gamma_S=0.088$ | 10.52 | 6.13 | 5.01 | 4.48 | 3.94 | 3.73 | 3.64 | 3.59 | 3.12 | 2.56 | 2.43 | 2.35 | 2.24 | 2.15 | 2.04 | 1.94 |
| $\kappa/(1-\kappa\varepsilon)=1.8$ | $\gamma_S=0$ | 9.25 | 3.99 | 2.91 | 2.45 | 1.98 | 1.81 | 1.74 | 1.71 | 1.73 | 1.16 | 1.05 | 1 | 0.95 | 0.92 | 0.9 | 0.87 |
| | $\gamma_S=0.028$ | 10.07 | 4.78 | 3.67 | 3.21 | 2.73 | 2.56 | 2.49 | 2.46 | 2.41 | 1.83 | 1.71 | 1.66 | 1.57 | 1.52 | 1.48 | 1.42 |
| | $\gamma_S=0.058$ | 10.88 | 5.54 | 4.43 | 3.96 | 3.48 | 3.31 | 3.22 | 3.2 | 3.06 | 2.48 | 2.35 | 2.29 | 2.2 | 2.11 | 2.04 | 1.96 |
| | $\gamma_S=0.088$ | 11.63 | 6.26 | 5.13 | 4.65 | 4.17 | 3.98 | 3.91 | 3.87 | 3.67 | 3.08 | 2.94 | 2.87 | 2.75 | 2.65 | 2.54 | 2.43 |

*Notes:* See Section 6 for discussion.

# Online Appendix

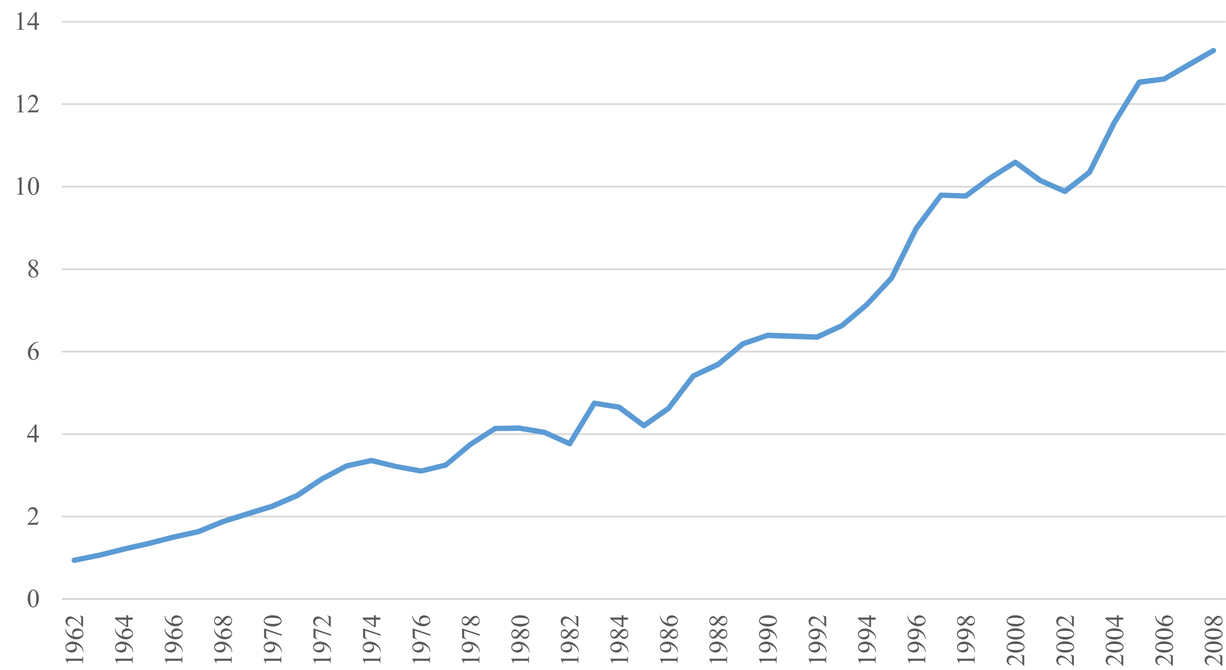
## Appendix 1: Additional Figures and Tables

Table A.1: Accommodation Share in Total Mexican Tourism Expenditure 2003-2013

| Year | Share of Accommodation in Total Tourism Expenditure |
|---|---|
| 2003 | 0.130 |
| 2004 | 0.125 |
| 2005 | 0.126 |
| 2006 | 0.124 |
| 2007 | 0.126 |
| 2008 | 0.126 |
| 2009 | 0.125 |
| 2010 | 0.127 |
| 2011 | 0.127 |
| 2012 | 0.127 |
| 2013 | 0.129 |
| Average 2003-13 | 0.127 |

*Notes:* The data source is the tourism satellite account of Mexico's national account statistics.

Figure A.1: International Tourist Arrivals in Mexico (in Millions)




*Notes:* The data source is the Mexican Secretariat for Tourism (SECTUR). The depicted time series of arrivals refers to "interior tourists" and excludes so called "border tourists" that cross the border but do not move on to Mexico's interior. For example in 2008, Mexico reported a total of 23 million international tourists of which 13.5 million were interior tourists (depicted) and 9.5 million were border tourists.

Table A.2: Wavelength Ranges Among the Top-Ranked Beaches in Mexico

| Beaches | Bandwidth 1 | | Bandwidth 2 | | Bandwidth 3 | | Bandwidth 4 | | Bandwidth 5 | | Bandwidth 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | min | max | min | max | min | max | min | max | min | max | min | max |
| Playa del Carmen | 72 | 125 | 67 | 110 | 79 | 120 | 119 | 175 | 69 | 142 | 41 | 93 |
| Tulum | 81 | 106 | 74 | 94 | 99 | 120 | 121 | 153 | 97 | 133 | 56 | 84 |
| Cozumel | 71 | 111 | 66 | 101 | 78 | 102 | 113 | 157 | 96 | 138 | 59 | 86 |
| Cancun | 81 | 111 | 72 | 101 | 74 | 102 | 38 | 149 | 15 | 125 | 7 | 71 |

*Notes:* See Section 3 for discussion. The table presents the wavelength ranges of the top four beaches in Mexico as identified by U.S. News. The data source are LandSat satellite data from 1980s and 90s at a resolution of 30x30 meters.

Table A.3: Tourism's Effect on Municipality Employment and Population: Not Using IHS Transformation

| Dependent variables: | Log Municipality Employment 2000, 2010 | | | Log Municipality Population 2000, 2010 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Baseline Specification (IHS Transformation) Both IVs | Log Hotel Sales (+1 for Zeroes) Both IVs | Log Hotel Sales (Ignore Zeroes) Both IVs | Baseline Specification (IHS Transformation) Both IVs | Log Hotel Sales (+1 for Zeroes) Both IVs | Log Hotel Sales (Ignore Zeroes) Both IVs |
| Log Hotel Sales | 0.275*** | 0.279*** | 0.255*** | 0.221*** | 0.225*** | 0.200** |
| | (0.0643) | (0.0632) | (0.0763) | (0.0686) | (0.0680) | (0.0859) |
| Log Distance to US Border | -0.00217 | -0.00742 | 0.0603 | 0.0444 | 0.0403 | 0.0871 |
| | (0.0486) | (0.0470) | (0.0569) | (0.0514) | (0.0502) | (0.0622) |
| Log Distance to Mexico City | -0.516*** | -0.537*** | -0.452*** | -0.568*** | -0.584*** | -0.487*** |
| | (0.0761) | (0.0695) | (0.0510) | (0.0809) | (0.0744) | (0.0566) |
| Log Municipality Area | 0.282*** | 0.308*** | 0.256*** | 0.343*** | 0.364*** | 0.289*** |
| | (0.0810) | (0.0726) | (0.0338) | (0.0863) | (0.0781) | (0.0376) |
| State Capital Dummy | 0.570* | 0.568* | 0.783*** | 0.540* | 0.538* | 0.728** |
| | (0.304) | (0.298) | (0.304) | (0.328) | (0.323) | (0.339) |
| Old City Dummy | 0.809** | 0.834*** | 0.980*** | 0.836** | 0.855** | 0.947*** |
| | (0.323) | (0.313) | (0.290) | (0.349) | (0.340) | (0.320) |
| Colonial Port Dummy | 0.483* | 0.463 | 0.714** | 0.589* | 0.572* | 0.824** |
| | (0.291) | (0.291) | (0.310) | (0.308) | (0.307) | (0.355) |
| Log Average Percipitation | 0.253*** | 0.263*** | 0.104* | 0.241*** | 0.249*** | 0.0968* |
| | (0.0425) | (0.0412) | (0.0536) | (0.0428) | (0.0419) | (0.0564) |
| Log Average Temperature | 0.212* | 0.213* | 0.200 | 0.273** | 0.275** | 0.220 |
| | (0.111) | (0.109) | (0.137) | (0.108) | (0.107) | (0.141) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 2,613 | 4,889 | 4,889 | 2,613 |
| Number of Municipalities | 2455 | 2455 | 1489 | 2455 | 2455 | 1489 |
| First Stage F-Stat | 11.59 | 12.09 | 16.95 | 11.59 | 12.09 | 16.95 |
| Over-ID Test P-Value | 0.617 | 0.642 | 0.528 | 0.533 | 0.550 | 0.485 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table A.4: Tourism's Effect on Municipality Population: Using 100% Census Samples

| Dependent variable: | Log Census Population 2000 and 2010 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | 10% Sample Data (IPUMS) | | 100% Sample Data (INEGI) | |
| | OLS | Both IVs | OLS | Both IVs |
| Log Hotel Sales | 0.200*** | 0.221*** | 0.200*** | 0.223*** |
| | (0.00564) | (0.0686) | (0.00563) | (0.0682) |
| Log Distance to US Border | 0.0341 | 0.0444 | 0.0300 | 0.0410 |
| | (0.0427) | (0.0514) | (0.0425) | (0.0511) |
| Log Distance to Mexico City | -0.592*** | -0.568*** | -0.590*** | -0.565*** |
| | (0.0284) | (0.0809) | (0.0283) | (0.0804) |
| Log Municipality Area | 0.370*** | 0.343*** | 0.369*** | 0.341*** |
| | (0.0171) | (0.0863) | (0.0170) | (0.0858) |
| State Capital Dummy | 0.627*** | 0.540* | 0.632*** | 0.540* |
| | (0.195) | (0.328) | (0.195) | (0.326) |
| Old City Dummy | 0.920*** | 0.836** | 0.920*** | 0.831** |
| | (0.233) | (0.349) | (0.233) | (0.347) |
| Colonial Port Dummy | 0.672*** | 0.589* | 0.673*** | 0.585* |
| | (0.143) | (0.308) | (0.143) | (0.306) |
| Log Average Percipitation | 0.245*** | 0.241*** | 0.246*** | 0.242*** |
| | (0.0407) | (0.0428) | (0.0407) | (0.0427) |
| Log Average Temperature | 0.282*** | 0.273** | 0.280*** | 0.271** |
| | (0.104) | (0.108) | (0.104) | (0.108) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 4,889 | 4,889 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 |
| First Stage F-Stat | | 11.59 | | 11.59 |
| Over-ID Test P-Value | | 0.533 | | 0.525 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table A.5: Coastal vs Inland Variation

| Dependent variable: | Log Municipality Employment 2000, 2010 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | All Municipalities | Coastal Municipalities Only | Interacted Controls |
| | Both IVs | Both IVs | Both IVs |
| Log Hotel Sales | 0.279*** | 0.279*** | 0.248*** |
| | (0.0511) | (0.0511) | (0.0708) |
| Year-By-Coast FX | ✓ | ✓ | ✓ |
| Full Set of Controls Interacted with Coast FX | | | ✓ |
| Observations | 4,889 | 297 | 4,889 |
| Number of Municipalities | 2455 | 150 | 2455 |
| First Stage F-Stat | 14.45 | 14.23 | 9.423 |
| Over-ID Test P-Value | 0.168 | 0.168 | 0.404 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table A.6: Tourism's Effect on Municipality Employment and Population: 2SLS vs LIML Estimates

| Dependent variable: | Log Municipality Employment 2000, 2010 | | Log Municipality Population 2000, 2010 | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | 2SLS | LIML | 2SLS | LIML |
| | Both IVs | Both IVs | Both IVs | Both IVs |
| Log Hotel Sales | 0.27479*** | 0.27581*** | 0.22148*** | 0.22223*** |
| | (0.06426) | (0.06549) | (0.06859) | (0.07096) |
| Log Distance to US Border | -0.00217 | -0.00168 | 0.04438 | 0.04473 |
| | (0.04856) | (0.04889) | (0.05143) | (0.05206) |
| Log Distance to Mexico City | -0.51589*** | -0.51477*** | -0.56771*** | -0.56689*** |
| | (0.07613) | (0.07739) | (0.08089) | (0.08334) |
| Log Municipality Area | 0.28172*** | 0.28046*** | 0.34301*** | 0.34210*** |
| | (0.08096) | (0.08245) | (0.08635) | (0.08921) |
| State Capital Dummy | 0.56995* | 0.56586* | 0.54029* | 0.53730 |
| | (0.30399) | (0.30792) | (0.32782) | (0.33553) |
| Old City Dummy | 0.80933** | 0.80538** | 0.83618** | 0.83329** |
| | (0.32251) | (0.32597) | (0.34902) | (0.35590) |
| Colonial Port Dummy | 0.48312* | 0.47922 | 0.58868* | 0.58582* |
| | (0.29110) | (0.29549) | (0.30760) | (0.31577) |
| Log Average Percipitation | 0.25285*** | 0.25267*** | 0.24079*** | 0.24066*** |
| | (0.04248) | (0.04257) | (0.04282) | (0.04296) |
| Log Average Temperature | 0.21193* | 0.21155* | 0.27348** | 0.27320** |
| | (0.11091) | (0.11108) | (0.10827) | (0.10852) |
| Year-By-Coast FX | ✓ | ✓ | ✓ | ✓ |
| Observations | 4,889 | 4,889 | 4,889 | 4,889 |
| Number of Municipalities | 2455 | 2455 | 2455 | 2455 |

*Notes:* See Section 3 for discussion. Standard errors are clustered at the level of municipalities. * 10%, ** 5%, *** 1% significance levels.

Table A.7: The Gains from Tourism Across Alternative Parameter Combinations

| | | Gains from Tourism | | | | | | | | Gains from International Tourism | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\sigma=\rho$ $=1.1$ | $\sigma=\rho$ $=1.3$ | $\sigma=\rho$ $=1.5$ | $\sigma=\rho$ $=1.7$ | $\sigma=\rho$ $=2.2$ | $\sigma=\rho$ $=2.7$ | $\sigma=\rho$ $=3.2$ | $\sigma=\rho$ $=3.7$ | $\sigma=\rho$ $=1.1$ | $\sigma=\rho$ $=1.3$ | $\sigma=\rho$ $=1.5$ | $\sigma=\rho$ $=1.7$ | $\sigma=\rho$ $=2.2$ | $\sigma=\rho$ $=2.7$ | $\sigma=\rho$ $=3.2$ | $\sigma=\rho$ $=3.7$ |
| $\kappa/(1-\kappa\varepsilon)=7.8$ | $\gamma_M=0$ | 12.12 | 7.56 | 6.42 | 5.9 | 5.35 | 5.14 | 5.04 | 5.01 | 4.45 | 3.86 | 3.71 | 3.63 | 3.47 | 3.32 | 3.16 | 2.99 |
| | $\gamma_M=0.024$ | 11.19 | 6.99 | 5.86 | 5.33 | 4.78 | 4.56 | 4.46 | 4.41 | 3.79 | 3.22 | 3.07 | 3 | 2.85 | 2.72 | 2.59 | 2.44 |
| | $\gamma_M=0.054$ | 10.11 | 6.38 | 5.33 | 4.81 | 4.23 | 4 | 3.9 | 3.84 | 3.05 | 2.51 | 2.38 | 2.3 | 2.18 | 2.06 | 1.96 | 1.84 |
| | $\gamma_M=0.084$ | 9.06 | 5.79 | 4.89 | $\boxed{4.42}$ | 3.85 | 3.62 | 3.49 | 3.42 | 2.29 | 1.79 | 1.67 | $\boxed{1.61}$ | 1.49 | 1.4 | 1.32 | 1.24 |
| $\kappa/(1-\kappa\varepsilon)=4.8$ | $\gamma_M=0$ | 12.79 | 7.78 | 6.63 | 6.12 | 5.6 | 5.41 | 5.32 | 5.29 | 4.84 | 4.24 | 4.09 | 4 | 3.83 | 3.68 | 3.51 | 3.33 |
| | $\gamma_M=0.024$ | 12.03 | 7.18 | 6.03 | 5.53 | 5 | 4.8 | 4.71 | 4.67 | 4.23 | 3.65 | 3.5 | 3.42 | 3.27 | 3.14 | 3 | 2.84 |
| | $\gamma_M=0.054$ | 11.22 | 6.6 | 5.45 | 4.94 | 4.41 | 4.2 | 4.11 | 4.07 | 3.63 | 3.05 | 2.91 | 2.83 | 2.7 | 2.59 | 2.47 | 2.34 |
| | $\gamma_M=0.084$ | 10.52 | 6.13 | 5.01 | 4.48 | 3.94 | 3.73 | 3.64 | 3.59 | 3.12 | 2.56 | 2.43 | 2.35 | 2.24 | 2.15 | 2.04 | 1.94 |
| $\kappa/(1-\kappa\varepsilon)=1.8$ | $\gamma_M=0$ | 13.71 | 8.19 | 7.04 | 6.56 | 6.08 | 5.9 | 5.82 | 5.79 | 5.46 | 4.84 | 4.69 | 4.59 | 4.42 | 4.25 | 4.06 | 3.84 |
| | $\gamma_M=0.024$ | 12.97 | 7.5 | 6.37 | 5.89 | 5.4 | 5.22 | 5.14 | 5.11 | 4.83 | 4.22 | 4.07 | 3.98 | 3.83 | 3.69 | 3.53 | 3.35 |
| | $\gamma_M=0.054$ | 12.23 | 6.81 | 5.69 | 5.2 | 4.71 | 4.54 | 4.46 | 4.43 | 4.19 | 3.58 | 3.44 | 3.37 | 3.24 | 3.11 | 2.99 | 2.84 |
| | $\gamma_M=0.084$ | 11.63 | 6.26 | 5.13 | 4.65 | 4.17 | 3.98 | 3.91 | 3.87 | 3.67 | 3.08 | 2.94 | 2.87 | 2.75 | 2.65 | 2.54 | 2.43 |

*Notes:* See Section 6 for discussion.

## Appendix 2: Counterfactuals With Less Aggregated Interior Regions

In this appendix we provide additional results to investigate the sensitivity of the welfare quantifications with respect to more or less regional aggregation for the interior regions of Mexico. As we discuss in Section 5, the 2455 regions case pushes the limits of the computational requirements when it comes to our indirect inference approach (involving a grid of parameter combinations) as well as for the bootstrapping of the computation of counterfactual equilibria several hundreds of times.

To this end, we document the welfare gains from tourism as well as the regional effects of tourism across regions when running counterfactuals based on the disaggregated 2455 regions case compared to the baseline 300 regions case that we work with in the main text. In particular, we use the same model parameter values as in our preferred counterfactuals, but solve the model for counterfactual no-tourism equilibria in both the more and less aggregated scenarios.

Tables A.8 and A.9 report the quantification results back-to-back. Reassuringly, we find very similar estimates of the welfare gains from tourism and international-only tourism, and we also find that the regional implications of tourism are remarkably similar across the two levels of regional aggregation. As discussed in the main text, these results are as expected, because the key source of variation that we use to inform the calibration of the model and its parameters stems from coastal municipalities. The aggregation of interior municipalities into larger regions that are centered around the 150 largest economic centers –while keeping the coastal geography as in the reduced-form analysis– is thus greatly convenient for computational power, but largely inconsequential for the estimated results.

Table A.8: The Gains from Tourism With Different Numbers of Regions

| | 300 Mexican Regions | | 2455 Mexican Regions | |
|---|---|---|---|---|
| | Estimated | No Spillovers | Estimated | No Spillovers |
| Parameters | $\gamma_S = 0.088$<br>$\gamma_M = 0.084$ | $\gamma_S = 0$<br>$\gamma_M = 0$ | $\gamma_S = 0.088$<br>$\gamma_M = 0.084$ | $\gamma_S = 0$<br>$\gamma_M = 0$ |
| Gains from All Tourism | 4.42% | 4.16% | 4.63% | 4.62% |
| Gains from International Tourism | 1.60% | 2.43% | 1.52% | 2.70% |

*Notes:* See Section 6 for discussion.

Table A.9: The Regional Implications of Tourism With Different Numbers of Regions

| Dependent variable: | Counterfactual Change in Log Total GDP | |
|---|---|---|
| | (1) | (2) |
| | 300 Mexican Regions | 2455 Mexican Regions |
| Log Tourism GDP | 0.409*** | 0.385*** |
| | (0.0928) | (0.0942) |
| Coast FX | ✓ | ✓ |
| Full Set of Controls | ✓ | ✓ |
| Observations | 300 | 2455 |
| **Welfare Gains** | **0.0442** | **0.0463** |
| Number of Clusters | 32 | 32 |

*Notes:* See Section 6 for discussion. The point estimates are from an IV regression using the island instrument and the first beach instrument. Standard errors are clustered at the level of Mexican states. * 10%, ** 5%, *** 1% significance levels.