Urban Welfare: Tourism in Barcelona[*]

Treb Allen

Dartmouth and NBER

Simon Fuchs

Federal Reserve Bank of Atlanta

Sharat Ganapati

Georgetown University and NBER

Alberto Graziano

CaixaBank Research

Rocio Madera

Southern Methodist Univ.

Judit Montoriol-Garriga

CaixaBank Research

November 30, 2021 (Preliminary and Incomplete)

Abstract

We develop a simple methodology to estimate the heterogeneous welfare effects of any small shock to residents within a city. The methodology relies only on modest assumptions regarding residents' choice of where to consume and work and delivers an expression that shows that the welfare elasticity to any small shock can be written as a function of (1) the spatial patterns of consumption and income; and (2) the price and wage effects of the shock. We then apply this methodology to ask the question: Is tourism good for locals? Using detailed spatial data on expenditure and income patterns of residents in Barcelona, we show that plausibly exogenous shifts in tourist expenditure due to compositional differences in their country of origin across time and over space in the city crowds out local expenditure by increasing prices but partially compensates through increases in wages. The incidence of the tourism shock, however, is highly heterogeneous across the city, with inner city residents bearing the the largest welfare losses and peripheral residents enjoying the greatest welfare gains.

JEL classification: R12, R13, R22, R23

Keywords: Tourism, Spatial Consumption, Spatial Equilibrium, Urban Spillovers, Urban Welfare.

1 Introduction

Over recent years, new data sources have become available that describe urban activity in an astonishing level of detail. At the same time, the development of powerful quantitative spatial equilibrium models enables the calculation the welfare impacts of policies while accounting for the observed micro-geographic distribution of economic activity. In order to remain tractable, however, these spatial models rely on strong parametric assumptions regarding e.g. the distributions of preferences and functional forms of production. To what extent do the quantitative implications of these models depend on the underlying assumptions? And – more importantly – is it possible to infer the welfare effects of urban shocks without such assumptions and without requiring the full estimation and the solution of a high-dimensional structural model?

This paper develops a new methodology for estimating the distribution of welfare effects across the city for (small) urban shocks. By combining an envelope-type argument for resident's optimization over consumption and commuting choices and a simple market clearing condition, we show how one can calculate the (short-run) welfare impacts of (small) urban shocks using a regression based approach and without the need for strong functional form assumptions or the estimation of a full structural model. We then combine our methodology with a new high-resolution spatial dataset on consumption and income patterns in Barcelona to ask the question: "Is tourism good for locals?". Using a novel identification strategy based on the fact that tourists from different origins vary both in when they visit Barcelona and where they consume, we show that tourist expenditure crowds out local expenditure by increasing prices but partially compensates through increases in wages. The incidence of the tourism shock, however, is highly heterogeneous across the city, with inner city residents bearing the the largest welfare losses and peripheral residents enjoying the greatest welfare gains.

Our proposed methodology is based on two insights: First, as long as residents optimally choose where to work and consume, a simple envelope result shows that the welfare effect of any (small) shock depends only on (1) the spatial patterns of consumption and income; and (2) the price and wage effects of the shock throughout the city. Intuitively, the welfare impact of a shock depends on whether it increases a residents' total income more than it increases the prices she pays for her goods, where the income and price shocks depend on the particular income and consumption patterns of the resident. Second, a simple market clearing condition allows one to trace out both

the direct effect of the shock on prices and wages as well as its indirect general equilibrium effects as it dissipates throughout the city. Intuitively, a shock that directly increases the income of one resident will indirectly increase the income of other residents working where she consumed (and so on). By combining these two insights, one can estimate the welfare effects of a shock – incorporating both the observed micro-geographic distribution of economic activity and the general equilibrium interactions across the city – using a simple regression based methodology.

We apply our methodology to analyze the impact of tourism on residents of Barcelona. In many locations around the world, tourism comprises a substantial and growing fraction of the economy. In Spain, tourism is the largest export sector and the second fastest growing sector of the economy, with tourism expenditures currently equal in value to half of all Spain's exports of goods (and 11% of GDP in total). In Barcelona, the number of tourists have approximately doubled over the past decade. While some in Barcelona argue that this increase in foreign demand has been a boon for local businesses, the increase in tourism – and resulting increases in prices and congestion – have also led to protests and calls for policies discouraging tourism, such as the imposition of new taxes.

To employ our proposed methodology, we assemble a new high-resolution spatial dataset on consumption and income patterns in the city of Barcelona based on hundreds of millions of credit and debit card transactions and covering roughly 3% of the entire Barcelona metro area GDP. The transactions are from two sources: (1) purchases made at a point-of-sale owned by the bank; and (2) purchases made by customers of the bank. The former allows us to construct monthly level expenditure for 20 different product categories across 1,095 locations within the Barcelona metro area for tourists from each non-Spanish country. The latter allows us to construct *bilateral* expenditure share matrices for Barcelona residents by both their location of residence and the location and category of purchase. (It also also allows us to construct expenditure data for non-local Spanish tourists). To account for the fact that not all purchases are made using a credit or debit card, we append additional housing rental data and re-weight the expenditures by product category to match aggregate expenditure surveys. From the same source we also obtain detailed information about residential month-by-month income. We combine this detailed expenditure data with commuting data at the same spatial resolution, which we construct by cross-referencing from cell-phone location data with commuting survey data.

Using this dataset, we document several new and salient facts. First, different tourist groups

arrive at different times in the city, and the relative popularity of locations depends importantly on the country of origin of the tourist. This fact proves particularly useful, as it allows us to use aggregate variation in the composition of tourists in the city at a given time – driven e.g. by differences in timings of school breaks in the origin countries – to generate variation in tourism expenditure that is plausibly orthogonal to unobserved changes in local conditions. Second, we document that both local expenditure and income has a strong spatial component, where locals are much more likely to purchase goods and work nearby their residence. Combined with the first stylized fact, this implies that residents living closer to places popular with tourists will be more exposed to tourism. Third, comparing the tourism "low" and "high" seasons within a year, we show that total sales increase more in locations popular with tourists where tourism expenditure increases but that local expenditure falls the most in these same locations, suggesting that tourism both increases incomes earned by locals but also increases the prices locals pay for goods.

We combine this source of plausibly exogenous tourist expenditure with our novel methodology to trace out the welfare implications on residents in different parts of the city. We document substantial and heterogeneous effects on prices from tourism that depend both on how important tourism is locally as well as indirect exposure to tourism shocks in other cells through the network structure of the economy. The impact of income is also sizable and heterogeneous across the city with the inner-city dwellers benefiting the most. We also examine house prices and show a substantial pass-through of tourist expenditures into housing prices.

Finally, we impose functional form assumptions on production, consumption, and commuting to build a structural Ricardo-Viner specific factors trade model embedded into an urban setting with a rich geographic patterns of consumption and commuting. The purpose of this quantitative framework is two-fold: first, it allows us to compare our welfare estimates from our regression based approach to those more typical in a quantitative literature; and, second it allows us to estimate the welfare effect of counterfactual policies. We show that the estimates from our proposed methodology closely align with the quantitative welfare results. We furthermore use the model to determine an optimal place-varying tourist tax that correct for the heterogeneous incidence of the shock.

This paper makes three primary contributions to the literature. First, we provide an estimate of the spatially heterogeneous welfare impact of tourism on locals throughout a city. While several recent papers have examined the impact of tourism on local housing markets and consumption

amenities (e.g. Almagro and Domínguez-Iino (2019) and García-López et al. (2019)), they have tended to abstract from spatial linkages (through either commuting or consumption) within the city, instead treating different neighborhoods as independent locations. Here, we explicitly model these linkages and show they play an important role generating heterogeneity in wage and price effects across the city. In this way, the paper is closely related to Faber and Gaubert (2019), who show that the welfare impact of tourism depends importantly on spatial and sectoral linkages, albeit across regions within a country instead of neighborhoods within a city.

Second, we characterize the welfare implications of urban shocks across connected neighborhoods building on and extending results from the trade and consumption literature that relies on first-order approximations of the indirect utility function to express welfare changes in terms of expenditure share weighted compensating variations (Houthakker, 1952; Atkin et al., 2018; Porto, 2006; Deaton, 1989; Kim and Vogel, 2020). Despite complex spatial consumption and commuting patterns, it turns out that all one needs to calculate the welfare impact of any economic shock is: (1) the spatial patterns of consumption and income of residents; and (2) how the shock affects prices and wages throughout the city. As large-scale spatial data sets become increasingly available, (see e.g. Athey et al. (2018) and Couture et al. (2020) for examples using mobile phone data, Davis et al. (2019) for an example using online review data, and Carvalho et al. (2020) and Agarwal et al. (2017) examples using credit card transaction level data), we expect the first ingredient will become increasingly attainable.

The third contribution of the paper is to propose a new empirical methodology that marries recent advances in the *quantitative* spatial literature with recent advances in the *applied* spatial literature. While the seminal paper of Ahlfeldt et al. (2015) introduced general equilibrium counterfactual analysis to urban models with complex geographies, retaining tractability required making particular functional form assumptions on preferences, commuting, and production. Here, instead we relax those assumptions. Imposing only market clearing conditions, we provide new and intuitive analytical expressions for how prices and wages respond to a shock in the short-run (relying on tools introduced by Allen et al. (2020) for trade models and closely related to work by (Baqaee and Farhi, 2019; Kleinman et al., 2020; Baqaee and Burstein, 2021)). These predictions can be constructed from observable data and integrated into our empirical regressions to trace out the heterogeneous treatment effect of an urban shock on prices and wages. Implicitly, in the spirit of Donaldson (2018)

and Monte et al. (2018), this examines the extent to which the theoretically predicted price and wage effects are able to capture the observed empirical variation in prices and wages.

The rest of the paper is organized as follows. The next section describes our result regarding welfare effects of small shocks that motivates our methodology. Section 3 introduces the data and our empirical strategy. Section 4 presents both the average and heterogeneous treatment effects. Section 5 introduces the quantitative model and Section 6 concludes.

2   A Tractable Urban Model for Welfare Evaluations of Small Shocks

This section describes our main theoretical results that underlie our empirical methodology. In section 2.2 we begin by deriving under minimal assumptions a general expression for the change in welfare as a function of an arbitrary urban shock. We then proceed to show in section 2.3 that market clearing implies closed-form expressions for the (short-run) price and wage elasticities, which can be readily constructed from data.

2.1   Setting

Consider a city consisting of many city blocks which are index by $i, n \in \{1, \ldots, N\}$. In the following n will refer to the location of residence, while i will refer to the location of production and consumption[1]. Residents in each city block are endowed with a total time endowment of T_n. The city blocks are fixed in their geography to each other.

2.2   A Simple Envelope Theorem for Welfare Evaluations

The representative resident makes two choices, first how to allocate labor across production locations in order to maximize her income and second, given income and prices, she chooses her consumption to maximize her utility. The representative resident has homothetic demand that can be represented by the following indirect utility function:

$$u_n = \frac{v_n}{G\left(\boldsymbol{p}_n\right)}$$

---

[1]This implies that residents choose which of N goods to consume and produce. We can similarly think of a good as a locations $\times$ sector pair, in which case we would have $N \times S$ goods, where S is the number of sectors. We abstract from explicit sector references and indices in this section, but the results are completely general to allow for location-sector-indexed goods.

where $G(\cdot)$ represents some price aggregator and v_n is the income of the resident and $\boldsymbol{p}_n$ refers to the vector of prices faced by residents in block n of the goods produced by each block i. Consumption is assumed to require physically traveling to a city block and incurring an iceberg variable trade costs such that prices are given by

$$p_{ni} = \tau_{ni} p_i$$

where $\tau_{ni} \geq 1$ refers to the iceberg variable trade costs and where we assume that $\tau_{nn} = 1$. From Roy's identity, demand in block n for the good produced in country i is given by,

$$q_{ni}\left(\boldsymbol{p}_n\right) = -G\left(\boldsymbol{p}_n\right) v_n \times \frac{\partial\left(1/G\left(\boldsymbol{p}_n\right)\right)}{\partial p_{ni}}$$

The labor supply decision by the representative resident in block n is defined by an income maximization problem. Each location has a total time endowment of T_n and the representative resident solves a problem of how to best allocate labor across the city given a constraint on his labor allocation $H(\cdot)$:

$$v_n = \max_{\{\ell_{ni}\}} \sum_{i=1}^{N} w_i \ell_{ni}$$

$$\text{s.t. } H_n\left(\{\ell_{ni}\}\right) = T_n$$

where $H(\cdot)$ is a weakly convex function that captures potential decreasing returns to scale to allocating more labor to the production of a particular product, which we refer to as commuting friction. Define the income function $y(\boldsymbol{w}_n, T_n)$ as the maximand of this problem and applying the envelope theorem, we have,

$$\frac{\partial y(\cdot)}{\partial w_i} = \ell_i \tag{1}$$

The dual of the problem is given by commuting cost minimization problem subject to a fixed income level, where the function $h\left(\boldsymbol{w}_n, \bar{Y}\right)$ is the minimand of the constrained optimization problem that minimizes the time input subject to reaching at least the income level $\bar{Y}$. Differentiating this expression and solving for the partial impact of wages on income, we can derive a Roy's identity for the commuting problem which is given by,

$$\frac{\partial y(\cdot)}{\partial w_i} = -\frac{\frac{\partial h(\boldsymbol{w}_n, y(\boldsymbol{w}_n, T_n))}{\partial w_i}}{\frac{\partial h(\boldsymbol{w}_n, y(\boldsymbol{w}_n, T_n))}{\partial y}} = \ell_i \tag{2}$$

where in the final equality we have applied the envelope theorem from equation 1. Under our assumptions on the function $H(\cdot)$ we can write the time use function as,

$$h_n = \frac{v_n}{J(\boldsymbol{w}_n)}$$

where $J(\boldsymbol{w}_n)$ is the commuting equivalent of the price aggregator on the consumption side and we refer to it as a wage aggregator and v_n denotes as before the disposable income. Applying Roy's identity, we have,

$$\ell_{ni}(\boldsymbol{w}_n) = v_n \frac{1}{J(\boldsymbol{w}_n)} \frac{\partial(J(\boldsymbol{w}_n))}{\partial w_i}$$

Assuming both homothetic demand and a homothetic income maximization problem allows us to write the indirect utility function as,

$$u_n = \frac{T_n J(\boldsymbol{w}_n)}{G(\boldsymbol{p}_n)}$$

where indirect utility is solely a function of the price and wage aggregator, as well as the exogenously given time endowment across locations. Fully differentiating and applying both the consumption and commuting Roy's identity we obtain,

$$d\ln u_i = \sum_{i=1}^{N} c_{ni} \partial \ln w_i - \sum_{i=1}^{N} s_{ni} \partial \ln p_{ni} \tag{3}$$

where $s_{ni} = p_i q_{ni}/(\sum_j p_j q_{nj})$ is the bilateral expenditure share and $c_{ni} = w_i \ell_{ni}/(\sum_j w_j \ell_{nj})$ is the income share derived from supplying labor that originates in residential tile n and works in location i. Equation (3) forms the basis of the analysis that follows. It shows that there are two necessary ingredients in order to determine the welfare impact of any shock (including tourism) on local residents. The first necessary ingredient is knowledge of existing income shares (i.e. $\{c_{ni}\}$) and expenditure shares (i.e. $\{s_{ni}\}$). This requires knowledge of disaggregated spatial expenditure patterns as well as the patterns of labor supply within cities. Fortunately for us, we have access to detailed commuting flow data that can be used to reconstruct income shares and detailed expenditure data that records the location of the residence as well as the location and sector of the purchase to reconstruct expenditure shares.

The second necessary ingredient is knowledge of how a shock changes wages (across locations

7

within the city) and prices (across all goods, i.e. locations×sectors, within the city). While in principal these wage and price changes are estimable, in practice it is infeasible to simultaneously estimate N distinct wage changes and $N \times S$ distinct price changes. In what follows, we pursue two complementary strategies to overcome this limitation: first, in results presented in Section 4.1, we empirically estimate average price and wage elasticities across all locations in the city, which allows us to recover an average welfare impact of tourism; second, we impose simple equilibrium market clearing conditions to derive expressions for all wage and price changes throughout the city as a function of observed data. We turn to these derivations next.

2.3   A General Expression for Price and Wage Effects of a Demand Shock

The results thus far have relied solely on the optimization on the part of residents; as a result, we have not needed to impose any general equilibrium conditions for the economy as a whole. Here, we show that imposing standard market clearing conditions allow us to trace out the impact of a demand shock on prices and wages throughout the city.

We model a demand shock as exogenous expenditures E^T on goods produced in the city. We now derive how an (exogenous) increase in E^T – in our setting a tourism shock – affects prices and wages throughout the city. For the time being, we hold labor allocations and expenditure shares fixed, which we refer to as the "short-run." This is perhaps an appropriate assumption given that our empirical context examines impacts of tourism by comparing expenditure across months within a year; however, in the next section below we extend the framework to allow for adjustments to local labor allocations as well. Denoting as before v_n as the income of residents and s_{ni} and s_i as the expenditure shares of residents and the demand shock respectively, we have for the market clearing condition,

$$y_i = p_i q_i = \sum_{n=1}^{N} s_{ni} v_n + \sum_{g=1}^{G} s_i^g E_g^T$$

Totally differentiating this market clearing condition and substituting for the expression of log changes in disposable income we obtain,

$$\mathrm{d}\ln \mathbf{p} = \mathbf{TC}\,\mathrm{d}\ln \mathbf{w} + \mathbf{T_T}\,\mathrm{d}\ln \mathbf{E^T}$$

where $\mathbf{T_T} \equiv \left[\frac{s_{gi} E_g^T}{y_i} \right]_{ig}$ still summarizes the direct effect from tourist expenditures on prices, and

where $\mathbf{TC} \equiv \left[\sum_{n=1}^{N} \frac{s_{ni}v_n}{y_i} \times c_{nj}\right]_{ij}$ takes commuting linkages into account and traces out to what extent demand in location i is dependent on disposable income at residential location n and in turn to what extent residential income in location n depends on wages in location j.

Similarly, we can derive an expression for the change of wages, keeping labor allocations and therefore supply fixed and assuming furthermore that expenditure shares do not adjust. Starting from the factor market clearing condition,

$$\frac{w_i \ell_i}{\theta_i^{\ell}} = \sum_{n=1}^{N} s_{ni}v_n + \sum_{g=1}^{G} s_{gi}E_g^T$$

where θ_i^{ℓ} represents the output elasticity to labor. Totally differentiating, we obtain, the matrix representation,

$$\mathrm{d}\ln\mathbf{w} = \mathbf{M}\mathrm{d}\ln\mathbf{w} + \mathbf{M_T}\mathrm{d}\ln\mathbf{E^T}$$

where $\mathbf{M_T} \equiv \left[\frac{s_{gi}E_g^T}{y_i}\right]_{gi}$ summarizes the direct effect from tourist expenditures on wages, and where $\mathbf{M} \equiv \left[\sum_{n=1}^{N} s_{ni} \times \frac{v_n}{y_i} \times c_{nj}\right]_{ij}$ captures to what extent local demand depends on income increases in residential location n and to what extent disposable income depends on wage increases in location j. We can solve for the net effect on wages using standard matrix algebra,

$$\mathrm{d}\ln\mathbf{w} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{M_T}\mathrm{d}\ln\mathbf{E^T}$$

Using this result, we can re-express the Leontief inverse in a Neumann series and obtain an expression for the elasticity of wages with regard to group specific tourist expenditure shocks,

$$\frac{\partial \ln p_{is}}{\partial \ln E^T} = \frac{s_{gi}E_g^T}{y_i} + \sum_{n} s_{nis} \frac{v_n}{y_i} \sum_{j} c_{nj} \frac{\partial \ln w_j}{\partial \ln E^T} \tag{4}$$

$$\left[\frac{\partial \ln w_i}{\partial \ln E_g^T}\right]_i = (\mathbf{I} - \mathbf{M})^{-1} \left[\frac{s_{gi}E_g^T}{y_i}\right]_i = \frac{s_{gi}E_g^T}{y_i} + \frac{1}{y_i} \sum_{j} \sum_{n} s_{ni} \times v_n \times c_{nj} \times \left(\frac{s_{gi}E_g^T}{y_j}\right) + \dots \tag{5}$$

Equation (4) shows that the direct impact of a tourism shock on wages is to increase demand for goods produced in location, which it does in proportion to tourist's initial share of expenditure. However, tourist shocks also have indirect effects on demand (and hence wages). The first degree

indirect effect is that some of the the direct impact on wages elsewhere translate in changes in demand; for example, a local resident who earns additional tourist income (the direct effect) will spend that dollar elsewhere in the city (the first degree effect). The first degree indirect effect then in turn generates a second degree indirect effect (as another resident paid by the first resident spends that additional income elsewhere), and the process repeats ad infinitum, convering to the expression in equation (5).
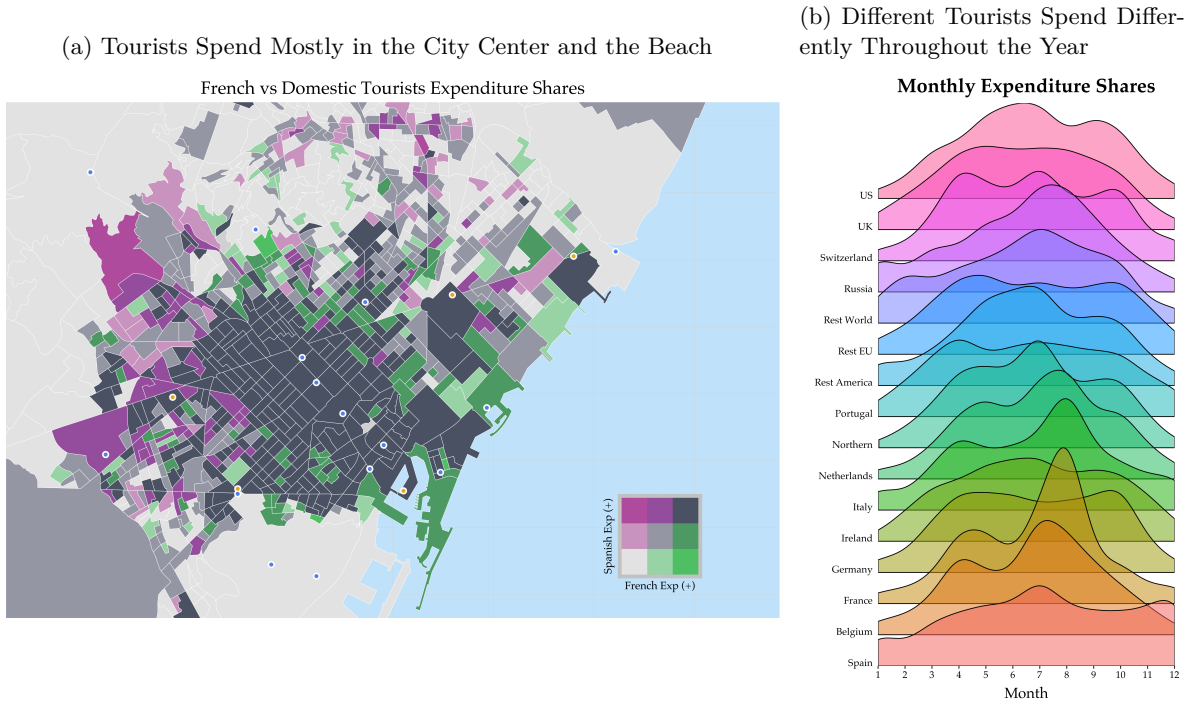
There are two key take-aways from equations (4) and (5). First, how exactly a tourism shock changes prices and wages throughout the city depend on the interaction of the spatial patterns of income and consumption of local residents. Second, given knowledge of these spatial patterns (along with knowledge of the spatial pattern of consumption by tourists), one can apply equations (4) and (5) to determine the price and wage impacts of a tourism shock. This, in turn, can be combined with equation (3) to allows us to determine the (short-run, first-order) welfare impacts of tourism solely as function of observed data, an approach we pursue below in Section 4.2.

3   Empirical Context and Data

Tourism is a key sector in Spain, and in Barcelona in particular. In 2018, 19.12 million foreign tourists visited the region of Catalonia (Idescat 2019), approximately doubling in a decade. On average, each tourist spent €185 per day, totaling €20.6 billion in declared expenditures. That tourism is a large and positive aggregate income shock for many urban enclaves like Barcelona seems uncontroversial. However, the incidence of such an income shock for the millions of residents in the city is still an open question, in part due to limited availability of economic activities at the microgeographic level. While it is safe to assume that the largest share of tourism expenditures is spent in a few hotspots of Barcelona, consumption and commuting patterns of residents create a network through which highly localized shocks interact and propagate heterogeneously.

For this project, we draw on multiple data sources that describe in fine geographical detail the economic activities of tourists and locals within the city of Barcelona. The core data source is a new expenditure database that is constructed from electronic payments processed by CaixaBank, the largest bank in Barcelona. In this section, we introduce this database that describes in much detail expenditures in Barcelona by census tract, both by tourists and locals. We combine the expenditure data with information on locals' checking accounts, commuting flows, and rental and housing prices

Figure 1: Tourist Expenditures

(a) Tourists Spend Mostly in the City Center and the Beach

(b) Different Tourists Spend Differently Throughout the Year

Notes: **(Panel A)** This map compares the spatial patterns of foreign and Spanish tourists, where expenditure is measured as total annual euros spent per scare meter. Areas with high Spanish tourist expenditure and low foreign tourist expenditure are indicated in purple, whereas areas with high foreign tourist expenditure and low Spanish tourist expenditure are indicated in green. Expenditure measured in average monthly expenditure per square meter of the underlying tile. The boundary points for the tertiles are given by ([0,2.9],[2.9,14.4],[14.4,218.8]) for tourist expenditure changes and by ([0,1.4],[1.4,4.9],[4.9,253.6]) for local expenditure changes. **(Panel B)** This ridge plot shows the density of tourist expenditure within the year, for each of the 16 tourist groups in our sample (15 foreign plus rest of Spain).

into a high-resolution spatial monthly panel dataset. We will use this novel data source to motivate and discipline our theoretical framework, our empirical strategy, and our quantitative exercise.

3.1 New High-Resolution Spatial Panel Data

Expenditure Data for Locals and Tourists

We use transaction-level data from the electronic payments that were submitted to CaixaBank's Payment Processing Service. CaixaBank is the leading bank among individuals and SMEs in Spain and is based in Barcelona, where it has close to a 40% market share. The underlying data contains each debit or credit card purchase at any merchant with a CaixaBank Point of Sale (PoS) in the city of Barcelona. For each transaction, the total euro amount, the exact merchant geo-localization, the expenditure category, the country of origin of the paying credit card, as well as the time and

date when it happened, are recorded. Importantly for us, if the customer is a CaixaBank client herself, her home address is additionally registered, allowing us to trace out the spatial expenditure pattern of a residential location in Barcelona. Our data of analysis consist of the total value of the full set of these transactions per month and census-tract in the Barcelona metropolitan area (Àmbit metropolità de Barcelona), further disaggregated by merchant category, type of customer (resident or tourist, and subgroups within), and origin location of the customer (census-tract if resident and country if tourist).

To put the scope of our data in context, we have over 165 million yearly observations adding up to a total value of 2,970 million euros. There are 1068 census tracts in the city of Barcelona proper and we further include the 27 municipalities that form the metropolitan area of Barcelona (AMB). Our data span from January 2017 through December 2019. We define five customer groups based on residence status and data availability (CaixaBank relation): (1) residents that are CaixaBank customers, (2) residents that are not CaixaBank customers, (3) domestic tourists that are CaixaBank customers, (4) domestic tourists that are not CaixaBank customers, (5) and foreign tourists. The latter group can be further disaggregated into 15 subgroups based on the country of origin of the credit card used. We document detailed destination-level descriptive analytics for all groups. In our benchmark analysis, however, we define resident transactions as those originating from members of group (1), and tourist transactions as those originating from members of groups (3) through (5).

To ease concerns about the representativeness of expenditures taken from electronically processed transactions only (as opposed to cash transactions), we have created a crosswalk from the original coding in terms of merchant categories to COICOP - a classification of expenditures into consumption categories commonly used for expenditure surveys. This allows us to directly compare the expenditure shares in our sample to the national expenditure survey in Spain as is down in table 7. Our dataset corresponds to 54.4pc of the expenditures observed in national expenditure survey. The weights on individual categories is roughly comparable to the results of the expenditure survey, but far from exactly matching it.

Housing Prices

While detailed in many aspects, our expenditure database has limited information on the housing expenditure of locals, a category of expenditures not commonly payed using debit or credit cards in Spain. We therefore employ an additional database on local rental rates and housing prices across Barcelona that we obtained from *Idealista*, a Spanish real estate marketplace comparable to *Zillow*. Idealista imputes rental rate trends at a monthly frequency for neighborhoods (Barrios) in Barcelona. Neighborhoods contain multiple census sections and therefore the geographical resolution of these data is coarser than our expenditure data, but nevertheless it is sufficient to capture key trends and cross-sectional variation in the Barcelona housing market, inlcuding rental and housing prices. The sample for the housing data covers the period between January 2010 and June 2020 at a monthly frequency.

Income Data for Locals

For locals, we combine the expenditure data with labor income information. The source of the income data is the anonymized monthly payrolls and unemployment benefits deposited in personal bank accounts in CaixaBank, aggregated at the census tract level and for the same time period for which we observe expenditures. CaixaBank has the highest market share in direct-deposit payrolls in the country (27.1%). See Aspachs et al. (2020) for more details on the underlying microdata, including sample representativeness.I

Commuting Patterns of Locals

Commuting data has traditionally been a cornerstone of applied urban analysis. Traditionally, surveys have been used to inform our understanding of employment linkages across the urban landscape. More recently, additional data on urban mobility has become available exploiting the spatial extent of the cell phone network. As cell phones and their owners chart their path through the city, they continue to switch between cell phone towers, logging into the closest one to optimize network coverage. This leaves a data trail of time stamps and associated towers for each cell phone which in turn can be used to impute the spatial path of cell phones through the day. In the light of the Covid-19 crisis, the national statistical agency in Spain has acquired data from the most important network operators in Spain to impute the mobility patterns of residents across Spain.

They also released a benchmark dataset that describes mobility patterns on November 18th, 2019. We use this data to inform our analysis. The data describes the flows between spatial units. The geographical aggregation of this data is somewhat coarser, being coded at the neighborhood (barrio) level. An additional challenge is that for privacy reasons INE does not report bilateral flows that are less than 100 in absolute magnitude.

To complement the coarser cell phone data in the task of tracing the propagation of the tourist income shocks from workplace to residence, we also construct commuting flows using a subset of the CaixaBank electronic-payments transactions described above. For the group of residents of which we observe the residential location of the account holder, we isolate their lunchtime (1-4pm) restaurant expenditures on weekdays. Assuming that lunchtime expenditures are very proximate to the place of work, this strategy can isolate commuting flows. An interesting advantage of this approach is that this can be done at the same geographic resolution as the expenditure data, i.e. it allows us to recover commuting patterns between census blocks.

3.2 Tourism in the City: Three Stylized Facts

In this section, we use our data sources to document three stylized facts: (1) tourism varies across space and time within the city; (2) locals' consumption and income exhibits strong spatial patterns localized around their place of residence; and (3) tourism appears to crowd out local consumption but increase total spending, consistent with it having both price and wage effects.

Fact 1: Tourism varies across space and time within the city

1(a) Tourism is spatial: the size of total tourist spending varies substantially across space. Figure 8 shows the intensity of tourist expenditures across individual locations for a given year. The intensity is normalized by the area of the underlying tile to account for heterogeneity in the size of individual census blocks. For convenience, we also show - with blue labels - the location of 15 of the most popular tourist sites in Barcelona. Not surprisingly the expenditure of tourists is closely correlated with the location of the main tourist attraction. The historical medieval core of Barcelona together with its extension towards Gracia forms an axis of intense touristic activity, with additional hotspots close to La Sagrada Familia and along the beachfront. Notice, that our data is sensitive to the intensity of commercial activity across different locations. We indicate with yellow

14

labels the largest shopping centers that tend to be associated with high levels both for tourists and - as we will show further below - for residents. Overall, however it is clear that tourist activity is fairly concentrated in the historical core of the city.

1(b) Tourism is seasonal: the incidence of tourist spending varies over time within the year, and differently so depending on origin of tourists. In Figure 4, we map as a bivariate chloropleth the quantiles of Spanish and foreign tourists across Barcelona. Spanish tourists are defined as Spanish account holders that visit from outside of the province of Catalonia. In the map locations that experience higher foreign tourist expenditure, but comparatively lower domestic tourist expenditures are market with green colors, while the reverse situation - low foreign but high domestic tourist expenditures - are marked in magenta tones. Locations that experience both high expenditures by domestic and foreign tourists are marked in dark grey, while low expenditure locations are market in light gray. The inner city is popular with both domestic and foreign tourists. While Baraceloneta and Montjuic are particularly popular with foreign visitors, some of the outer areas tend to be more popular with domestic visitors. Overall, there is evidence for a distinctive heterogeneity in preferences for locations between different groups of tourists.

The spatial heterogeneity of different tourist groups interacts with their importance across time, in particular their seasonality. In table6 we demonstrate the expenditure composition in our data. What is probably most striking is the seasonal variation in total tourist expenditures. Between February and July on average, combined expenditures of domestic and foreign tourists increase by a stunning 70 percent. More interestingly, there is substantial heterogeneity across types of tourists, based on ther country of origin. The panel (b) of figure 1 summarizes the spending patterns of the most important groups of tourists. Note how the seasonal pattern changes: some have a market summer preference, some are more smoothed out over the year, the extreme being domestic tourists, who are highly stable throughout the year except for the low season months (January-March). This variation will prove enormously helpful when we seek to identify price and wage effects in using a shift-share instrument in Section 4 below.

Fact 2: Locals' spending and income are spatially determined by residence

2(a) Local consumption is spatial: residents are more likely to spend nearby their home. Residential consumption experiences spatial decay, with the own location often being the destination

15

of a clear majority of expenditure of residents and with other expenditures strongly declining as distance and travel cost increases. In Figure 9 we plot the expenditure shares for a resident of the historical urban core. There are two clear take aways from the figure: The first is that there is a substantial share of expenditures in close vicinity to the residential location. A large fraction of the expenditures take place in less than 1km distance from the home location. The second observation is that expenditure patterns are widely spread across the city, reaching into almost all areas.

2(b) Local income is spatial: residents are more likely to earn nearby their home. As is more commonly known, and as has previously been explored in the urban economics literature, transport cost are prohibitive and prevent long commutes, inducing people to choose locations close to their workspace, resulting in a localized employment pattern. This is similarly apparent in our data. In Figure 2b, we compare commuting patterns from the lunchtime location data for residents of the historical urban core. The majority of the commuting trips lead to neighboring locations in the city. This suggests a very strong distance coefficient in the gravity regression for commuting, as we confirm in Table 8.

Fact 3: Tourist spending affects locals

3(a) Tourist spending spatially crowds out local consumption: In Figure 4, we map the bivariate chloropleth simultaneously showing the change in tourist and local expenditure between February (the tourist low season) and August (the tourist high season) in 2019. It is clear from the map that particularly in the (foreign) tourist hotspots close to the beach as well as the lower part of the historical inner city tourist expenditures increase while local expenditures decrease or only grow weakly. We interpret this to be suggestive evidence that tourism increased prices of local goods, crowding out local consumption. Instead residential expenditure grows much stronger in the predominantly non-tourist residential locations in the northern part of the city. The map also points towards the identification problem that underlies the challenge of estimating the crowding out effect. While some areas seem to indicate crowding out, there are also central areas where expenditures strongly co-move, indicating that - despite possible price increases effects - these areas become more attractive for both tourists and locals in certain seasons.

16

Figure 2: Local Income and Expenditures

(a) Residents Are More Likely to Spend Nearby Their Home

(b) Residents Are More Likely to Earn Nearby Their Home.
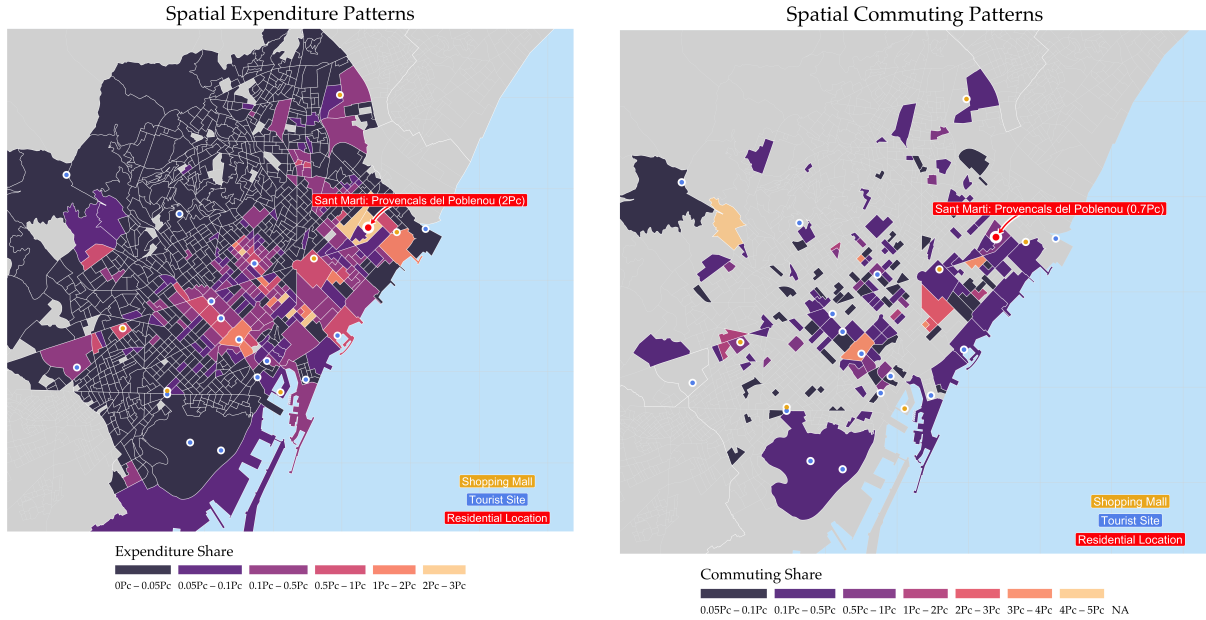
Notes: **(Panel A)** This figure shows the expenditure shares of residents from the census tract denoted by the label in the picture, in the city center. For comparison, we include the equivalent graph for the residents of a different neighborhood further from the center in 9. **(Panel B)** This figure shows the work commuting shares of residents of the same location, as measured by the share of unique cards owned by residents in this location used for lunchtime purchases across the city.

3(b) Tourist spending increases total income in a large part of the city: In Figure 3 we compare income changes of locals between low and high season with the implied market access by tourist spending (tourist spending that reaches a location given commuting patterns of locals). Despite the suggestive evidence of tourist expenditure crowding out local expenditure from Fact #3(a), tourist spending increases from low to high season correlate with income increases of locals. Together with Fact #3(a), this is consistent with tourism both increasing local prices (and crowding out local spending) and increasing local incomes (by increasing total spending).

We now turn to developing a theoretical framework to assess the welfare impacts of tourism. Given the evidence of substantial consumption and income heterogeneity of residents across space from Fact #2, our framework incorporates a flexible urban geography in order to incorporate the complex observed patterns of local consumption and expenditure. Given the suggestive evidence of both price and income effects from Fact #3, our framework allows for tourism to have arbitrary impacts on prices and wages throughout the city.

17

Figure 3: Tourist Spending Increases Income in a Large Part of the City
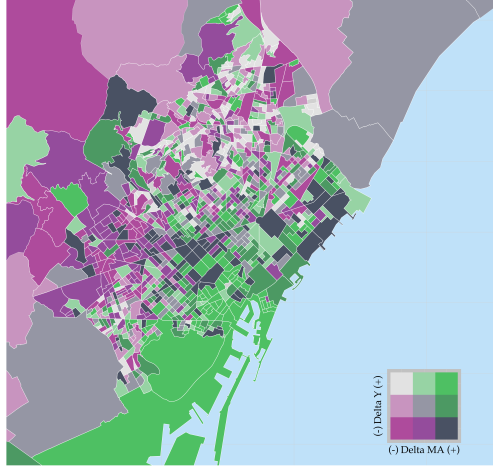
Notes: This figure compares income changes of locals between low and high season with the implied market access (MA) by tourist spending (tourist spending that reaches location given commuting patterns of locals). Greener colors denote locations where both market access and residents' income increase between low and high season. Purple colors denote locations where both market access and residents' income decrease between low and high season. White and dark gray areas are those where market acess and income do not comove.

3.3 A Shift-Share Instrument Exploiting Seasonal Variation in Tourist Origin Composition and Spatial Preferences

Our goal is to estimate the local price and wage effects of tourism. A generic regression of price or wages on tourist activity, however, would be inappropriate, as there may be correlated preference shocks between tourists and locals. We therefore build an instrument to identify β. We make use of the stylized facts outlined in Subsection 3.2 and exploit shifts in tourist origin composition over 16 countries of origin, including domestic tourists from Spain, as well as the spatial distribution of each group's expenditure across all the location in the city. Intuitively, we rely on two facts (see Stylized Fact 1): (1) tourists from different countries of origin allocate their expenditure differently within Barcelona; and (2) the composition of tourists from different counties of origin in Barcelona changes throughout the year.

Specifically, we use the identity

$$X_{it}^T = \sum_{g \in T} \underbrace{\frac{X_{igt}^T}{\sum_j X_{jgt}^T}}_{\pi_{igt}^T} \sum_j X_{jgt}^T,$$

where the shares $\left\{\pi_{igt}^T\right\}_{i=1}^J$ capture the spatial distribution for tourist-group g's spending in

Figure 4: Tourists Crowd Out Locals

(a) Low-to-High Season Change

(b) Instrument and Tourist Expenditures are Highly Correlated
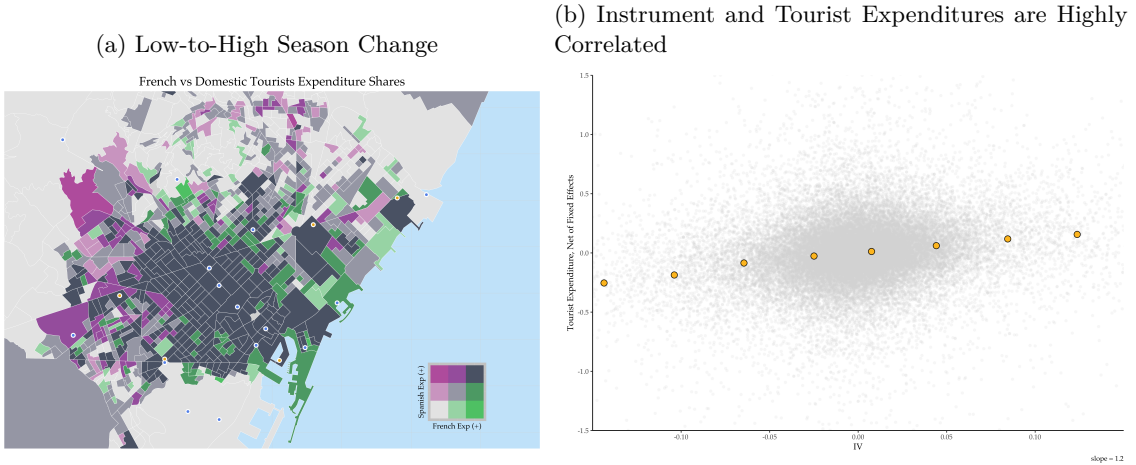
Notes: **(Panel A)** This figure compares the percentage change in tourist and local expenditure between the tourist high season (August) and the low season (January) average across 2017-2019. Locations where tourism expenditures increases and local expenditure decrease or grow weakly are demarked with green colors, while magenta colors mark the locations where the reverse is true. Expenditure changes are measured as level differences in average monthly expenditure per square meter of the underlying tile. Dark gray denotes locations where expenditures comove. The boundary points for the tertiles are given by ([-4.36,0],[0,0.09],[0.09,19,65]) for tourist expenditure changes and by ([-0.9,-0.02],[-0.02,0.01],[0.01,1.13]) for local expenditure changes. **(Panel B)** This scatter plot shows the correlation between the log tourist expenditures and the shift-share instrument $B_{it}^{T,low}$. The black dots forming a cloud are the observations at the location-sector-month level. The yellow dots correspond to a binned scatter. The blue line is the linear fit.

period t across all locations i in the city, and the shift originates from changes in total tourist-group spending in the whole city. Appendix A.4 derives the structural interpretation of our instrument.

We then define two Bartik-style instruments with fixed shares and leave-one-out shifts:

$$
\begin{aligned}
B_{it}^{T} &= \sum_{g \in T} \pi_{ig}^{T,0} \sum_{j \neq i} X_{jgt}^{T} \\
B_{it}^{T,low} &= \sum_{g \in T} \pi_{ig}^{T,low} \sum_{j \neq i} X_{jgt}^{T},
\end{aligned}
$$

depending on the reference period over which the fixed shares are calculated. B_{it}^{T} uses the average share over the twelve months in the first year available—2017 in our data—, while the reference for $B_{it}^{T,low}$ is only over the low season months in the first year—January through March in 2017—. In all cases, the shift in total tourist-group spending across the city leaves out the endogenous location i. Figure 4b shows the fit of our first stage using $B_{it}^{T,low}$.

Table 1: Consumption Gravity

Dependent Variables:	X_ij		lX_ij		lX_ij0	
Model:	(1) Poisson	(2) Poisson	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Variables						
$\ln(Travel\ Time)$	-2.18***	-2.18***	-1.36***	-1.37***	-1.36***	-1.37***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed-effects						
Residence Loc. n (1,095)	✓		✓		✓	
Consumption Loc. i (1,096)	✓		✓		✓	
Residence Loc. $n \times t$ (39,420)		✓		✓		✓
Consumption Loc. $i \times t$ (39,384)		✓		✓		✓
Fit statistics						
Observations	43,204,320	43,125,480	6,566,622	6,566,622	6,566,622	6,566,622
R^2			0.370	0.384	0.370	0.384
Adjusted R^2			0.370	0.377	0.370	0.377

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Aggregate version of the gravity regresions described in 6 using total expenditure between origin n and destination i in year and month t, aggregated across all sectors. Columns (1) and (4) are estimated using PPML. Columns (2) and (4) are estimated using OLS with dependent variable $X + 1$ to include zeros. Columns (3) and (6) are estimated using OLS with dependent variable X, hence excluding zeros.

4 Empirics

Armed with the data described in Section 3.1 and the methodology presented in Section 2, we now turn to an empirical analysis of the welfare impacts of tourism on locals in Barcelona. We pursue two complementary approaches. In the first approach, we seek to estimate average impacts of tourism on prices and wages across the city using our shift-share instrument only. In a second step we then use the short-run predictions for wage and price elasticities from section 2.3 to estimate heterogeneous treatment effects across tiles.

4.1 Average Welfare Effects

We first pursue an empirical regression-based approach to identify the average elasticities of wages and sector-specific prices. We then combine these estimated average elasticities in a very general welfare expression to calculate the welfare impact of tourism. This analysis has the advantage, under the appropriate identification assumptions, of providing estimates that are independent from the theoretical structure. The downside is that we will be able to identify only average price effects and wage effects across all locations (and hence average welfare effects). We begin by discussing first the price regressions and then turn towards the effects on income.

Expenditure Gravity Regressions

We first examine the impact of tourism on local prices. Our dataset does not contain prices directly, but is informative about resident's expenditure shares across tiles. To obtain a clean measure of the attractiveness of a location, we first estimate gravity regressions and isolate the destination fixed effect, which will become our dependent variable for the price regressions. In section A.1 in the appendix we show how a parametric gravity can be derived as a local approximation of the non-parametric model introduced in section 2. Our main specifications are being estimated using pseudo poisson maximum likelihood to address concerns about bias from heteroskedasticity and zeros in the data Silva and Tenreyro (2006). Specifically we estimate,

$$X_{ni,s,t} = \exp\left(\gamma_{n,s,t} + \delta_{is,t} + \beta \log dist_{ni}\right)\epsilon_{ni,s,t} \tag{6}$$

where $X_{ni,s,t}$ refers to the expenditures of residents residing in location n in destination location i in sector s at time t. We regress this bilateral flow on the travel time obtained from the HERE API, averaging times between public transit and driving times. There is stark heterogeneity across sectors, emphasizing the importance of sectoral data to understand the spatial component of urban consumption. In our original specification on the raw data maintained by CaixaBank we include the full set of origin-sector-time as well destination-sector-time fixed effects. The resulting coefficient for this estimation on distance is visualized in Figure 10. In 1 we present the results replicating the gravity estimation on aggregated bilateral flows that abstract from sectoral heterogeneit. The table also presents the log gravity formulation as a reference point for the improved poisson estimation. In what follows, we use the full set of destination fixed effects obtained from the ppml specification which was run on CaixaBank's raw data.

Price Regressions

We now turn towards analyzing the impact of tourist expenditures on residential expenditure shares. To do so we regress the destination fixed effect in sector s in location i on changes in tourist expenditure in the same location, instrumenting tourist expenditure with the Bartik instrument introduced in Subsection 3.3:

$$\ln \delta_{ismt} = \gamma_{is} + \gamma_{ist} + \gamma_{ism} + \gamma_{mts} + \beta^p \times \ln E_{itm}^T + \epsilon_{ismt} \tag{7}$$

Table 2: A Tourist Demand Shock Increases Prices on Average

Dependent Variable:	Residents Expenditure (Gravity): δ_{ist}^{R}								
	OLS			IV - Ref: 2017 Average			IV - Ref: 2017 Low Season		
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
Tourists Expenditure: $\ln E_{it}^{T}$	0.159***	0.152***	0.091***	-0.437***	-0.477***	-0.576***	-0.469***	-0.512***	-0.668***
	(0.011)	(0.014)	(0.010)	(0.096)	(0.108)	(0.189)	(0.100)	(0.111)	(0.223)
Fixed-effects									
Month-Year×Sector (480)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Location×Sector (21,920)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Location×Sector×Year (43,840)		✓	✓		✓	✓		✓	✓
Location×Sector×Month (263,040)			✓			✓			✓
Fit statistics									
Observations	526,080	526,080	526,080	526,080	526,080	526,080	526,080	526,080	526,080
Adjusted R^2	0.992	0.993	0.994	0.99	0.991	0.993	0.99	0.991	0.992
F-test $= t^2$ (1st Stage)				145.4	138.2	38.4	153.3	148.7	30.7

Heteroskedasticity-robust standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Notes: Estimates from regressing our measure of (inverse) prices (gravity destination fixed obtained as in 6) in sector s, location i, and year-month t on total location i's tourist expenditure in the same year-month. To obtain price elasticities from these estimates, we will rescale using a function of the elasticity of substitution, which results in a proportional negative factor. Hence, a negative coefficient indicates a positive price elasticity. Columns (4) through (6) use the Bartik instrument B^T that is build with shares fixed to the average in 2017, while columnts (7) through (9) use $B^{T,low}$, using an average over the low-season only as a reference.

where δ_{ismt} refers to the destination fixed effect for expenditures in sector s in year t and month m at location i obtained from the previous gravity regression and where $\ln E_{itm}^{T}$ refers to the overall level of tourist expenditures across all groups appropriately instrument in a first stage with the shift-share instrument of section 3.3. The coefficient of interest is β^p which measures the responsiveness of residential expenditure shares to plausibly exogenous increases in tourist expenditures. From our discussion in A.1, we know that the residual ϵ_{ismt} includes any time varying changes in local preferences for a good (i, s). The exclusion restriction requires that aggregate variation in the composition of tourists is uncorrelated with location-specific changes in local preferences within Barcelona, conditional on fixed effects. In our specifications, a sector-year-month fixed effect is always necessary since the gravity regression is only identified up-to-scale and separately run for each time period and sector. The equation indicates the most stringent specification and additionally adds three groups of fixed effects: Location-sector fixed effects control for cross-sectional heterogeneity and effectively transform the specification into panel regressions. Location-sector-month control for seasonal variation in the appeal of different locations (e.g. beach-side locations during

the summer). Finally, location-sector-year fixed effects control for aggregate demand shocks.

The regressions results are reported in table 2. The table presents the OLS results and then six different specifications for the 2SLS estimation. Column (1) through (3) show the OLS results, adding incrementally the fixed effects discussed above. Across all specifications, the OLS shows a significant positive relationship between tourist expenditures and local expenditure shares. This illustrates the bias inherent in the raw data: Any residual change in the amenity or attractiveness jointly shifts tourist and residential expenditure into the same direction, but cannot be directly observed or controlled for, thus making it difficult to obtain the causal effect of tourist expenditures on residential expenditure shares. As discussed before, we construct two different types of instruments, both of which exploit differences in the group composition of tourism throughout the year and between years, but differing in the baseline expenditure shares for tourist groups that are being used. Column (4) through (6) present the results where the shares are constructed using the average spatial expenditure share in 2017. Column (7) through (9) use the average expenditure share in the low season of 2017 - effectively January through March. Again we present the results adding incrementally the fixed effects discussed above. Using 2SLS flips the sign of the estimates and we consistently get estimates in the range of -.4 and -.7 across specifications, indicating a substantial negative effect of tourism on residential expenditure shares. To obtain the price efffect from the point coefficient we need to take a stance on the demand elasticity. Calibrating to a typical value from the literature ($\sigma = 5$), we obtain a price effect in the range of $[.1, .175]$.

In the appendix, we furthermore present separate results for regressions on housing prices. Table 12 presents the results. Our data provides two different series of data, one representing changes across tiles for rental rates, the other for housing prices. The data provides prices directly and not expenditure shares, which implies that a theoretically consistent estimate would be positive - one where additional tourist demand increases prices. Again, we present two different sets of specifications for each data series exploiting the two different instruments that we constructed. All specifications control for location fixed effects absorbing cross-sectional variation. Furthermore, we separately introduce either location-year or location-month fixed effects to absorb either aggregate or seasonal common variation. Throughout we find a significant positive elasticity ranging between 0.01 and 0.06, somewhat smaller than the implied price effect from electronic transactions.

Table 3: A Tourist Demand Shock (Weakly) Increases Income on Average

Dependent Variable:				ln Income (Mean)		
	Lunch	Cell		Cell Phone		Lunchtime
	OLS		IV - 2017 Average	IV - 2017 Low Season	IV - 2017 Average	IV - 2017 Low Season
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
ln MA_{nt}	0.006	0.012	0.032	0.035	0.038***	0.040**
	(0.004)	(0.012)	(0.021)	(0.025)	(0.015)	(0.018)
Fixed-effects						
Location	✓	✓	✓	✓	✓	✓
Month	✓	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	26,472	1,776	1,776	1,776	26,472	26,472
Adjusted R^2	0.888	0.93	0.93	0.93	0.888	0.888
F-test $= t^2$ (1st Stage)			204.5	142.8	1,267.2	927.0

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Estimates from regressing income of residents on our measure of Market Access discussed in 9. Columns (3) through (4) use the Bartik instrument B^T that is build with shares fixed to the average in 2017, while columnts (5) through (6) use $B^{T,low}$, using an average over the low-season only as a reference.

Income Regressions

We now examine the impact of tourism on local incomes. To do so we regress in residential income on changes in their exposure to tourist shocks. From section 2 we know that changes in spatial income can be expressed as,

$$d \ln v_n = \sum_{i=1}^{N} c_{ni} \partial \ln w_i \tag{8}$$

where the overall effect on tourist expenditures on residential income depends on the income shares across locations as well as the effect of tourism on cell-specific wages. We create a proxy for the right hand side of this equation by spatially summing over the tourist shock across locations,

$$\ln \mathrm{MA}_{ntm}^T = \sum_i c_{ni} \times \ln E_{itm}^T \tag{9}$$

where $\ln \mathrm{MA}_{ntm}^T$ refers to the log of (labor) market access to tourist activity in residential location n in year t and month m. Notice that 9 is an imperfect proxy for equation 8, since it does not correct for how tourist expenditures affect wages in the destination location. This will affect the

24

interpretation of the coefficient in our regression. With this proxy in hand we then proceed to examine the impact of tourist expenditures on residential income. Specifically, we run the following regression,

$$\ln v_{nmt} = \gamma_{it} + \gamma_{im} + \gamma_{tm} + \beta^w \times \ln \mathrm{MA}_{ntm}^T + \epsilon_{imt} \tag{10}$$

where v_{nmt} refers to the (mean) income of residents in residential tile n in year t and in month m and where furthermore $\ln \mathrm{MA}_{ntm}^T$ refers to the (labor) market access measure defined above, appropriately instrument in a first stage with its equivalent constructed from the shift-share instrument of section 3.3. The coefficient of interest is β^w which measures the responsiveness of residential income to plausibly exogenous increases in tourist expenditures across accessible labor markets. The coefficient therefore measures the pass-through of tourist income into wages and therefore income and is the welfare relevant measure that is needed to implement the welfare analysis indicated by equation 3.

In section A.2 in the appendix we provide a more detailed discussion, but it is intuitive that ϵ_{imt} includes time-varying changes in the productivity or amenity of different locations that a resident draws income from. The exclusion restriction requires that aggregate variation in the composition of tourists is uncorrelated with location-specific changes productivity or amenities within Barcelona, conditional on the fixed effects. The equation indicates the most stringent specification and additionally adds three groups of fixed effects: Location fixed effects control for cross-sectional heterogeneity and effectively transform the specification into panel regressions. Month fixed effects control for seasonal variation in income that is common to all locations. Finally, year fixed effects control for aggregate demand shocks.

The regressions results are reported in table 3. The table presents the OLS results and then four different specifications for the 2SLS estimation. Column (1) and (2) show the OLS results, fully saturating the regressions with the fixed effects discussed above. The OLS indicate a small insignificant relationship between tourist expenditures and residential income. This might be due to measurement error, it might also be caused by a negative relationship between the tourist shock and the productivity of individual cells. As discussed before, we construct two different types of instruments, both of which exploit differences in the group composition of tourism throughout the year and between years, but differing in the baseline expenditure shares for tourist groups that are

being used. To construct the market access term we rely on our two data sources for commuting patterns and present separate results for cell phone data and our imputed lunchtime expenditure commuting dataset. Column (3) through (4) present the results using cell phone data, while Column (5) through (6) use our lunchtime expenditure data. Using 2SLS provides consistently positive results. The cell phone data reduces our number of our observations due to its higher aggregation and therefore our precision. Only the results from the lunchtime data are significant, but the magnitudes are comparable across both approaches. We get estimates of a pass-through from tourist expenditures to residential incomes of approximately 0.04. Compared to the price effects this is substantially lower.

Is tourism good for the locals (on average)?

Combining the price and wage estimates, we now assess whether or not tourism is good for locals on average using welfare equation (3). Since the price and wage effects are averaged across all locations, the spatial patterns of commuting do not affect the welfare estimates and the spatial patterns of consumption affect the welfare estimates only inasmuch as locals differ in their sectoral expenditure shares. Evaluating the price effects at the aggregate (city-wide) sectoral expenditure shares, we find that the consumer price index elasticity to tourism expenditure is 0.1 to 0.175. Combining this estimate with our preferred housing rental price elasticity of 0.06 and weighting by the Spanish expenditure share on housing which according to the household budget survey (HBS) is approximately .31, we obtain an overall impact in the range of 0.08 and 0.14. Since the wage elasticity to tourism expenditure is approximately 0.04, this means that the average welfare impact of tourism was -0.04 to -0.1. To put these number in context, consider the increase in tourist expenditures between February and July, on average. Using our data, this number is 70.3%. Using the average for our elasticities above this seasonal increase would translate into an income gain of 2.8% and a price-index welfare deterioration of 9.9%, with a net welfare deterioration of 7.7% for residents on average. This average effect turns out to mask substantial heterogeneity in welfare effects across residents of Barcelona, which we now turn to examining.

Table 4: A Tourist Demand Shock Increases Income More in Those Locations *Predictably More Exposed*

Dependent Variable:				ln Income (Mean)		
	Cell	Lunch		Cell Phone		Lunchtime
			IV - 2017 Average	IV - 2017 Low Season	IV - 2017 Average	IV - 2017 Low Season
	OLS					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$\ln \mathrm{CiE}_{nt}$	0.012	0.006	0.007	0.008	-0.005	-0.009
	(0.012)	(0.004)	(0.029)	(0.037)	(0.021)	(0.025)
$\ln \widehat{\mathrm{CiE}}_{nt}(\eta_{it}^{0})$			0.045	0.046	0.086***	0.092***
			(0.030)	(0.033)	(0.027)	(0.027)
Fixed-effects						
Location	✓	✓	✓	✓	✓	✓
Month	✓	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	1,776	26,472	1,776	1,776	26,472	26,472
Adjusted R^2	0.93	0.888	0.93	0.93	0.888	0.888
F-test $= t^2$ (1st Stage)			204.5	142.8	1,267.2	927.0

Heteroskedasticity-robust standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Notes: Estimates from running specification 12.

4.2 Heterogeneous Welfare Effects

The advantage of the reduced-form approach of Section 4.1 was that it did not require any theoretical assumptions to identify the average price and wage effects of tourism across the city. In this section we refine this approach using the theoretically predicted heterogeneous effects from section 2.3 to construct regressions that identify heterogeneous treatment effects of tourist expenditure shocks.

Heterogeneous Income Effects

Our approach is simple and straightforward: we adapt the basic specification of Section 4.1 to allow for heterogeneous treatment effects, where the source of the heterogeneity is given by the distribution of the model-implied elasticities calculated in Section 2.3. Consider first the income regressions. Let $\eta_i^w \equiv \frac{\partial \log w_i}{\partial \log X_i^T}$ be the model-predicted elasticity of the wage in location i to an increase in tourism expenditure. We can use the "zero-degree" elasticity for wages to refine our (labor) market access proxy to account for the heterogeneous incidence of tourist shocks across tiles, i.e.

$$\ln \mathrm{MA}_{ntm}^{T}\left(\eta_{itm}^{0}\right) = \sum_i c_{ni} \times \eta_{itm}^{0} \times \log E_{itm}^{T} \tag{11}$$

where $\ln \mathrm{MA}_{ntm}^{T}\left(\eta_{itm}^{0}\right)$ refers to the log of (labor) market access to tourist activity in residential location n in year t and month m, but crucially the incidence of the shock is adjusted by $\eta_{itm}^{0} = \frac{E_i^T}{X_i}$, that is the tourist intensity of the location sector. To calculate this measure we take averages in 2017, symmetrically to how we obtain expenditure shares for the shift-share instrument. We can use this improved measure to determine heterogenous treatment effects that depend on how much income a given location derives from tourist-intensive locations. Specifically we run,

$$\ln v_{nmt} = \gamma_{it} + \gamma_{im} + \gamma_{tm} + \beta^w \times \ln \mathrm{MA}_{ntm}^{T} + \beta^{w,het} \times \ln \mathrm{MA}_{ntm}^{T}\left(\eta_{itm}^{0}\right) + \epsilon_{imt}, \qquad (12)$$

where, as before, v_{nmt} refers to the (mean) income of residents in residential tile n in year t and in month m and where furthermore $\ln \mathrm{MA}_{ntm}^{T}\left(\eta_{itm}^{0}\right)$ refers to the data counter-part to the instrument discussed above and is appropriately instrumented in a first stage with its equivalent constructed from the shift-share instrument of section 3.3. We also include the level term of the market access proxy. The coefficient of interest is $\beta^{w,het}$ which measures the responsiveness of residential income to plausibly exogenous increases in tourist expenditures and their heterogeenous incidence across accessible labor markets. The combination of the level coefficient and the heterogeneous coefficient is the welfare relevant effect of a tourist shock on income and operationalizes the welfare calculations from equation 3. Identification relies on the same assumptions as in the previous section and fixed effects are introduced in a symmetric manner.

The regressions results are reported in table 4. The table presents the OLS results and then four different specifications for the 2SLS estimation. Column (1) and (2) show the OLS results, fully saturating the regressions with the fixed effects discussed above. The OLS indicate a small positive relationship between tourist expenditures and residential income. As before, for the 2SLS results that follow in Column (3) through (6), we report the results for lunch and cell phone data as well as the results for the two different instruments. And as before, using 2SLS provides consistently positive results. However, the regression loads heavily on the heterogeneous market access proxy rather than the mean term. Our preferred estimate is the specification in column (6) which presents a $\beta^{w,het} = 0.092$ and a $\beta^w \approx 0$. Indicating substantially heterogeneous income effects across Barcelona.

Table 5: A Tourist Demand Shock Increases Prices More in Those Locations *Predictably More Exposed*

Dependent Variable:	δ_{ist}^R					
	IV - Ref: 2017 Average			IV - Ref: 2017 Low Season		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$\widehat{\ln E_{it}^T}$	0.019	-0.029	-0.059	0.011	-0.037	-0.062
	(0.059)	(0.059)	(0.060)	(0.064)	(0.064)	(0.065)
$\widehat{\ln E_{it}^T} \times \eta_{it}^0$	-0.523***	-0.467***	-0.357***	-0.628***	-0.555***	-0.448***
	(0.087)	(0.087)	(0.096)	(0.091)	(0.091)	(0.102)
$\widehat{\ln E_{it}^{GE}}(\bar{\eta}_i^0)$		-0.004***	-0.009***		-0.005***	-0.009***
		(0.0005)	(0.002)		(0.0005)	(0.002)
$\widehat{\ln E_{it}^{GE}}(\bar{\eta}_i^0) \times \bar{\eta}_i^{0,Res}$			0.007***			0.006**
			(0.003)			(0.003)
Fixed-effects						
Month-Year×Sector (480)	✓	✓	✓	✓	✓	✓
Location×Sector (21,840)	✓	✓	✓	✓	✓	✓
Location×Sector×Year (43,680)	✓	✓	✓	✓	✓	✓
Location×Sector×Month (262,080)	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	524,160	524,160	524,160	524,160	524,160	524,160
Adjusted R^2	0.975	0.975	0.975	0.975	0.975	0.975

Normal standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Estimates from running specification 13. $\widehat{\ln E_{it}^T}$ denotes the tourist shock across the city. $\widehat{\ln E_{it}^T} \times \eta_{it}^0$ captures heterogeneity between locations, as measured by tourist shares over total expenditure. $\widehat{\ln E_{it}^{GE}}$ captures the general equilibrium effects inspired by our model and calculated in our data.

Heterogeneous Price Effects Similarly, let $\eta_{is}^p \equiv \frac{\partial \log p_{is}}{\partial \log X_i^T}$ be the model-predicted elasticity of the price in location i and sector s to an increase in tourism expenditure. We can use the predicted elasticity to create interaction terms with the tourist expenditure shock that adjusts for the heterogeneous incidence across cells. Specifically we run,

$$\ln \delta_{ismt} = \gamma_{is} + \gamma_{ist} + \gamma_{ism} + \gamma_{mts} + \beta^p \times \ln E_{itm}^T + \beta^{p,Het} \times \eta_{itm}^0 \times \ln E_{itm}^T + \beta^{p,GE} \times \ln \mathrm{E}_{ntm}^{T,GE} \left(\eta_{itm}^0 \right) + \epsilon_{ismt},$$
(13)

where, as before, δ_{ismt} refers to the destination fixed effect for expenditures in sector s in year t and month m at location i obtained from the previous gravity regression, and where $\ln E_{itm}^T$ refers to the overall level of tourist expenditures across all groups appropriately instrument in a first stage with the shift-share instrument of section 3.3. An additional assumption that is often being ignored when estimating causal effects in urban networks, is the assumed absence of spillovers or what is sometimes called the Stable Unit Treatment Value Assumption (SUTVA). Indeed as our theory

shows spillovers - in the sensen of general equilibrium adjustments - are a likely occurrence and theoretically predicted. To adjust for this we introduce an additional term that captures GE spillover effects via labor market linkages. Specifically, we instrument for indirect tourist expenditures with the indirect expenditure shocks, i.e. $\ln \mathrm{E}_{ntm}^{T,GE}\left(\eta_{itm}^0\right) = \sum_n s_{ni} \times \ln \widehat{\mathrm{v}}_{nmt}$, where income data is being instrument with the (labor) market access proxy discussed in the previous paragraph. This addition creates regressions that are robust to GE interactions, and furthermore test the model prediction that such interactions exist and also affect the heterogeneity of price adjustments across locations.

The coefficients of interest are β^p, $\beta^{p,Het}$, $\beta^{p,GE}$ which respectively measure the responsiveness of residential expenditure shares to plausibly exogenous increases in tourist expenditures, on average, heterogeneously across tiles and indirectly via labor market linkages. The exclusion restriction still requires that aggregate variation in the composition of tourists is uncorrelated with location-specific changes in local preferences within Barcelona, conditional on fixed effects. We add fixed effects to our specification as before.

The regressions results are reported in table 5. The table presents the first stage results and then different specifications for the 2SLS estimation, reproducing the analysis for different inclusions of tourist expenditure terms and for the two different instruments that we constructed. In column (2), (3), (8) and (9) we can see that the heterogeneous "zero-degree" interaction is important and explains almost all of the variation that the instrument induces into residential expenditure patterns. However, column (4) through (6) and columns (10) through (12) indicate that the inclusion of different GE predicted spillover terms lowers the point coefficient the direct tourist effect in absolute magnitude, indicating that a violation of the Stable Unit Treatment Value Assumption biased the coefficient downward, likely because the specification without the GE term wrongly attributes the indirect effect of a spatially correlated shock to the direct effect. Nevertheless we still obtain robustly negative coefficients for direct effect and in our preferred specification the indirect effect carries an additional negative effect on price, consistent with our theoretical predictions.
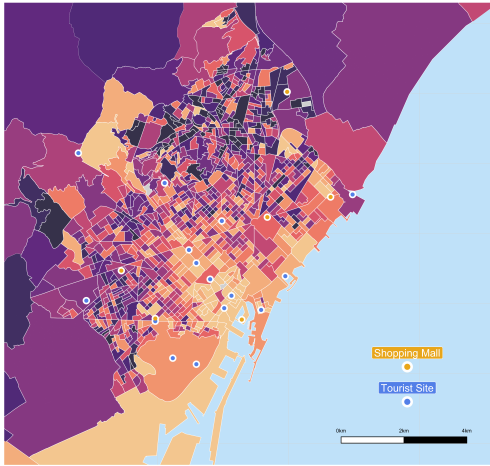
Implied welfare effects

Armed with estimates of $\beta^{w,het}, \beta^w, \beta^{p,het}, \beta^p, \beta^{p,GE}$ we can examine the predicted wage and income effect taking the heterogeneous incidence of the shock into account. Specifically, we predict the income and wage effect using the difference in the shift-share instrument between January and

August, simulating a move from low to high season. The maps in 5 present separately the predicted effects on income, price indices and overall welfare.
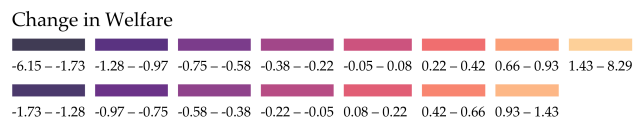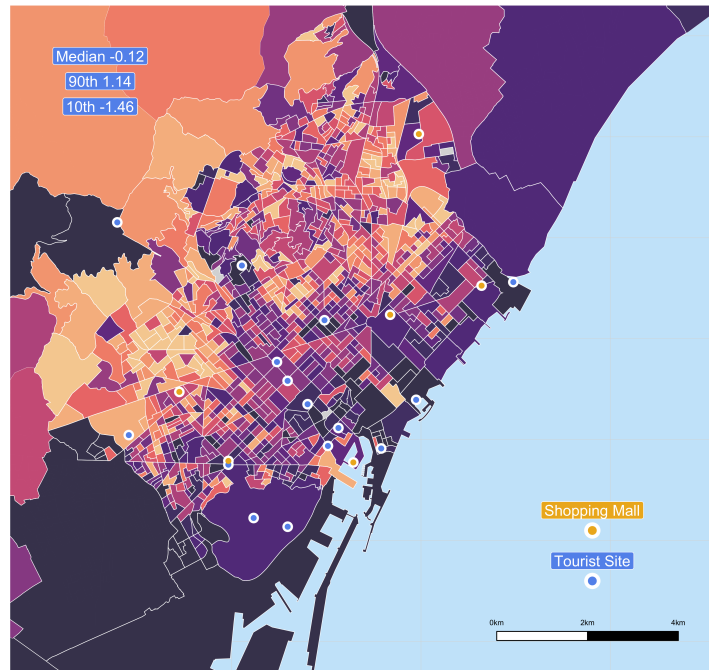
We find that tourism has a modest negative effect; however, these average small losses mask substantial heterogeneity all across the city: The welfare changes from moving between low and high tourist seasons range from a *negative* 1.46 percent (10th percentile in the welfare changes' distribution over locations) to a *positive* 1.14 percent (percentile 90th). Dissecting this net welfare changes into price and income effects, residents in the city center and those near tourist locations bear the largest price changes but also enjoy substantial income gains. In contrast, residents of peripheric neighborhoods suffer lower but still sizable price changes, with the income gains varying between different outer city locations: some experience none and some get moderate income benefits from tourism.

Figure 5: Is tourism good for locals?

(a) Changes in Income

Change in Income

0.29 − 0.8 0.92 − 0.98 1.04 − 1.1 1.17 − 1.23 1.29 − 1.36 1.44 − 1.55 1.65 − 1.78 2.01 − 3.69

0.8 − 0.92 0.98 − 1.04 1.1 − 1.17 1.23 − 1.29 1.36 − 1.44 1.55 − 1.65 1.78 − 2.01

(b) Changes in Price Index

Change in Price Index

0.64 − 0.92 1.04 − 1.23 1.35 − 1.42 1.51 − 1.62 1.73 − 1.84 1.98 − 2.15 2.33 − 2.56 3.08 − 4.7

0.92 − 1.04 1.23 − 1.35 1.42 − 1.51 1.62 − 1.73 1.84 − 1.98 2.15 − 2.33 2.56 − 3.08

Change in Welfare

-6.15 − -1.73 -1.28 − -0.97 -0.75 − -0.58 -0.38 − -0.22 -0.05 − 0.08 0.22 − 0.42 0.66 − 0.93 1.43 − 8.29

-1.73 − -1.28 -0.97 − -0.75 -0.58 − -0.38 -0.22 − -0.05 0.08 − 0.22 0.42 − 0.66 0.93 − 1.43
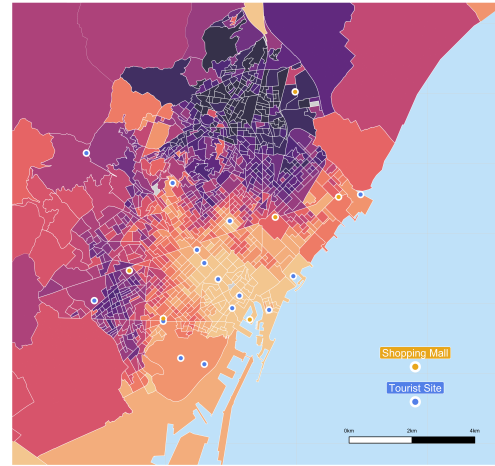
(c) Changes in Welfare

Notes: This figure shows the "short-run" impact of the observed increase in tourism between high and low season on wages (panel a), the price index (panel b), and total welfare (panel c), where by "short-run" we mean that local labor allocations and expenditure shares are kept fixed.

5 Quantitative Evaluation

The advantage of the framework presented in the previous subsection is its generality: the expressions derive hold for any homothetic preferences, any constant returns to scale production functions, and any commuting function. The disadvantage of the framework is that the welfare expressions are valid only for small tourist shocks and they hold only when expenditure shares and labor allocations are held constant (the "short run"). In this section, we present a complementary framework where we pursue the opposite tactic: we assume particular set of preferences and production functions and then derive welfare expressions that hold for arbitrarily sized tourist shocks and account for changes in labor allocations.

5.1 Quantitative Model

5.1.1 Setting

Consider a city consisting of many city blocks which are index by $i, n \in \{1, \ldots, N\}$. In the following n will refer to the location of residence, while i will refer to the location of production and consumption. Residents in each city block are endowed with a total time endowment of T_n. Each city block is also endowed with M_i units of a specific factor. The city blocks are fixed in their geography to each other. Each city block produces across multiple sectors indexed by $s \in \{0, \ldots, S\}$.

5.1.2 Residential Preferences

The representative resident makes two choices, first how to allocate labor across sectors in order to maximize her income and second, given income and prices, she chooses her consumption to maximize her utility. The representative resident's utility is represented by a nested CES utility function, with a constant elasticity of substitution within sectors across location-differentiated varieties and a different constant elasticity of substitution between sectors. The indirect utility function is given by,

$$u_n = \frac{v_n}{\left(\sum_{s=0}^{S} \alpha_s \left(\left(\sum_{i=1}^{N} \tilde{p}_{nis}^{1-\sigma_s} \right)^{\frac{1}{1-\sigma_s}} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}}}$$

where the denominator corresponds to the price aggregator $G(\cdot)$, v_n is the income of the resident and $\tilde{p}_{nis}$ refers to the prices in block n of the goods produced by each block i in sector s. Consumption

is assumed to require physically travelling to a city block and incurring an iceberg variable trade costs such that prices are given by

$$\tilde{p}_{nis} = b_{ni}p_{nis} = b_{ni}\tau_{ni}p_{is}$$

where $\tau_{ni} \geq 1$ refers to the iceberg variable trade costs and where we assume that $\tau_{nn} = 1$. We furthermore, parameterize the amenity spillovers,

$$b_{ni} = \tilde{b}_{ni}(x_{is})^{\gamma_{rr}} \Pi_{g=1}^{G}(x_{is}^{g})^{\gamma_{rg}}$$

From Roy's identity, demand in block n for the good produced in country i is given by,

$$q_{nis}(\tilde{\boldsymbol{p}}_n) = \left(\frac{\tilde{p}_{nis}^{-\sigma_s}}{\left(\sum_{i=1}^{N}\tilde{p}_{nis}^{1-\sigma_s}\right)}\right) \left(\frac{\alpha_s\left(\left(\sum_{i=1}^{N}\tilde{p}_{nis}^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}\right)^{1-\eta}}{\sum_{s=0}^{S}\alpha_s\left(\left(\sum_{i=1}^{N}\tilde{p}_{nis}^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}\right)^{1-\eta}}\right) v_n$$

The labor supply decision by the representative resident in block n is defined by an income maximization problem. Each location has a total time endowment of T_n and the representative resident solves a problem of how to best allocate labor across the city and sectors subject to a time constraint. The maximization problem of the representative resident in block n is given by,

$$v_n = \max_{\{\ell_{ni}\}} \sum_{i=1}^{N} w_{ni}\ell_{ni}$$

$$\text{s.t. } H_n(\{\ell_{ni}\}) = T_n$$

where we assume that the $H_n(\{\ell_{ni}\})$ takes on a constant elasticity of substitution functional form, i.e.

$$H_n(\{\ell_{ni}\}) = \left(\sum_i (\ell_{ni})^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}$$

where ϵ is the elasticity of substitution. We require H_n to be convex and thus we require $\epsilon < 0$. This implies that we can write the indirect time use function as,

$$h_n = \frac{v_n}{\left(\sum_i (w_{ni})^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}}$$

where the denominator corresponds to the wage aggregator $J(\boldsymbol{w}_n)$, v_n is the income of the resident and w_{ni} refers to the wages in block i that can be accessed to a resident in block n. Labor supply is assumed to require physically travelling to a city block and incurring an iceberg variable commuting cost costs such that prices are given by $w_{ni} = \mu_{ni} w_i$, where $\mu_{ni} \geq 1$ refers to the iceberg variable trade costs and where we assume that $\mu_{nn} = 1$. From the commuting equivalent of Roy's identity 2, labor supply from residents in block n to production in block i is given by,

$$\ell_{ni}(\boldsymbol{w}_n) = \frac{w_{ni}^{-\epsilon}}{\left(\sum_i (w_{ni})^{1-\epsilon}\right)} v_n$$

where labor supply is increasing in the wage in the destination location, but decreasing in the wage index, which summarizes the alternative outside options to the resident in location n.

5.1.3   Tourist Preferences

The city is visited by tourists from many countries, indexed by $g, h \in \{1, \dots, G\}$. The representative tourist from country g has homothetic demand that can be represented by the following indirect utility function:

$$u_g = \frac{E_g^T}{G(\tilde{\boldsymbol{p}})}$$

where $G(\cdot)$ represents some price aggregator and E_g^T is the fixed expenditure of tourist from country g and $\boldsymbol{p}$ refers to the vector of prices across all city blocks. Consumption is assumed to require physically travelling to a city block and incurring an iceberg variable trade costs such that prices are given by $\tilde{p}_{gis} = b_{gi} p_{is}$ where $\tau_{ni} \geq 1$ refers to the iceberg variable trade costs and where we assume that $\tau_{nn} = 1$. We furthermore, parameterize the amenity spillovers,

$$b_{gi} = \tilde{b}_{gi} (x_{is})^{\gamma_{gr}} \Pi_{g=1}^{G} (x_{is}^g)^{\gamma_{gh}}$$

From Roy's identity, demand for group g for the good produced in block i is given by,

$$q_{gis}(\tilde{\boldsymbol{p}}) = \left(\frac{\tilde{p}_{gis}^{-\sigma_s}}{\left(\sum_{i=1}^{N} \tilde{p}_{gis}^{1-\sigma_s}\right)}\right) \left(\frac{\alpha_s \left(\left(\sum_{i=1}^{N} \tilde{p}_{gis}^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}\right)^{1-\eta}}{\sum_{s=0}^{S} \alpha_s \left(\left(\sum_{i=1}^{N} \tilde{p}_{gis}^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}\right)^{1-\eta}}\right) E_g^T$$

5.1.4 Production

Production in either sector requires the labor of the resident ℓ_{is} and a specific factor m_{is} with some constant returns to scale production function

$$Q_{is} = F_{is}\left(\ell_{is}, m_{is}\right) = z_{is}\ell_{is}^{\beta_s} m_{is}^{1-\beta_s}.$$

Competitive markets imply that factor returns are equal to marginal products, i.e.

$$w_{is} = p_{is}\frac{\partial F_{is}\left(\ell_{is}, m_{is}\right)}{\partial \ell_{is}} \qquad r_{is} = p_{is}\frac{\partial F_{is}\left(\ell_{is}, m_{is}\right)}{\partial m_{is}}$$

5.1.5 Expenditure and Commuting Shares

For a resident in block n the expenditure shares on goods produced in block i can be written as,

$$s_{nis} = \frac{p_{nis}q_{nis}\left(\boldsymbol{p}_n\right)}{\sum_{\ell=1}^{N} p_{n\ell s}q_{n\ell s}\left(\boldsymbol{p}_n\right)} \equiv \frac{x_{nis}\left(\boldsymbol{p}_n\right)}{\sum_{\ell=1}^{N} x_{n\ell s}\left(\boldsymbol{p}_n\right)}$$

Similarly, for a tourist from country g the expenditure shares on goods produced in block i can be written as,

$$s_{gis} = \frac{p_{is}q_{gis}\left(\boldsymbol{p}\right)}{\sum_{\ell=1}^{N} p_{\ell s}q_{g\ell s}\left(\boldsymbol{p}\right)} \equiv \frac{x_{gis}\left(\boldsymbol{p}\right)}{\sum_{\ell=1}^{N} x_{g\ell s}\left(\boldsymbol{p}\right)}$$

Finally, we can define income shares that residents in block n derive from supplying workers to location i, which can be written as,

$$c_{ni} = \frac{w_{ni}\ell_{ni}\left(\boldsymbol{w}_n\right)}{\sum_{\ell=1}^{N} w_{n\ell}\ell_{n\ell}\left(\boldsymbol{w}_n\right)} \equiv \frac{y_{ni}\left(\boldsymbol{w}_n\right)}{\sum_{\ell=1}^{N} y_{n\ell}\left(\boldsymbol{w}_n\right)}$$

5.1.6 Market Clearing

Market clearing requires that income in country i equals the expenditure on goods produced by that country:

$$y_{is} = w_{is}\ell_{is} + r_{is}m_{is} = \sum_{n=1}^{N} s_{nis}v_n + \sum_{g=1}^{G} s_{gis}E_g^T$$

We can separately define the market clearing condition for each factor. For the specific factor we obtain,

$$\frac{r_{is} m_{is}}{\theta_s^m} = \sum_{n=1}^{N} s_{nis} v_n + \sum_{g=1}^{G} s_{gis} E_g^T$$

For the labor market clearing condition we assume that labor is freely mobile between sectors within locations such that,

$$w_i \ell_i = \sum_{s=0}^{S} \theta_s^{\ell} \sum_{n=1}^{N} s_{nis} v_n + \sum_{s=0}^{S} \theta_s^{\ell} \sum_{g=1}^{G} s_{gis} E_g^T$$

Disposable income is given by the aggregate income derived from all labor markets,

$$v_n = \left(\sum_i (w_{ni})^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \times T_n$$

5.1.7 Counterfactual Equilibrium in Hat Algebra

TBD

5.2 Parameterization

TBD

5.3 Results

5.3.1 Welfare Effect of Tourism

TBD

5.3.2 Optimal Tourist Tax

In this section we propose to analyze the impact of tax on consumption by tourists, or a tourist tax. Specifically, we consider a proportional tax to consumer prices for tourists only, i.e.

$$p_{is}^t = (p_{is} (1 + t_{is}))$$

We consider a welfare problem where a social planner is maximizing weighted indirect utility

Figure 6: Optimal Tourist Tax

(a) Mapping the Optimal Tourist Tax

(b) Mapping Welfare under Optimal Tourist Tax
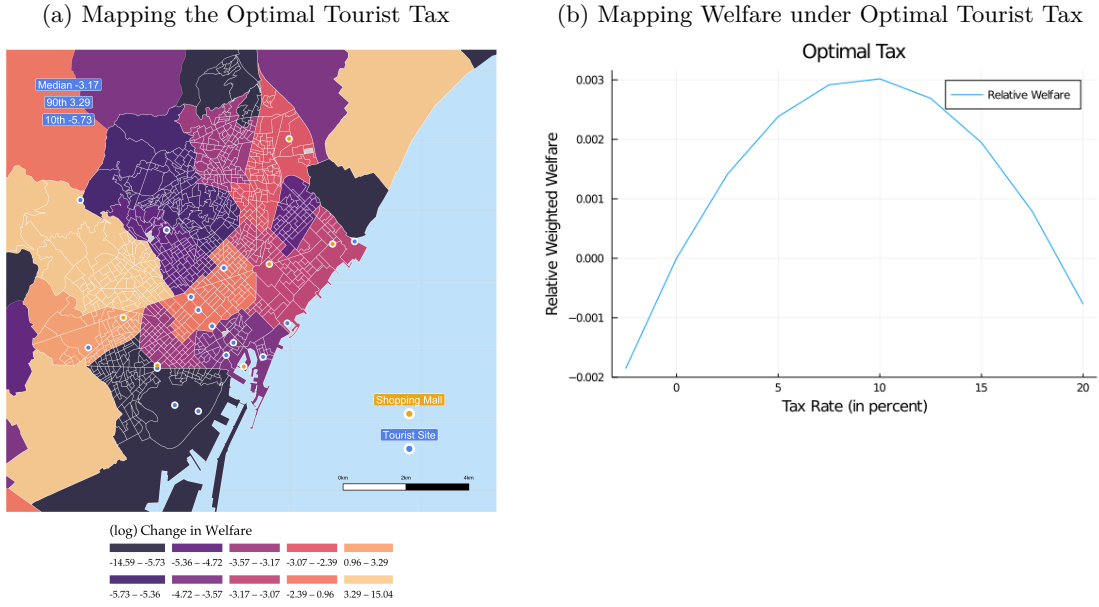
Notes: **(Panel A) (Panel B)**

across locations, where the weights are given by the population size of the different cells, i.e.

$$\arg\max_{t_{is}} \sum \lambda_n U_n = \sum \lambda_n \frac{v_n}{P_n}$$

where $\lambda_n = Population_n = T_n$. Any revenues from the tourist tax are being lump sum distributed to residents, that is,

$$\tilde{v}_n = v_n + transfer_n$$

The solution solves the hat algebra counterfactual equilibrium. The welfare implications are mapped figure 6 as well as the optimal level of the tax.

6 Conclusion

TBD

References

Sumit Agarwal, J. Bradford Jensen, and Ferdinando Monte. Consumer Mobility and the Local Structure of Consumption Industries. NBER Working Papers 23616, National Bureau of Economic Research, Inc, July 2017. URL https://ideas.repec.org/p/nbr/nberwo/23616.html.

Gabriel M Ahlfeldt, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015.

Treb Allen, Costas Arkolakis, and Yuta Takahashi. Universal gravity. *Journal of Political Economy*, 128(2):393–433, 2020.

Milena Almagro and Tomás Domínguez-Iino. Location sorting and endogenous amenities: Evidence from amsterdam. 2019.

Simon P. Anderson, Andre de Palma, and Jacques-Francois Thisse. The ces and the logit: Two related models of heterogeneity. *Regional Science and Urban Economics*, 18(1):155–164, 1988. ISSN 0166-0462. doi: https://doi.org/10.1016/0166-0462(88)90009-9. URL https://www.sciencedirect.com/science/article/pii/0166046288900099.

Oriol Aspachs, Ruben Durante, Alberto Graziano, Josep Mestres, Jose Garcia Montalvo, and Marta Reynal-Querol. Real-time inequality and the welfare state in motion: Evidence from COVID-19 in Spain. Economics Working Papers 1734, Department of Economics and Business, Universitat Pompeu Fabra, July 2020. URL https://ideas.repec.org/p/upf/upfgen/1734.html.

Susan Athey, David Blei, Robert Donnelly, Francisco Ruiz, and Tobias Schmidt. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. *AEA Papers and Proceedings*, 108:64–67, May 2018. doi: 10.1257/pandp.20181031. URL https://www.aeaweb.org/articles?id=10.1257/pandp.20181031.

David Atkin, Benjamin Faber, and Marco Gonzalez-Navarro. Retail Globalization and Household Welfare: Evidence from Mexico. *Journal of Political Economy*, 126(1):1–73, 2018. doi: 10.1086/695476. URL https://ideas.repec.org/a/ucp/jpolec/doi10.1086-695476.html.

David Baqaee and Ariel Burstein. Welfare and output with income effects and demand instability.

Working Paper 28754, National Bureau of Economic Research, May 2021. URL http://www.nber.org/papers/w28754.

David Rezza Baqaee and Emmanuel Farhi. The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem. *Econometrica*, 87(4):1155–1203, July 2019. doi: 10.3982/ECTA15202. URL https://ideas.repec.org/a/wly/emetrp/v87y2019i4p1155-1203.html.

Vasco Carvalho, Juan R. Garcia, Stephen Hansen, Álvaro Ortiz, Tomasa Rodrigo, Sevi Rodriguez Mora, and José Ruiz. Tracking the covid-19 crisis with high-resolution transaction data. WorkingPaper 2020/16, April 2020.

Victor Couture, Jonathan Dingel, Allison Green, and Jessie Handbury. Quantifying social interactions using smartphone data. 2020.

Donald R. Davis, Jonathan I. Dingel, Joan Monras, and Eduardo Morales. How Segregated Is Urban Consumption? *Journal of Political Economy*, 127(4):1684–1738, 2019. doi: 10.1086/701680. URL https://ideas.repec.org/a/ucp/jpolec/doi10.1086-701680.html.

Angus Deaton. Rice Prices and Income Distribution in Thailand: A Non-parametric Analysis. *Economic Journal*, 99(395):1–37, Supplemen 1989. URL https://ideas.repec.org/a/ecj/econjl/v99y1989i395p1-37.html.

Dave Donaldson. Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934, 2018.

Benjamin Faber and Cecile Gaubert. Tourism and economic development: Evidence from mexico's coastline. *American Economic Review*, 109(6):2245–93, June 2019. doi: 10.1257/aer.20161434. URL https://www.aeaweb.org/articles?id=10.1257/aer.20161434.

Miquel-Àngel García-López, Jordi Jofre-Monseny, Rodrigo Martínez Mazza, and Mariona Segú. Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona. Working Papers 2019/05, 2019. URL https://ideas.repec.org/p/ieb/wpaper/doc2019-05.html.

H. S. Houthakker. Compensated Changes in Quantities and Qualities Consumed. *The Review of Economic Studies*, 19(3):155–164, 01 1952. ISSN 0034-6527. doi: 10.2307/2296018. URL https://doi.org/10.2307/2296018.

Ryan Kim and Jonathan Vogel. Trade and Welfare (Across Local Labor Markets). NBER Working Papers 27133, National Bureau of Economic Research, Inc, May 2020. URL https://ideas.repec.org/p/nbr/nberwo/27133.html.

Benny Kleinman, Ernest Liu, and Stephen J Redding. International friends and enemies. Working Paper 27587, National Bureau of Economic Research, July 2020. URL http://www.nber.org/papers/w27587.

Ferdinando Monte, Stephen J. Redding, and Esteban Rossi-Hansberg. Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90, December 2018. doi: 10.1257/aer.20151507. URL https://www.aeaweb.org/articles?id=10.1257/aer.20151507.

Guido G. Porto. Using survey data to assess the distributional effects of trade policy. *Journal of International Economics*, 70(1):140–160, 2006. ISSN 0022-1996. doi: https://doi.org/10.1016/j.jinteco.2005.09.003. URL https://www.sciencedirect.com/science/article/pii/S0022199606000195.

J. M. C. Santos Silva and Silvana Tenreyro. The Log of Gravity. *The Review of Economics and Statistics*, 88(4):641–658, November 2006. URL https://ideas.repec.org/a/tpr/restat/v88y2006i4p641-658.html.

A   Additional Derivations

In this part of the appendix we provide additional derivations that are omitted in the main text. First in subsection A.1 we show how a log-linear gravity regression design can be derived from the non-parametric model as a first order approximation. Second, in subsection A.2 we show similarly, how the income regressions can be motivated non-parametrically. The subsection furthermore provides insights on how to think about the residual in the income regression. Third, in subsection A.3 we show the isomorphism between the more commonly used multinomial choice model with extreme value distributed preference shocks and our income maximization problem. Finally, in subsection A.4 we show how our shift-share instrument can be derived non-parametrically.

A.1   Non-Parametric Expenditure Gravity

In this section we show that the linear log gravity specification can be derived as a first order approximation to its non-linear counterpart. We first start by deriving the gravity function for the non-parametric model of section 2, where indirect utility is given by,

$$u_n = \frac{v_n}{G\left(\boldsymbol{p}_n\right)}$$

where $G(\cdot)$ represents some price aggregator and v_n is the income of the resident and $\boldsymbol{p}_n$ refers to the vector of prices in block n of the goods produced by each block i. Consumption is assumed to require physically travelling to a city block and incurring an iceberg variable trade costs such that prices are given by

$$p_{ni} = \tau_{ni} p_i$$

where $\tau_{ni} \geq 1$ refers to the iceberg variable trade costs and where we assume that $\tau_{nn} = 1$. From Roy's identity, demand in block n for the good produced in country i is given by,

$$q_{ni}\left(\boldsymbol{p}_n\right) = -G\left(\boldsymbol{p}_n\right) v_n \times \frac{\partial(1/G\left(\boldsymbol{p}_n\right))}{\partial p_{ni}}$$

We can derive expenditure shares as,

$$\frac{p_{ni}q_{ni}(\boldsymbol{p}_n)}{v_n} = -p_{ni} \times G(\boldsymbol{p}_n) \times \frac{\partial(1/G(\boldsymbol{p}_n))}{\partial p_{ni}}$$

$$= \frac{\partial \ln G(\boldsymbol{p}_n)}{\partial \ln p_{ni}}$$

Taking logs we obtain,

$$\ln s_{ni} = \ln\left(\frac{\partial \ln G(\boldsymbol{p}_n)}{\partial \ln p_{ni}}\right)$$

A first order Taylor expansion around $p_{ni} = 1$ yields:

$$\ln s_{ni} = \left(\sum_j \frac{\partial \ln}{\partial \ln p_{nj}}\left(\frac{\partial \ln G(\boldsymbol{p}_n)}{\partial \ln p_{ni}}\right)\ln(p_{nj})\right) + \varepsilon_{ni},$$

where ϵ_{ni} is an approximation error that summarizes higher order terms of the Taylor expansion. Then applying $p_{ni} = \tau_{ni} \times p_i$ we have:

$$\ln s_{ni} = \sum_j \left(\alpha_{nj}\ln(\tau_{nj}) + \alpha_{nj}\ln p_j\right) + \varepsilon_{ni} \iff$$

$$\ln s_{ni} = \underbrace{\alpha_{ni}\ln(\tau_{ni})}_{\text{pair FE } \delta_{ni}} + \underbrace{\alpha_{ni}\ln(p_i)}_{\text{destination FE } \delta_n} + \underbrace{\sum_{j\neq i}\left(\alpha_{nj}\ln(\tau_{nj}) + \alpha_{nj}\ln(p_j)\right) + \alpha_i\ln(p_i) - \alpha_i\ln(p_i)}_{\text{origin FE } \delta_n} + \varepsilon_{ni} \iff$$

$$\ln s_{ni} = \underbrace{\alpha_{ni}\ln(\tau_{ni})}_{\text{pair FE } \delta_{ni}} + \underbrace{\alpha_i\ln(p_i)}_{\text{destination FE } \delta_i} + \underbrace{\sum_{j\neq i}\left(\alpha_{nj}\ln(\tau_{nj}) + \alpha_{nj}\ln(p_j)\right) + (\alpha_{ni} - \alpha_i)\ln(p_i)}_{\text{origin FE } \delta_n} + \varepsilon_{ni}$$

where $\alpha_{nj} \equiv \frac{\partial \ln}{\partial \ln p_{nj}}\left(\frac{\partial \ln G(\boldsymbol{p}_n)}{\partial \ln p_{ni}}\right)$ measures the local substitutability of different goods in the consumer's utility maximization problem.

A.2 Non-Parametric Income Changes

We first start by deriving the gravity function for the non-parametric model of section 2, where the indirect time use function is given by,

$$h_n = \frac{v_n}{J(\boldsymbol{w}_n)}$$

where $J(\cdot)$ represents some wage aggregator and v_n is the income of the resident and $\boldsymbol{w}_n$ refers to the vector of wages that are accessible via commuting linkages to residents in block n. Supplying labor to a location is assumed to require physically travelling to a city block and incurring an iceberg variable commuting costs such that wages are given by,

$$w_{ni} = \mu_{ni} w_i$$

where $\mu_{ni} \geq 1$ refers to the iceberg variable commuting costs and where we assume that $\mu_{nn} = 1$. From the commuting equivalent of Roy's identity 2, labor supply from residents in block n to production in block i is given by,

$$\ell_{ni}(\boldsymbol{w}_n) = -J(\boldsymbol{w}_n) v_n \times \frac{\partial(1/J(\boldsymbol{w}_n))}{\partial w_{ni}}$$

We can derive income shares as,

$$\begin{aligned}
\frac{w_{ni} \ell_{ni}(\boldsymbol{w}_n)}{v_n} &= -w_{ni} \times J(\boldsymbol{w}_n) v_n \times \frac{\partial(1/J(\boldsymbol{w}_n))}{\partial w_{ni}} \\
&= \frac{\partial \ln J(\boldsymbol{w}_n)}{\partial \ln w_{ni}}
\end{aligned}$$

Total income changes are given by,

$$v_n = \sum_i c_{ni} \times w_i,$$

Totally differentiating,

$$dv_n = \sum_i c_{ni} \frac{dw_i}{dz_i} dz_i + \sum_i c_{ni} \frac{dw_i}{dE_i^T} dE_i^T,$$

A.3 Isomorphism with Multinomial Choice with Frechet Preference Shocks

Parallel to Anderson et al. (1988), individual weibit consumer has indirect utility

$$v_n = \frac{y_n}{P_n}$$

where y_n is the disposable income of an individual in location n and P_n is the local price index.

Disposable income depends on the individual's labor supply decision. Conditional on providing labor to location i the individual derives the following income,

$$y_i = T_n w_{ni} = \frac{T_n}{c_{ni}} w_i$$

where T_n is the time endowment, c_{ni} is the bilateral migration cost, and w_i is the wage rate in location i. The indirect utility of an individual in n deciding to supply labor to location i is given by,

$$v_{ni} = \frac{1}{P_n} \left(\frac{T_n}{c_{ni}} w_i \right)$$

An individual is making a stochastic choice,

$$\max_{v_{ni}} v_{ni} \times \epsilon_{ni}$$

where $\epsilon_{ni} \sim \text{Frechet}(\theta, 1)$ is the stochastic preference shock. By the standard properties of the Frechet distribution, the choice probability is given by,

$$P_{ni} = \frac{(w_{ni})^{\theta}}{\sum_k (w_{nk})^{\theta}}$$

Comparing with the time use share above,

$$\frac{(w_{ni})^{1-\epsilon}}{\sum_i (w_{ni})^{1-\epsilon}}$$

we notice that this coincides with the time use share above, if $\theta = 1 - \epsilon$.

A.4 Non-Parametric Bartik Instrument exploiting Group-specific Heterogeneity

The city is visited by tourists from many countries, indexed by $g \in \{1, \ldots, G\}$. The representative tourist from country g has homothetic demand that can be represented by the following indirect utility function:

$$u_g = \frac{E_g^T}{G(\tilde{\boldsymbol{p}})}$$

where $G(\cdot)$ represents some price aggregator, E_g^T is the fixed expenditure of tourist from country g, and $\boldsymbol{p}$ refers to the vector of prices across all production city blocks. Consumption has both a pecuniary and a non-pecuniary cost and benefit, where we parameterize the non-pecuniary benefit to be multiplicative,

$$\tilde{p}_{gi} = b_{gi} p_i$$

where b_{gi} is the non-pecuniary benefit of consuming in location i for group g and p_i refers to the price that is common to all groups. From Roy's identity, demand for group g for the good produced in block i is given by,

$$q_{gi}(\tilde{\boldsymbol{p}}) = -G(\tilde{\boldsymbol{p}}) E_g^T \times \frac{\partial(1/G(\tilde{\boldsymbol{p}}))}{\partial \tilde{p}_{gi}}$$

$$s_{gi} = \frac{b_{gi} p_{gi} \left(\frac{\partial(G(\boldsymbol{p}_g))}{\partial p_{gi}}\right)}{\sum_{\ell=1}^{N} b_{g\ell} p_{n\ell} \left(\frac{\partial(G(\boldsymbol{p}_n))}{\partial p_{nl}}\right)}$$

Multiplying and dividing by $\frac{p_{ni}}{G(\boldsymbol{p}_n)}$ we obtain,

$$s_{gi} = \frac{b_{gi} G(\boldsymbol{p}_g) \left(\frac{p_{gi}}{G(\boldsymbol{p}_g)} \frac{\partial(G(\boldsymbol{p}_g))}{\partial p_{gi}}\right)}{\sum_{\ell=1}^{N} b_{g\ell} G(\boldsymbol{p}_g) \left(\frac{p_{g\ell}}{G(\boldsymbol{p}_g)} \frac{\partial(G(\boldsymbol{p}_g))}{\partial p_{gl}}\right)}$$

$$s_{gi} = \frac{b_{gi} \epsilon_{gi}}{\sum_{\ell=1}^{N} b_{gi} \epsilon_{gl}}$$

where $\frac{p_{ni}}{G(\boldsymbol{p}_n)} \frac{\partial(G(\boldsymbol{p}_n))}{\partial p_{ni}} \equiv \epsilon_{ni}$ is the demand elasticity, locally measured at the point of consumption. Total tourist spending in a given location can thus be expressed as:

$$X_i^T = \sum_g E_g^T \times s_{gi}, \tag{14}$$

Taking the sources of exogenous variation to be the total spending of each group, group-specific consumption shares, and preferences, we totally differentiate equation (14) to obtain an explicit expression for the sources of exogenous variation in terms of initial shares and group-specific spending variation:

$$dX_i^T = \sum_g s_{gi} dE_g^T + \sum_g s_{gi} db_{gi}$$

Taking it to the data, we construct changes in location i's tourist expenditures $g_{imt}^T \equiv \frac{\Delta E_{imt}^T}{E_i^T}$ as:

$$\Delta E_{imt}^T \quad = \quad \underbrace{\sum_g s_{gi} \times g_{E_g t}^T}_{\text{Group Composition}} \quad + \epsilon_{imt}^T \tag{15}$$

where $g_{E_g t}^T \equiv \Delta E_{gt}^T$ denote changes in total group's income; the $s_{gi} \equiv \frac{E_{igt}^T}{E_{gt}^T}$ captures spatial shares of tourist expenditure for group g. Notice, that our structural derivation implies that the unobserved confounder corresponds to group specific amenity changes, i.e.

$$\epsilon_{imt}^T = \sum_g s_{gi} db_{gi}$$

We use (15) to define our instrument exploiting group composition. We define the *initial* shares to be orthogonal to seasonal demand shocks by building averages of the shares over the full period available in our data leaving out the current month for which the change is calculated.
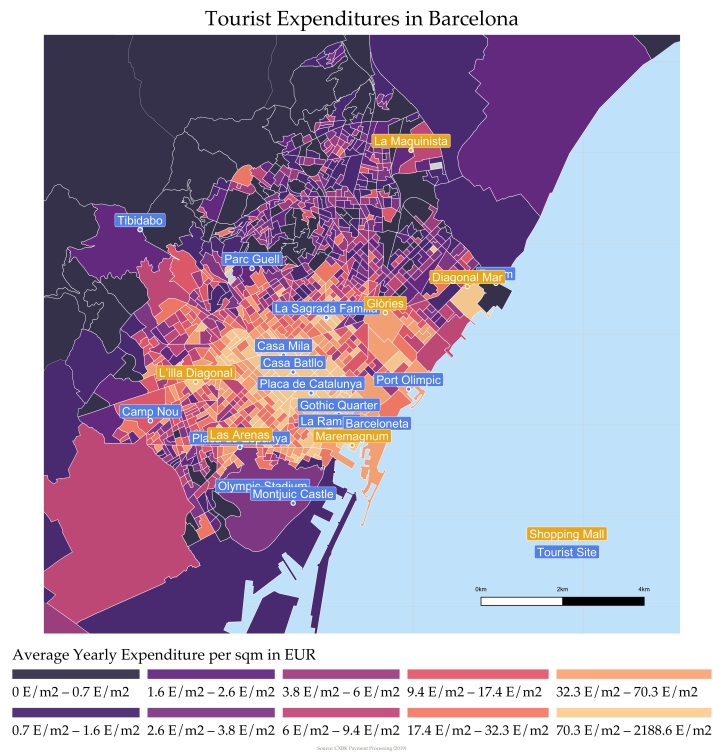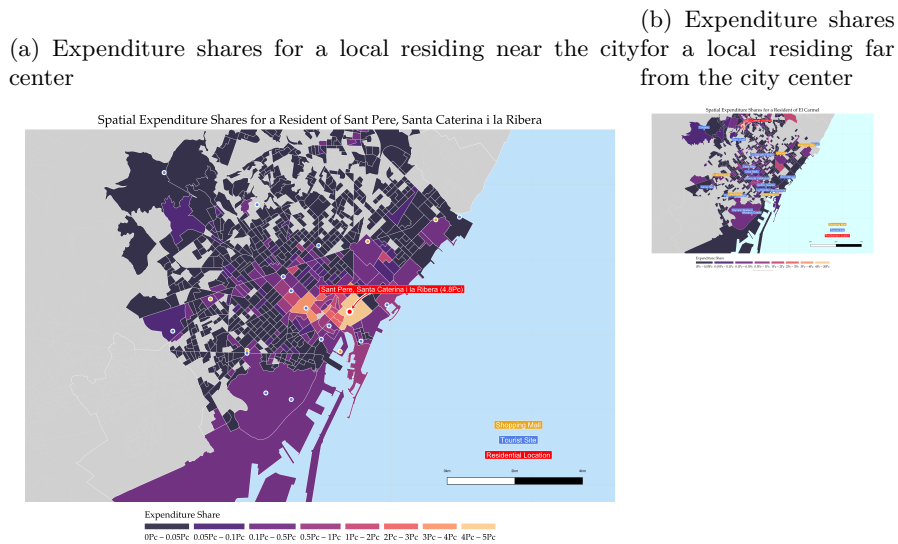
B   Figures

Figure 7: Signs in Barcelona

Notes: Source: "Why Barcelona Locals Really Hate Tourists." *The Independent.* August 9, 2017.

Figure 8: Tourists spend disproportionately more in the city center

Notes: This figure shows the average yearly expenditure (normalized per square meter) in euros by tourists throughout the city of Barcelona.

Figure 9: Locals spend more near their home

(a) Expenditure shares for a local residing near the city center

(b) Expenditure shares for a local residing far from the city center

Notes: This figure compares expenditure patterns for locals residing in different areas of the city. The left panel is the expenditure shares for a resident of Sant Pere, Santa Caterina i la Ribera (near the city center). The first panel is the same included in 2a, we repeat it here for comparison with different areas of the city. The right panel is for El Carmel (far from the city center).

Figure 10: Residents spend more near their home, although the impact of distance is heterogeneous across sectors

Distance Elasticity

Source: CXBK Payment Processing (2019)
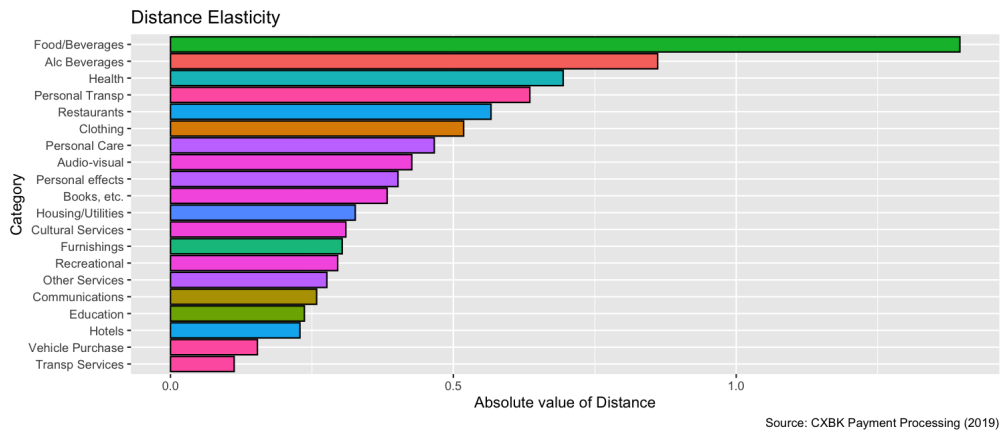
Notes: This figure shows the impact of distance on expenditure by sector. The distance coefficient is estimated using sector-specific gravity regression of local expenditure shares on bilateral travel times with origin-sector-month and destination-sector-month fixed effects.

Figure 11: Mean Income across Barcelona
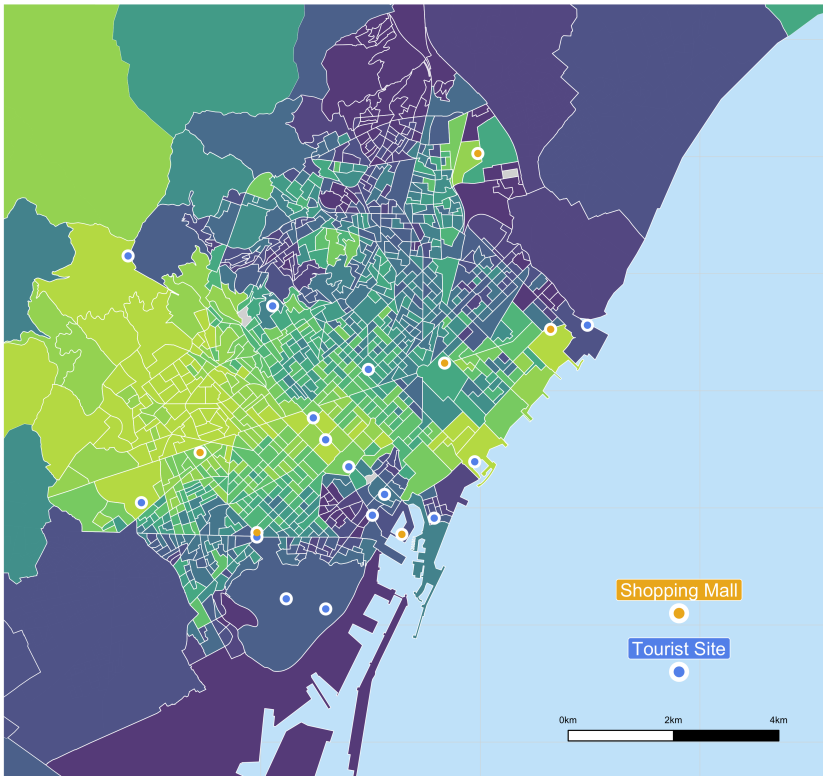
Mean Income

1039.61 – 1260.88	1421.98 – 1486.94	1585.91 – 1623.15	1705.59 – 1767.53	1956.66 – 2132.63
1260.88 – 1352.46	1486.94 – 1541.06	1623.15 – 1662.96	1767.53 – 1859.12	2132.63 – 2396.31
1352.46 – 1421.98	1541.06 – 1585.91	1662.96 – 1705.59	1859.12 – 1956.66	2396.31 – 11806.33

Notes:

Figure 12: Change in Housing Prices between low and high Season

Change in House Price

| -27.47 – -2.9 | -1.41 – -0.78 | -0.19 – 0.12 | 0.51 – 0.83 | 1.17 – 1.48 | 1.84 – 2.16 | 2.54 – 3.18 | 3.96 – 18.3 |

| -2.9 – -1.41 | -0.78 – -0.19 | 0.12 – 0.51 | 0.83 – 1.17 | 1.48 – 1.84 | 2.16 – 2.54 | 3.18 – 3.96 |
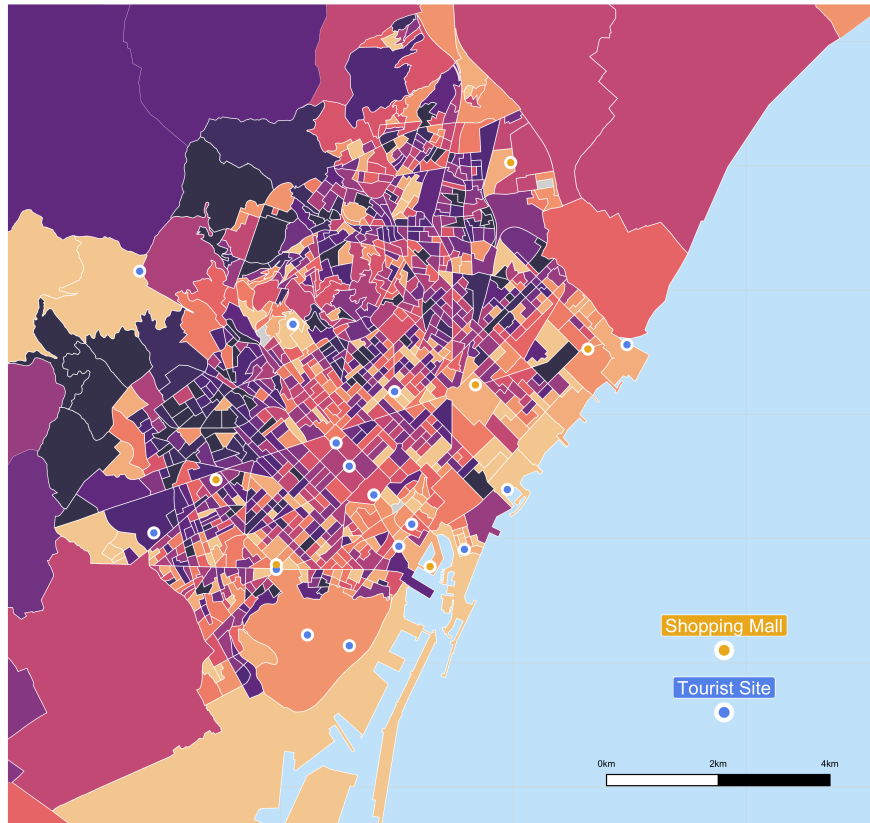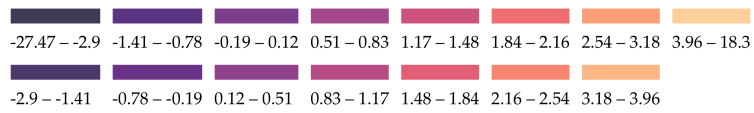
Notes:

Figure 13: Gravit BinScat

Notes:

C Additional Tables

Table 6: Sum Stats

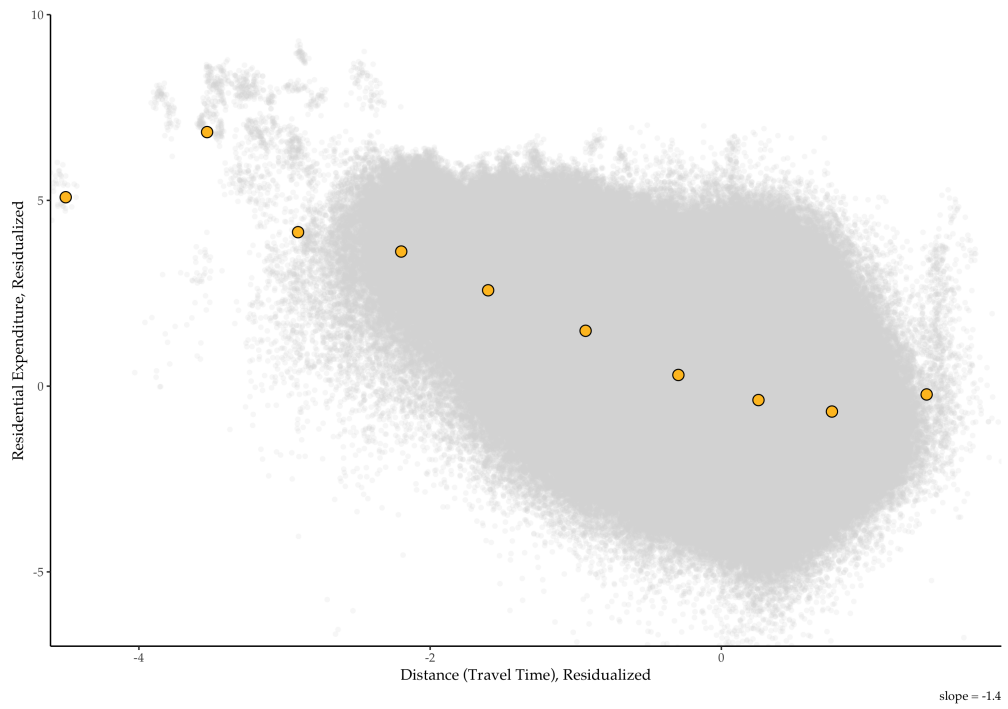|	Local	Spanish Tourists	Foreign Tourists	Total
Total	1593.60 (54)	314.07 (11)	1062.589 (36)	2970.26
Jan	142.31 (63)	24.06 (11)	60.53 (27)	226.90 (100)
Feb	125.63 (59)	21.81 (10)	66.93 (31)	214.36 (100)
Mar	143.02 (58)	25.57 (10)	79.38 (32)	247.97 (100)
Apr	135.99 (52)	26.98 (10)	97.05 (37)	260.02 (100)
May	146.34 (53)	28.16 (10)	104.01 (37)	278.50 (100)
Jun	145.43 (53)	28.05 (10)	101.05 (37)	274.54 (100)
Jul	149.24 (50)	32.83 (11)	118.40 (39)	300.47 (100)
Aug	101.74 (41)	27.83 (11)	116.46 (47)	246.03 (100)
Sep	117.89 (49)	23.97 (10)	96.55 (40)	238.41 (100)
Oct	122.80 (51)	23.77 (10)	93.40 (39)	239.97 (100)
Nov	124.67 (57)	24.04 (11)	68.46 (32)	217.17 (100)
Dec	138.55 (61)	27.01 (12)	60.37 (27)	225.92 (100)

Notes:

Table 7: Summary Statistics: Total Average Total Expenditure 2-Digit COICOP

COICOP (2D)	COICOP (2D)	Local	Spanish Tourists	Foreign Tourists	Total	Survey (INE)	Survey Adj (INE)
11	Food/Beverages	32.82 (24.72)	1.32 (5.04)	4.51 (5.10)	38.66	12.96	23.82
21	Alc Beverages	1.97 (1.48)	0.07 (0.28)	0.60 (0.68)	2.64	0.71	1.31
31	Clothing	11.58 (8.72)	1.94 (7.39)	12.00 (13.55)	25.51	3.39	6.23
41	Housing/Utilities	2.81 (2.12)	0.78 (3.00)	0.59 (0.67)	4.19	5.33	9.80
51	Furnishings	10.03 (7.55)	3.32 (12.67)	2.01 (2.27)	15.35	0.88	1.62
61	Health	10.76 (8.10)	1.94 (7.40)	1.82 (2.06)	14.52	2.24	4.12
71	Vehicle Purchase	3.14 (2.36)	0.18 (0.67)	0.32 (0.36)	3.63	3.78	6.95
72	Personal Transp	7.27 (5.47)	2.06 (7.89)	0.70 (0.79)	10.03	6.38	11.73
73	Transp Services	10.13 (7.63)	6.52 (24.90)	9.61 (10.85)	26.26	1.90	3.49
81	Communications	0.30 (0.23)	0.02 (0.09)	0.08 (0.09)	0.40	0.33	0.61
91	Audio-visual	5.06 (3.81)	0.57 (2.17)	1.78 (2.01)	7.40	0.58	1.07
93	Recreational	2.62 (1.97)	0.27 (1.03)	1.21 (1.37)	4.09	1.43	2.63
94	Cultural Services	4.29 (3.23)	0.62 (2.38)	2.79 (3.15)	7.70	0.57	1.05
95	Books, etc	1.64 (1.23)	0.22 (0.85)	0.53 (0.60)	2.39	1.30	2.39
101	Education	1.11 (0.84)	0.10 (0.39)	0.61 (0.69)	1.82	0.77	1.41
111	Restaurants	17.73(13.35)	3.79 (14.46)	19.04 (21.50)	40.56	7.83	14.39
112	Hotels	1.13 (0.85)	1.49 (5.69)	23.12 (26.11)	25.75	1.21	2.22
121	Personal Care	4.84 (3.64)	0.32 (1.23)	0.97 (1.10)	6.14	2.53	4.65
123	Other	2.49 (1.88)	0.36 (1.37)	5.69 (6.42)	8.54	0.32	0.59
Total		131.72 (100)	25.88 (100)	87.97 (100)	245.58	54.4	100

Notes: The table shows the average total expenditures (in million Euros) per COICOP category and across groups. The groups are aggregated to reflect our notion of locals (CXBK and non-CXBK customers), foreign tourists (transcation utilizing a credit card with a foreign issuer) and domestic Spanish tourists (cards that have their largest expenditure outside of the province of Barcelona). We also report the corresponding expenditure share in the expenditure survey by INE for Catalonia. Since our consumption categories only add up to 54.4pc of total expenditures observed in the INE surveys, we construct an adjusted expenditure share measure from the surveys that accounts for this and is directly comparable to our expenditure shares.

Table 8: Commuting Gravity

Dependent Variables:	commuters	log(commuters+1)	log(commuters)	transactions	log(transactions+1)	log(transactions)
	Cell Phone			Lunchtime		
Model:	(1) Poisson	(2) OLS	(3) OLS	(4) Poisson	(5) OLS	(6) OLS
Variables						
ldist	-4.48***	-1.51***	-1.17***	-1.53***	-0.134***	-0.411***
	(0.107)	(0.037)	(0.054)	(0.028)	(0.002)	(0.012)
Fixed-effects						
Origin	✓	✓	✓			
Destination	✓	✓	✓			
Origin (CT)				✓	✓	✓
Destination (CT)				✓	✓	✓
Fit statistics						
Observations	24,025	24,025	2,162	1,051,159	1,216,609	42,086
Pseudo R^2	0.798	0.117	0.193	0.598	0.343	0.091

Heteroskedasticity-robust standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Notes: Standard errors are two-way clustered at origin and destination. Columns (1) and (3) are estimated using Pseudo-Poisson Maximum Likelihood, where commuting flows are: $E(\lambda_{ni}) = \exp(\alpha \log(d_{ni}) + \gamma_n + \delta_i)$. Columns (2) and (4) are estimated using Ordinary Least Squares, where commuting flows are: $\log(\lambda_{ij}) = \alpha \log(d_{ni}) + \gamma_n + \delta_i + \epsilon_{ni}$. Distances in minutes are computed using the simple average of transit times over commuting hours using a car and public transit. Travel times within a location is normalized to 2 minutes.

Table 9: Housing Price and Rental Elasticities

Dependent Variables:	HPRICE				RENT			
	IV - Ref: 2017 Average		IV - Ref: 2017 Low Season		IV - Ref: 2017 Average		IV - Ref: 2017 Low Season	
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
$\widehat{\log E_{it}^T}$	0.059***	0.028***	0.059***	0.028***	0.043***	0.008*	0.044***	0.009*
	(0.016)	(0.005)	(0.016)	(0.005)	(0.013)	(0.005)	(0.013)	(0.005)
Fixed-effects								
i (108)	✓	✓	✓	✓	✓	✓	✓	✓
i×month (1,296)	✓		✓		✓		✓	
i×year (216)		✓		✓		✓		✓
Fit statistics								
Observations	2,592	2,592	2,592	2,592	2,592	2,592	2,592	2,592
Adjusted R^2	0.983	0.993	0.983	0.993	0.933	0.952	0.933	0.952

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes:

Table 10: By Sector

Notes: