March 2026

Urban Welfare: Tourism in Barcelona[1]

Preliminary and Incomplete, Do Not Circulate

Treb Allen
Dartmouth and NBER

Alberto Graziano
CaixaBank Research

Simon Fuchs
Federal Reserve Bank of Atlanta

Rocio Madera
Southern Methodist Univ.

Sharat Ganapati
Georgetown University and NBER

Judit Montoriol-Garriga
CaixaBank Research

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.

KEYWORDS: Sufficient Statistic, Tourism, Quantitative Spatial Economics, Heterogeneous Effects, SUTVA
JEL CLASSIFICATION CODES: E31, F14

[1]Contact: treb.allen@dartmouth.edu, simon.fuchs@atl.frb.org (corresponding author), sg1390@georgetown.edu, romadera@smu.edu.

The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System or CaixaBank. All errors are our own.

1 Introduction

New data sources describe urban activity in astonishing detail, tracking activity, spending, and income over space and time. Simultaneously, powerful quantitative spatial equilibrium models enables the calculation of 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, for example, 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 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 residents' optimization over consumption and commuting choices with a simple market-clearing condition, we show how one can calculate the (short-run) welfare impacts of urban shocks using a regression-based approach, 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 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 initial 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 resident's 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 has 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—has 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, 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 expenditures 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 (and also 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 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 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, for example, by differences in the timing 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 have a strong spatial component, where locals are much more likely to purchase goods and work near 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 when tourism expenditure rises, but that local expenditure falls the most in these same locations—suggesting that tourism both increases incomes earned by locals and 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 for 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 on income is also sizable and heterogeneous across the city, with inner-city dwellers benefiting the most. We also examine housing 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 quantitative Ricardo–Viner model of the city with rich geographic patterns of consumption and commuting. This framework serves two purposes. First, it lets us compare the welfare estimates from our regression-based approach to those implied by a standard structural model, and we find that the two are closely aligned. Second, it provides a disciplined environment for counterfactual policy analysis. We use it to characterize optimal place-varying tourist taxes that internalize the heterogeneous incidence of tourism across locations. The linear and nonlinear tax formulas produce similar spatial rankings of where taxes should be higher, but the nonlinear model determines their equilibrium scale and captures the curvature that arises once diversion and general-equilibrium feedback are taken into account.

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 in 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 by building on and extending results from the trade and consumption literature that rely on first-order approximations of the indirect utility function to express welfare changes in terms of expenditure-share-

weighted compensating variations (Houthakker, 1952; Atkin, Faber and Gonzalez-Navarro, 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 datasets 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, Jensen and Monte (2017) for 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[2]. 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, Arkolakis and Takahashi (2020) for trade models and closely related to work by (Baqaee and Farhi, 2019; Kleinman, Liu and Redding, 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, Redding and Rossi-Hansberg (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.

2 A Tractable Urban Model for Welfare Evaluations of Small Shocks

We now develop the theoretical framework that underpins our empirical strategy. The goal is to derive a transparent mapping from tourist demand shocks to resident welfare. The first step is to show that, under standard optimization and market-clearing conditions, welfare can be written in terms of wage and price pass-through alone, without imposing ex ante distributional assumptions on preferences or behavior. The second step is to show how tourist spending shocks propagate through the local economy, generating a general-equilibrium object that summarizes their effects on wages, prices, and welfare. This same object will guide the empirical implementation below.

[2] Adão, Arkolakis and Ganapati (2020) show how critical function form assumptions are for both identification and in linking cross-sectional heterogeneity to aggregate outcomes.

2.1 Setting

Consider a city partitioned into N blocks, indexed by $i \in \{1, \ldots, N\}$. Each block i produces a differentiated variety that sells at the producer price p_i.

A resident of origin block n has a time endowment T_n each period and can work in any production block i. Let $\ell_{ni} \geq 0$ denote labor supplied by residents of n in workplace i, and let w_i denote the wage paid in block i. We summarize commuting and work time with a function

$$H_n(\ell_{n1}, \ldots, \ell_{nN}) \leq T_n,$$

where $H_n(\cdot)$ is continuous, weakly convex, and homogeneous of degree one. These assumptions imply a well-behaved income maximization problem. Let

$$V_n(\mathbf{w}, T_n) \equiv \max_{\ell_{n1}, \ldots, \ell_{nN} \geq 0} \left\{ \sum_{i=1}^{N} w_i \ell_{ni} : H_n(\ell_{n1}, \ldots, \ell_{nN}) \leq T_n \right\},$$

where $\mathbf{w} = (w_1, \ldots, w_N)'$ collects wages across blocks. Because the commuting technology is homogeneous of degree one, this value function is linear in the time endowment. We can therefore write

$$V_n(\mathbf{w}, T_n) = T_n J_n(\mathbf{w}),$$

where $J_n(\mathbf{w})$ is the commuting-adjusted wage index for residents of n. Disposable income is then

$$v_n = T_n J_n(\mathbf{w}).$$

The envelope theorem implies that the derivative of V_n with respect to wage w_i equals optimal labor supplied to workplace i. This lets us define

$$c_{ni} \equiv \frac{w_i \ell_{ni}}{v_n},$$

the share of resident n's labor income earned in block i. The vector $\mathbf{c}_n = (c_{n1}, \ldots, c_{nN})$ therefore describes where the resident earns income across the city.

On the consumption side, residents must physically access the places where goods are sold. Let

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

denote the delivered price paid by a resident of n for the variety produced in block i, where $\varphi_{ni} \geq 1$

captures shopping frictions such as travel time, convenience, or access.[3] Preferences are homothetic, so indirect utility takes the form

$$u_n = \frac{v_n}{G(\mathbf{p}_n)},$$

where $\mathbf{p}_n = (p_{n1}, \ldots, p_{nN})'$ is the vector of delivered prices faced by residents of n, and $G(\mathbf{p}_n)$ is the associated price index. Let q_{ni} denote consumption of the variety from block i. Shephard's lemma or Roy's identity then implies

$$s_{ni} \equiv \frac{p_{ni} q_{ni}}{\sum_{j=1}^{N} p_{nj} q_{nj}},$$

the share of resident n's expenditure directed to block i. The vector $\mathbf{s}_n = (s_{n1}, \ldots, s_{nN})$ therefore describes where that resident shops across the city.

The household side of the model is thus summarized by two observable baseline share vectors. The income-share vector $\mathbf{c}_n$ tells us where resident n earns income; the expenditure-share vector $\mathbf{s}_n$ tells us where resident n spends. These two objects will be the weights that translate wage and price changes into welfare.

2.2 From tourist shocks to wages, prices, and welfare

We study small, short-run tourist demand shocks, holding technology, capital rental rates, shopping frictions, and baseline shares fixed. Let E_i^T denote tourist spending in block i, and let

$$\mathbf{g} \equiv d \ln \mathbf{E}^T = \left(d \ln E_1^T, \ldots, d \ln E_N^T \right)'$$

denote the vector of log changes in tourist spending across blocks. Our objective is to map this shock vector into resident welfare.

Because resident n's utility is $u_n = v_n / G(\mathbf{p}_n)$, a first-order log change in welfare can be written as the log change in nominal income minus the log change in the price index. Homotheticity and the envelope theorem translate those derivatives into the observed income and expenditure shares introduced above. As a result,

$$d \ln u_n = \sum_{i=1}^{N} c_{ni} \, d \ln w_i - \sum_{i=1}^{N} s_{ni} \, d \ln p_i. \tag{T.1}$$

Equation (T.1) is the key welfare envelope. A wage increase in block i raises resident n's welfare in proportion to the share of income that resident earns there, c_{ni}. A price increase in block i lowers welfare in proportion to the share of expenditure directed there, s_{ni}. The problem is therefore reduced

[3] We use φ_{ni} for shopping frictions to reserve τ_i for tourist taxes later in the paper.

to characterizing how tourist shocks change wages and prices.

To do so, let $y_i = p_i q_i$ denote nominal sales in block i. Goods-market clearing requires those sales to equal spending by residents plus spending by tourists:

$$y_i = \sum_{n=1}^{N} s_{ni} v_n + E_i^T.$$

The term v_n is resident income, and E_i^T is tourist spending in block i. Differentiating this expression and using the fact that resident income changes with the wages residents earn across workplaces yields a linear relationship between block sales and wages. It is useful to summarize that relationship with two objects.

The first is the matrix $\mathbf{M}$, defined element by element as

$$M_{ij} \equiv \sum_{n=1}^{N} \left(\frac{s_{ni} v_n}{y_i} \right) c_{nj}.$$

The element M_{ij} measures a one-round GE feedback: it tells us how a wage increase in block j raises resident income for those who work there, how that extra income is spent across shopping destinations, and how much of it ends up increasing sales in block i. For this reason, $\mathbf{M}$ is the one-round propagation matrix.

The second object is the diagonal matrix

$$\mathbf{D}_T \equiv \operatorname{diag}\left(\frac{E_1^T}{y_1}, \ldots, \frac{E_N^T}{y_N} \right).$$

Its diagonal entries measure the tourist intensity of each block: the share of local sales accounted for by tourists. This matrix tells us where the tourist shock enters the local economy directly.

To connect sales to wages and prices, we impose a short-run production structure. Under Cobb–Douglas technology with labor share $\beta \in (0, 1)$, producer prices move with wages according to

$$d \ln p_i = \beta \, d \ln w_i.$$

With fixed local labor in the short run, a proportional increase in sales raises the local wage bill proportionally, so wage changes move one-for-one with sales changes. Stacking these relationships across blocks gives

$$d \ln \mathbf{w} = \mathbf{M} \, d \ln \mathbf{w} + \mathbf{D}_T \, \mathbf{g}.$$

If the spectral radius satisfies $\rho(\mathbf{M}) < 1$, the matrix $I - \mathbf{M}$ is invertible. This lets us solve explicitly for the GE effect of tourist shocks on wages. Define

$$\mathbf{\Pi} \equiv (I - \mathbf{M})^{-1}\mathbf{D}_T.$$

We refer to $\mathbf{\Pi}$ as the percolation operator. It maps a vector of tourist-demand shocks into the equilibrium wage response after summing the direct effect and all higher-order rounds of income and spending feedback. Wage and price pass-through can then be written compactly as

$$d\ln \mathbf{w} = \mathbf{\Pi}\,\mathbf{g}, \qquad d\ln \mathbf{p} = \beta\,\mathbf{\Pi}\,\mathbf{g}. \tag{T.2}$$

The intuition is straightforward. The matrix $\mathbf{D}_T$ injects a tourist shock where it is initially spent. The matrix $\mathbf{M}$ propagates that shock through the city by tracing how higher income in one location becomes higher spending in another. The operator $\mathbf{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T$ accumulates all of those rounds.

Substituting equation (T.2) into the welfare envelope in equation (T.1) gives the effect of tourist shocks on the welfare of residents in origin n:

$$d\ln u_n = \big(\mathbf{c}_n - \beta\,\mathbf{s}_n\big)\mathbf{\Pi}\,\mathbf{g}. \tag{T.3}$$

Equation (T.3) is the central theoretical result. The vector $\mathbf{\Pi g}$ is the citywide GE index of the tourist shock: it summarizes how the shock changes wages across blocks, and, through pass-through, prices as well. Resident n loads on that index with the net-incidence vector $\mathbf{c}_n - \beta\,\mathbf{s}_n$. The resident benefits where the shock raises wages in places she works and loses where it raises prices in places she shops.

This expression also aggregates naturally to the citywide welfare level. Let λ_n denote the population weight on residents of origin n, and define aggregate welfare as $W = \sum_n \lambda_n \ln u_n$. Then

$$dW = \boldsymbol{\psi}\,\mathbf{\Pi}\,\mathbf{g}, \qquad \boldsymbol{\psi} \equiv \sum_n \lambda_n\big(\mathbf{c}_n - \beta\,\mathbf{s}_n\big). \tag{T.4}$$

The row vector $\boldsymbol{\psi}$ is the citywide net-incidence vector. It summarizes which locations matter for aggregate welfare once we combine residents' exposure on the income side and the consumption side.

The theory therefore reduces the welfare consequences of tourism shocks to a small set of economically interpretable objects. The vectors $\mathbf{c}_n$ and $\mathbf{s}_n$ describe where residents earn and spend. The matrix $\mathbf{M}$ captures one round of GE feedback across neighborhoods. The matrix $\mathbf{D}_T$ measures where tourism enters the local economy directly. The percolation operator $\mathbf{\Pi}$ then combines these ingredients to map tourist shocks into wages, prices, and welfare. In the empirical sections that follow, we take this map-

ping to the data and use it to measure welfare changes before turning to the fuller quantitative model and the tax analysis.

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 185EUR per day, totaling 20.6 billion euros 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 an open question, in part due to limited availability of economic activities at a detailed level. While the largest share of tourism expenditures may be 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.

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 we construct from electronic payments processed by CaixaBank. CaixaBank is the largest bank in Barcelona and the leading bank among individuals and SMEs in Spain. It is based in Barcelona, where it has close to a 40% market share. This database describes 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 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 comparative full quantitative exercise.

3.1 A New High-Resolution Spatial Panel of Expenditure and Income

In this section, we describe how we build a spatial network of monthly income flows for locals and tourists combining data from CaixaBank and official statistics.

3.1.1 Expenditures for Locals and Tourists: Sources

Our spending data comes from CaixaBank's Payment Processing Service. We use transaction-level data from the electronic payments aggregated to census-tract totals for the period 2017-2019 at a monthly frequency. 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 (ambito metropolitano 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).

3.1.2 Income of Locals: Sources

We combine two sources of disaggregated income statistics: (1) Our main source of monthly income is the anonymized monthly payrolls and unemployment benefits deposited in personal bank accounts in CaixaBank, aggregated at the census tract level, for the same residents, and for the same time period for which we observe expenditures. From this source, we use average income for each census tract and month. (2) We compare and combine our data with official disaggregated accounts from the National Statistics Institute (INE)'s *Income Distribution Atlas*.[4] The two sources yield very similar average annual incomes (See Appendix A.2). While this is not entirely surprising (CaixaBank has the highest market share in direct-deposit payrolls in the country, with almost a third), it confirms that our sample of residents is not particularly selected within a census-tract.

[4] Atlas de Distribución de Renta de los Hogares.

3.1.3 Spatial Network of Locals

Moving beyond cross-sectional spatial distributions, to understand the propagation of income flows for locals we need two networks: spending and commuting. Using the location of residence in the data described above, we can identify origin-destination spending shares for locals. This is our spending network.

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.

3.1.4 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, including rental and housing prices. The sample for the housing data covers the period between January 2010 and June 2020 at a monthly frequency.

3.1.5 Distance Measures

We calculate travel times using scraped data from the HERE API, averaging reported durations for both public transit and driving. The resulting raw measure of distance is expressed in seconds. We normalize distance within the same census tract to two minutes (120 seconds) and add this constant to all cross-tract distances to ensure strictly positive values.

In addition, we construct a measure of *effective distance* that is more appropriate from a structural perspective—one that combines travel time, local income, and the cost of consumption trips. Using transaction-level data, we first estimate the elasticity of per-trip spending with respect to travel time by regressing

$$\log(x_{ni}) = \alpha_n + \delta_i + \tau \log(\text{distance}_{ni}) + \varepsilon_{ni},$$

where x_{ni} denotes average spending per trip from origin n to destination i, and distance_{ni} is the average travel time in hours between the two tracts. The fitted values $\hat{x}_{ni}$ capture expected per-trip spending conditional on travel time, origin, and destination effects.

We then compute a structural measure of distance as

$$\text{ldist}_{ni}^{\text{struct}} = \log\left(1 + \frac{y_n \cdot \text{distance}_{ni}}{\hat{x}_{ni}}\right),$$

where y_n is the average hourly income at the origin, computed as mean monthly income divided by 160 hours (corresponding to 8 hours per day and 20 working days per month). This measure reflects the time cost of travel in units of forgone consumption, linking physical distance and opportunity cost in a manner consistent with gravity and commuting models.

The resulting $\text{ldist}_{ni}^{\text{struct}}$ replaces simple log distance in the network matrices used in subsequent regressions, ensuring that the model captures both spatial frictions implied by geography and economic frictions implied by income and spending behavior.

3.1.6 Final Panel: Combining and Scaling to Administrative Aggregates

Finally, we describe how we integrate the various datasets into a unified panel. Our starting point is the representative tax-based statistics on gross and net income described in subsection 3.1.2. A similarly

representative and granular expenditure data from administrative sources is not available. We thus anchor our panel around the well-measured annual spatial income distribution and use our CaixaBank data for spending and commuting shares, as well as for monthly variation.

The procedure is straightforward. We begin with our CaixaBank monthly income panel described in Section 3.1.2, and scale income to match totals using census population data and official annual tax records. This adjustment is minimal given how highly correlated the averages in the two income sources are and mostly just involves scaling average income to totals using total population from the census. We now have a measure of local ("national") income. We then allocate this local income to consumption across the city using our spending shares from the spending data described in subsection 3.1.1. In this step, we implicitly make the assumptions that all income is spent on the subset of goods covered by our sample. We do, however, have information on spending *outside* our spatial network, including spending in the bigger metropolitan area of Barcelona but outside the city, as well as spending outside Barcelona as long as it is in the CaixaBank network. This step yields a spatial distribution of resident spending.

We next measure total local spending, which we assume comes from two sources: residents and tourists. Or, in national-accounting terms: domestic (own census-tract) consumption, resident exports to other census tracts, and exports to rest of the world. Resident spending at each consumption location is calculated in the previous step. We next calculate a representative measure of tourist spending. Using total income in the city of Barcelona as a proxy of city-wide *domestic consumption*, we make an adjustment so the aggregate share of tourism in GDP matches the Barcelona official number of 13.5% per year. Of this total per year, we allocate tourist spending to a location and country of origin using the expenditure shares by country of origin described in Subsection 3.1.1, generating a cross-sectional profile of tourist spending that reflects observed patterns in CaixaBank's data.

In Appendix G, we include summary statistics for the final linked panel, as well as an illustration of the full spatial structure of local income and spending flows in G.1.1. Section 3.2 discusses three stylized facts that emerge from this dataset and form the basis for our empirical strategy.

3.2 Tourism in the City: Three Stylized Facts

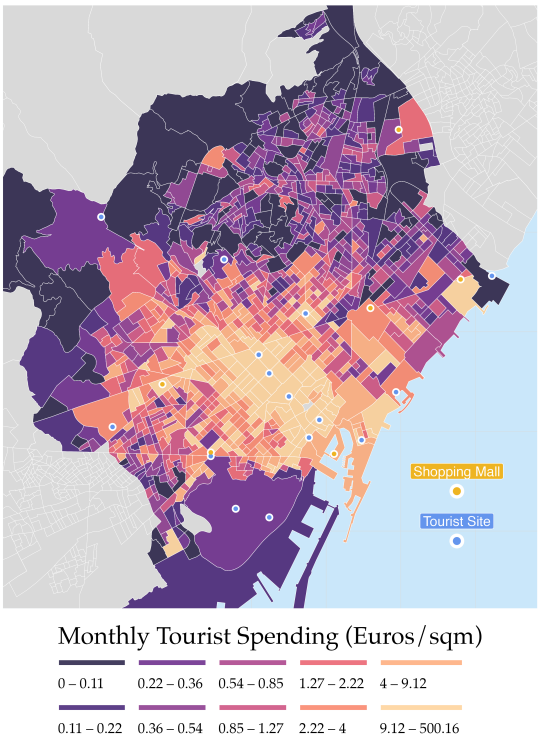
In this section, we use our data sources to document three stylized facts: (1) tourist activity varies across space and time, (2) locals' consumption and income exhibit strong spatial patterns localized around their place of residence, and (3) tourism appears to crowd out local consumption but increase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

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

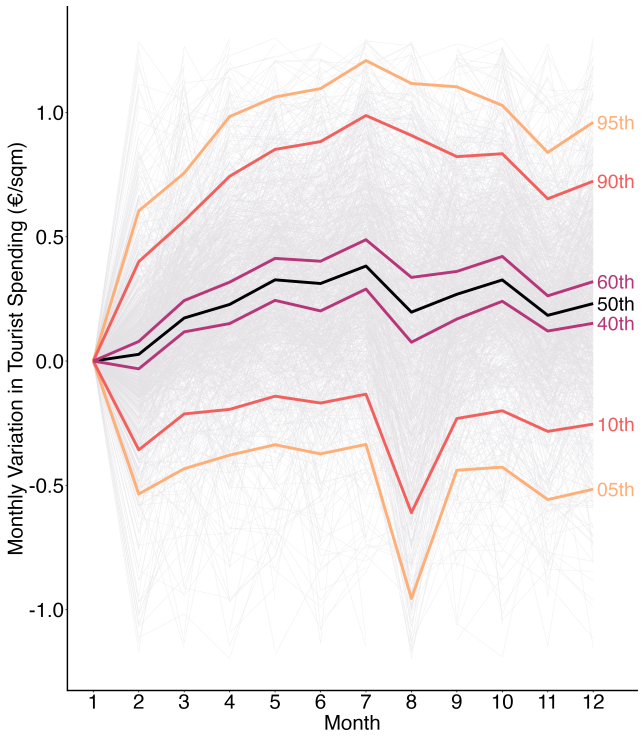
1(a) Tourist spending is spatial: the size of total tourist spending varies substantially across space. Figure 1 shows the intensity of tourist expenditures across individual locations for a given year. The intensity is normalized by the area of the underlying tile to account for heterogeneity in the size of individual census blocks. For convenience, we also show–with blue dots–the location of 15 popular tourist sites in Barcelona. Not surprisingly the expenditure of tourists is closely correlated with the location of main tourist attractions. 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 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 and along the coast.

Figure 1. Tourists spend disproportionately more in the city center and in summer

(a) Tourists spend more in the city center



(b) Tourist spending peaks in July



Notes: **(Panel a)** This figure shows the average monthly expenditure (normalized per square meter) in euros by tourists throughout the city of Barcelona. **(Panel b)** Log changes in monthly tourist spending relative to January, across different locations in the city and quantiles of the distribution. Each light gray line represents the time path for each location in the sample. The bold colored lines indicate selected quantiles of the cross-sectional distribution in each month.

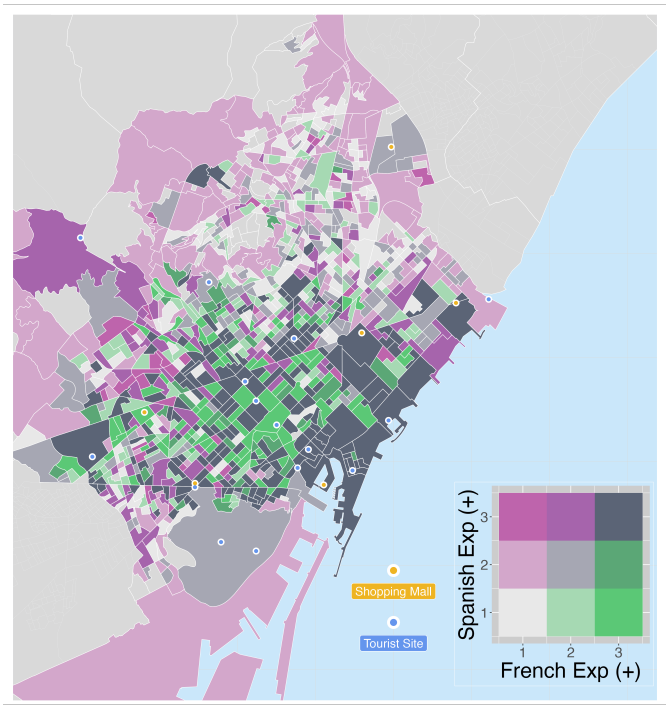
1(b) Tourism is seasonal: the incidence of tourist spending varies both over time and across locations within the city, as can be seen in Figure 1. Tourist spending rises sharply from January to the summer months, peaking in July, with marked seasonal booms in the most tourist-heavy areas. This is reflected in the upper quantiles of the distribution, which display much larger relative gains—highlighting the strong seasonal and spatial concentration of tourism. Starting around the 60th percentile and below, we observe a noticeable dip in August, likely driven by local factors such as business closures and co-ordinated holidays, which are common in Spain, outside the tourist-intensive areas. Toward the end of the year, tourist spending declines but generally remains above January levels in most locations. The figure thus illustrates both the temporal dynamics and spatial heterogeneity of tourism within the city.

1(c) Spatial concentration and seasonality are heterogeneous across types of tourists: the patters of spatial concentration and seasonal variation are heterogenenous across types of tourists, as defined by their region of origin. In Figure 2, we map as a bivariate chloropleth the quantiles of Spanish and French 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 gray, while low expenditure locations are market in light gray. The inner city is popular with both domestic and foreign tourists. While Barceloneta 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. Appendix G.3.1 compares other groups of tourists.

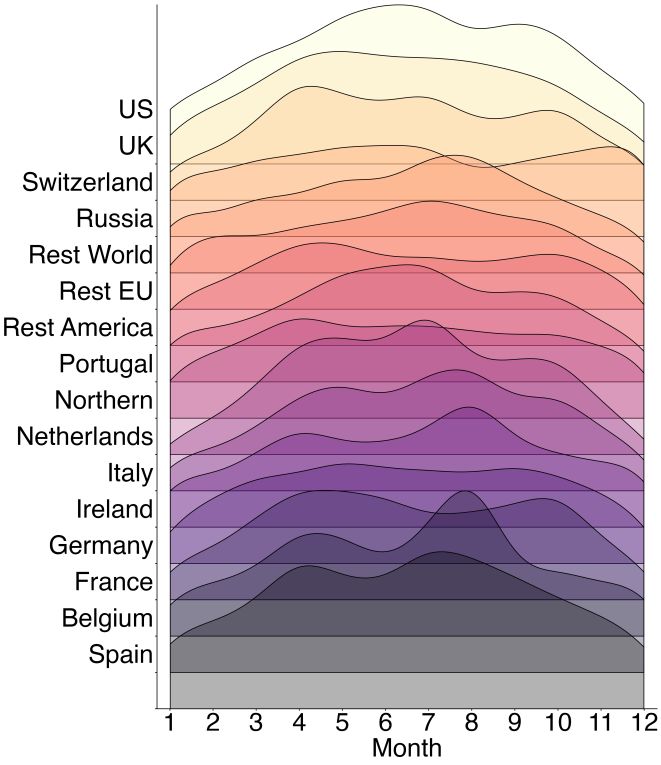
The spatial heterogeneity of different tourist groups interacts with their importance across time, in particular their seasonality. The panel (b) of Figure 2 summarizes the spending patterns of the most important groups of tourists. Note how the seasonal pattern changes: some have a market summer preference, 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.

Figure 2. Different tourists have different preferences within the city and year

(a) Tourists' Spending Across Space



(b) Tourists' Spending Across Seasons



Notes: **(Panel a)** This figure compares the percentage change in Spanish and French tourist expenditure between the tourist high season (July) 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 increase for both groups. **(Panel b)** This panel shows the monthly density of tourist spending by country of origin averaged over 2017–2019. Each curve represents the distribution of total spending across months for one nationality. The vertical axis indicates relative expenditure intensity within each group, with peaks corresponding to periods when that nationality accounts for a higher share of total tourist spending.

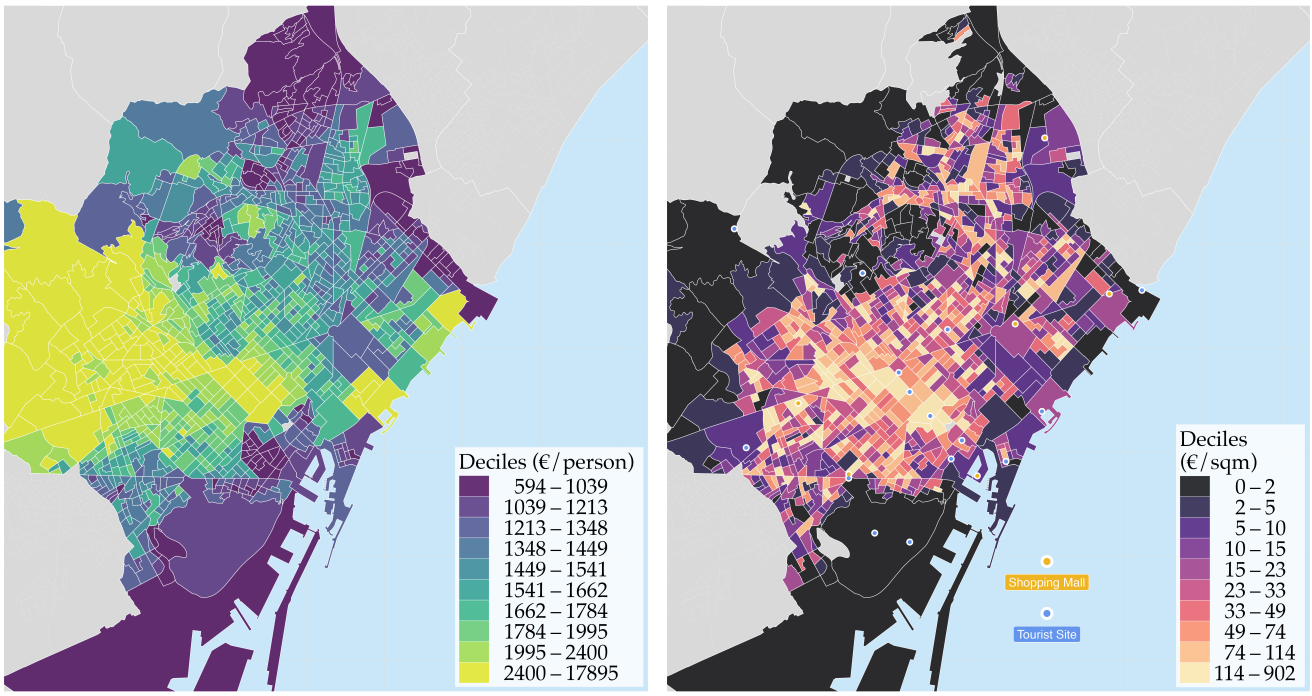
Fact 2: Locals' spending and income are spatial

2(ba). Locals' spending patterns across the city are shaped by residential proximity: While residents allocate spending across much of the city, expenditure is concentrated near their place of residence. In Figure 4 we plot the expenditure shares for the average resident of two sample locations. There are two clear takeaways from the figure: The first is that there is a substantial share of expenditures in close vicinity to the residential location. A la large share of spending takes place within 1 km of the home location. The second takeaway is that consumption is not purely local: while he spatial decay is strong, residents do spend in neighborhoods far from their own, with positive shares observed across nearly all areas and systematic concentration in central areas and major shopping centers. This combination of local concentration and broad reach reflects both spatial frictions and heterogeneous

consumption motives across the city.

Figure 3. Spending and Income Are Hyperlocal

(a) Residents' Income **(b) Residents' Spending**



Notes: This figure maps residents' monthly income and monthly spending by census tract, in euros. Panel (a) reports gross monthly labor income per person, computed as total labor income divided by the resident population in each tract. Panel (b) reports total monthly spending per square meter. Both variables are averaged across all months in 2017 and categorized in deciles.

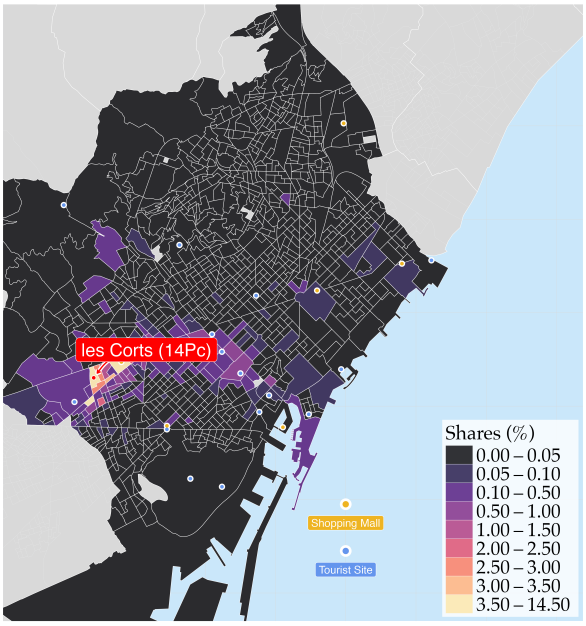
2(b) Local income is spatial: residents are more likely to earn nearby their home. As established in the urban economics literature, commuting involves time and monetary costs, which limit the willingness to travel long distances and lead individuals to locate near their workplace. This results in a spatially localized employment pattern. This can be observed in our data. In Figure 4, we illustrate commuting patterns from the lunchtime location data for residents of two sample locations. Most commutes are short-range, with the majority of trips directed to neighboring locations within the city.

Fact 3: Tourist spending affects locals

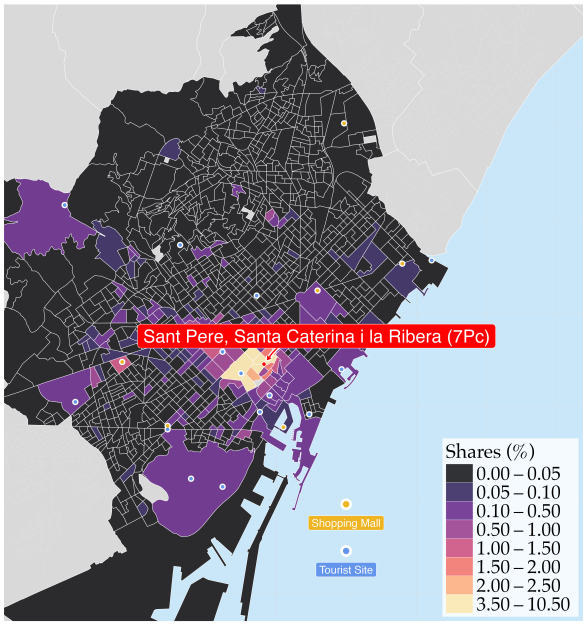
3(a) Tourist spending spatially crowds out local consumption: In Figure 5, we map the bivariate chloropleth simultaneously showing the change in tourist and local expenditure between February (the tourist low season) and July (the tourist high season) in 2019. It is clear from the map that particularly in the (foreign) tourist hotspots close to the beach as well as the lower part of the historical

Figure 4. Spending and Income Are Hyperlocal

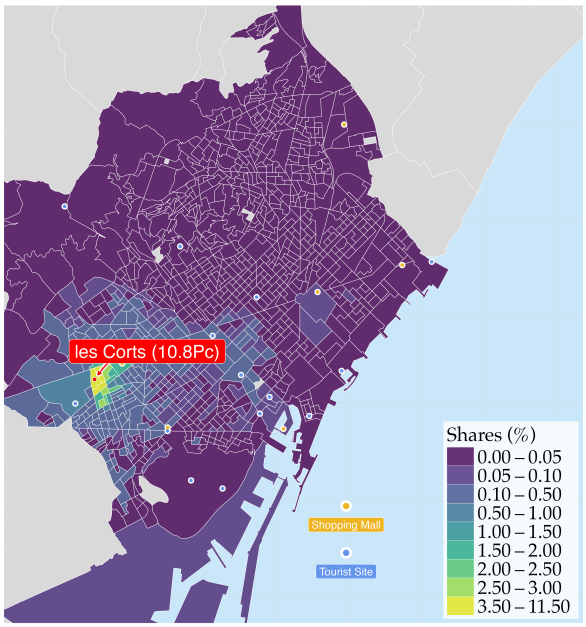
(a) Spending Shares - Example location 1



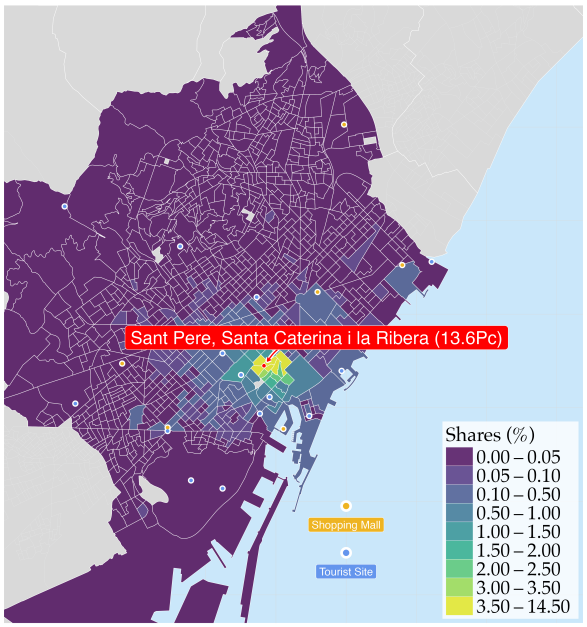
(b) Spending Shares - Example location 2



(c) Commuting Shares - Example location 1



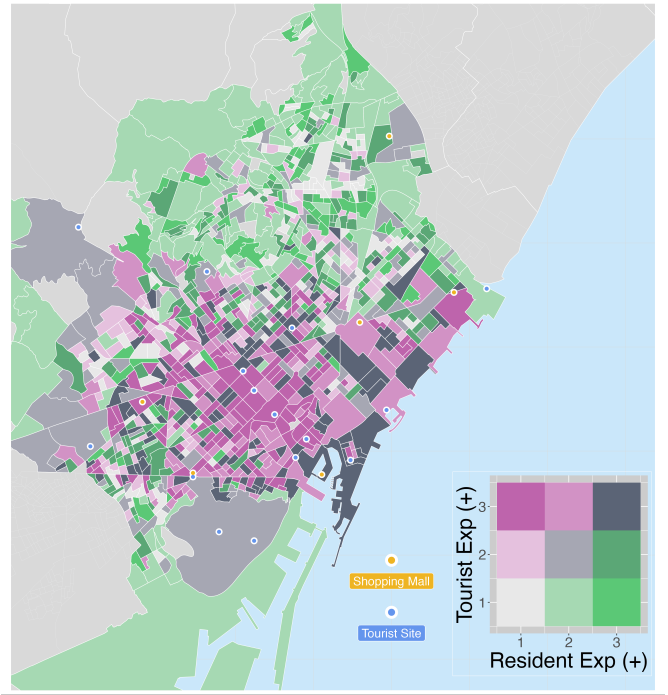
(d) Commuting Shares - Example location 2



Notes: This figure maps spending shares across the city for residents of a given location. Each panel maps the distribution of spending and commuting shares across the city for residents of two example locations: Les Corts (Panels a and c) and Sant Pere, Santa Caterina i la Ribera (Panels b and d). Lighter colors indicate tracts where a larger share of total spending or commuting by residents of the selected location occurs. Darker colors indicate smaller shares. Shares are expressed as percentages of total spending or commuting flows originating from each example location. Spending and commuting shares are described in details in Sections 3.1.1 and 3.1.3.

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.

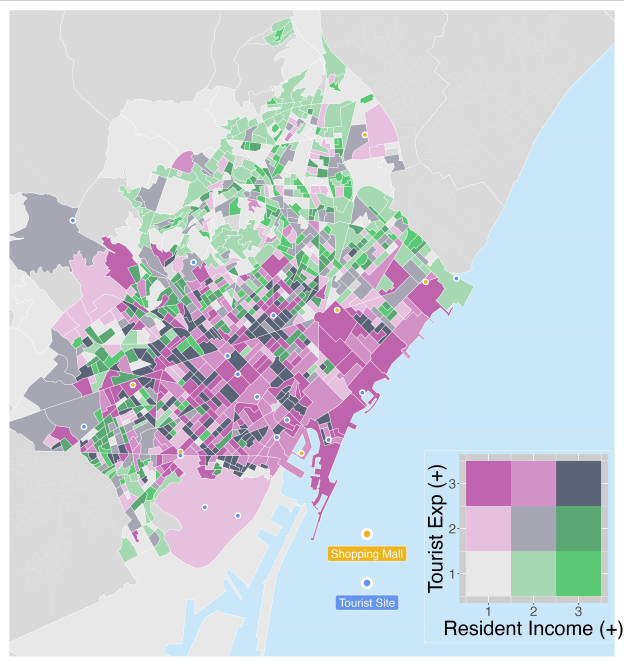
Figure 5. Tourist and Resident Spending Changes Are Negatively Correlated in a Large Part of the City



Notes: This figure maps the joint spatial distribution of changes in resident and tourist spending between the baseline month (January) and the summer peak (July). Each census tract is colored according to its position in the joint distribution of spending changes: dark (light) gray areas indicate joint increases (decreases) in both resident and tourist spending, purple areas correspond to increases in tourist spending accompanied by declines in resident spending, and green areas show the opposite pattern.

3(b) Tourist spending increases total income in a large part of the city: In Figure 6 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

Figure 6. 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 a 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 access and income do not comove.

local incomes (by increasing total spending).

We now turn to the empirical strategy used to assess the welfare impact of tourism across the city. The nonparametric framework introduced above provides the key insight for identification in an urban network: tourism shocks can generate distinct destination price effects and origin income effects, and both can propagate across space. Motivated by Fact #2, the strategy allows for rich heterogeneity in where residents live, work, and shop. Motivated by Fact #3, it separately identifies price and income responses and traces their diffusion through the city.

4 Tourist Shocks And Welfare

We now estimate the empirical objects defined in Section 2 and implement them in the data. Leveraging rich bilateral expenditure flows between census tracts in Barcelona, we recover reduced-form destination price effects, resident income effects, and tract-level real-income incidence. The section proceeds in four steps: first, we describe our empirical strategy built for urban networks; second, we estimate the bilateral expenditure responses and their propagated components; third, we inspect identification;

fourth, we build the price, income, and welfare objects.

4.1 Empirical Strategy and Causal Inference in Urban Networks

We aim to exploit the bilateral nature of our residents' income and spending panel. In an ideal world, we would estimate wage and price elasticities using a tract-level tourism shock $\{z_{it}\}_{i=1,\ldots,N}$ that is both exogenous and exclusively local. If we observed shocks that were orthogonal to contemporaneous local spending conditions and did not propagate through the urban network, the bilateral expenditure response could be summarized by

$$\log E_{nit} = \beta_1 z_{it} + \beta_2 z_{nt} + \gamma_{ni} + \delta_t + \varepsilon_{nit},$$

where β_1 captures the direct effect of tourism exposure on resident spending at destination (shopping) locations and β_2 captures the direct income or budget effect at residential location. In the data, however, neither condition is automatically satisfied.

The first problem is endogeneity of tourism exposure. A regression of resident spending on raw tourist inflows would not isolate the causal effect of tourism if the most exposed tracts are also changing along other dimensions that matter for local expenditure. One source of bias is omitted local demand shocks: tourist inflows may be higher in places that are simultaneously experiencing improvements that can raise resident spending directly—amenities and infrastructure, for example. A second source is endogenous spatial sorting and specialization: tracts that are historically more attractive to visitors may also attract firms, restaurants, and transport services that make them systematically different from less exposed tracts even absent a contemporaneous tourism shock. A third source is predictability in tourist demand itself: aggregate tourist inflows from a given origin country contain regular seasonal patterns and trends that are not necessarily unexpected or exogenous from the point of view of residents and firms.

To address these endogeneity issues, we construct a shift-share instrument for tourism exposure. We make use of our monthly panel of expenditure flows between origin tracts n and destination tracts i over the period 2017–2019. For each destination i, we construct a tourism shifter by interacting the pre-determined spending shares of tourist groups by country of origin across tracts with time-varying aggregate inflows of tourists from these groups:

$$\text{IV}_{it} = \sum_{g} s_{ig,2017} \times Z_{gt},$$

where $s_{ig,2017}$ is the share of group g's total spending allocated to tract i in the base year 2017, and Z_{gt} is the aggregate inflow of tourists from group g at time t during 2018–2019. The key identifying assumption is that the initial cross-tract allocation of tourist spending reflects persistent geographical and amenity characteristics rather than responses to contemporaneous local demand shocks, so that the shares $s_{ig,2017}$ are orthogonal to time-varying unobserved determinants of local economic activity. Variation in the instrument therefore comes from differential exposure to common tourist-origin shocks through the pre-existing spatial pattern of tourism. Analogously, for each origin tract n, we construct an exposure measure IV_{nt} using the same shift-share logic. In Subsection 4.2, origin n variation is used to recover resident income responses while destination i variation will serve to revover prices.

A remaining endogeneity concern discussed above is that even the shift-share instrument may still partly capture forecastable variation in tourist demand, rather than genuinely unexpected shocks. To separate these two components, we write $\text{IV}_{it} = \widehat{IV}_{it} + \widetilde{IV}_{it}$, where $\widehat{IV}_{it}$ is the predicted or forecastable component and $\widetilde{IV}_{it}$ is the residual unexpected component. Following Borusyak, Hull and Jaravel (2022), we control for $\log \widehat{IV}_{it}$ and $\log \widehat{IV}_{nt}$ directly in the regression, so that identification comes from $\widetilde{IV}$ rather than from systematic seasonal or trend variation in tourism demand. We will discuss this approach and inspect identification in Subsection 4.3.

Network propagation and SUTVA. The second departure from a simple OLS identification is that tourism shocks do not remain local. To address this network propagation problem, we allow shocks to tourism demand to diffuse throughout the urban economy rather than remaining confined to the initially exposed tracts. We construct network-based general equilibrium (GE) exposure measures using matrices of bilateral spending shares S and income-link or commuting shares C. Matrix S captures the fraction of spending by residents of origin n in destination i, while matrix C captures the share of resident income in origin n that is earned in destination i through commuting links.

To formalize diffusion using our nonparametric model of Section 2, we define the one-round matrix

$$\mathbf{M} \equiv (D^y)^{-1} S' D^v C.$$

The associated percolation operator is

$$\mathbf{\Pi} \equiv (I - \mathbf{M})^{-1} \mathbf{D}_T = \sum_{h \geq 0} \mathbf{M}^h \mathbf{D}_T.$$

For any destination-level variable X, the h-th order destination exposure is

$$\text{ge}_h^i(X) = \mathbf{M}^h \mathbf{D}_T X.$$

The corresponding origin (residence) exposure is obtained by mapping through commuting links:

$$\text{ge}_h^n(X) = C\mathbf{M}^h \mathbf{D}_T X.$$

Here D^y scales destination sales, D^v scales resident income, C maps destination-side income back to resident origins, S' maps origin spending back to destinations through shopping flows, and $\mathbf{D}_T$ is the tourist-intensity diagonal.

For the empirical analysis, we use cumulative propagation up to order k:

$$\text{GE}_k^i(X) = \sum_{h=0}^{k} \text{ge}_h^i(X), \qquad \text{GE}_k^n(X) = \sum_{h=0}^{k} \text{ge}_h^n(X).$$

We construct these GE terms for orders $k = 0, 1, 2, 3, 4, \infty$, allowing us to track the progressive diffusion of tourism shocks throughout the city. This propagation problem also has a causal-inference implication. A naive shift-share exposure index would summarize treatment at tract n with a single weighted average of shocks, implicitly imposing a Stable Unit Treatment Value Assumption (SUTVA). In the urban network, SUTVA fails because shocks do not remain local: once they propagate through commuting and spending links, outcomes depend on the entire shock vector rather than only on own exposure. In the linear notation of Section 2, the correct model-consistent exposure is the percolated object

$$\widetilde{z}_n = \mathbf{s}_n \mathbf{\Pi} \mathbf{g},$$

not the raw shift-share itself. Our empirical design addresses this issue by including both direct exposure terms and their propagated counterparts, summarized by the cumulative operators $\text{GE}_k^i(\cdot)$ and $\text{GE}_k^n(\cdot)$, so that the empirical specification allows shocks to diffuse through the urban network rather than forcing them to remain local. The next subsection 4.2 describes the final specification. Note that we apply the same predictability correction to both the direct and propagated exposure measures, so that identification comes from unexpected network propagation rather than from predictable diffusion of tourism demand through the city.

4.2 Empirical Estimates

Building on the constructed exposure measures, we estimate the effect of tourism shocks on local economic activity by regressing bilateral expenditures on both direct and propagated exposure terms. The key empirical separation is between destination-side crowd-out and origin-side budget expansion. Destination-family coefficients identify the price or cost-of-living channel. Origin-family coefficients identify the resident income or budget channel. The baseline regression specification is:

$$\begin{aligned}\log E_{nit} \;=\; & \beta_1 \log \widetilde{IV}_{it} + \beta_2 \log \widetilde{IV}_{nt} + \\ & \beta_{1k} \mathrm{GE}^i_k(\log \widetilde{IV}_{it}) + \beta_{2k} \mathrm{GE}^n_k(\log \widetilde{IV}_{nt}) + \\ & \boldsymbol{\alpha}_i' \widehat{\mathbf{v}}_{it} + \boldsymbol{\alpha}_n' \widehat{\mathbf{v}}_{nt} + \gamma_{ni} + \delta_t + \varepsilon_{nit},\end{aligned}$$

where $\log E_{nit}$ denotes the logarithm of total spending from residents of origin n in destination i at month t. The vectors $\widehat{\mathbf{v}}_{it}$ and $\widehat{\mathbf{v}}_{nt}$ collect all predicted-component controls—direct and network-propagated—at the destination and origin respectively:

$$\widehat{\mathbf{v}}_{it} \equiv \big(\log \widehat{IV}_{it},\ \mathrm{GE}^i_0(\log \widehat{IV}_{it}),\ \ldots,\ \mathrm{GE}^i_\infty(\log \widehat{IV}_{it})\big)',$$

and analogously $\widehat{\mathbf{v}}_{nt} \equiv \big(\log \widehat{IV}_{nt},\ \mathrm{GE}^n_0(\log \widehat{IV}_{nt}),\ \ldots,\ \mathrm{GE}^n_\infty(\log \widehat{IV}_{nt})\big)'$, with corresponding coefficient vectors $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_n$. That is, each vector stacks the forecastable component of the direct instrument together with its GE-propagated counterparts at every order. Controlling for $\widehat{\mathbf{v}}_{it}$ and $\widehat{\mathbf{v}}_{nt}$ therefore ensures that β_{1k} and β_{2k} are identified from unexpected network propagation only. The specification includes a full set of origin-destination fixed effects γ_{ni}, controlling for all time-invariant bilateral heterogeneity, including persistent economic ties or geographical proximity. Month-year fixed effects δ_t absorb aggregate city-wide shocks and seasonal variation. Standard errors are clustered two-way at the origin-month and destination-month levels to correct for potential spatial and temporal correlations in the error structure.

Our empirical strategy fully exploits the bilateral structure of the data, and explicitly models general equilibrium spillovers through a city's spatial network. This analysis moves beyond simple aggregation and enables a detailed tracing of shock propagation across space, providing a richer characterization of the urban impacts of tourism shocks than conventional shift-share designs.

Results. Table 1 presents the baseline bilateral expenditure regressions. Across all columns, the dependent variable is $\log E_{nit}$. Column (1) reports the raw OLS relationship using direct tourist spending at origin and destination; these coefficients are descriptive and reflect shared local demand conditions.

Column (2) replaces spending with the unexpected shift-share shocks $\widetilde{IV}$, controlling for the predicted component $\widehat{IV}$. The destination coefficient is large and negative ($\beta_1 = -0.1538$), indicating that an unexpected increase in tourism at destination i reduces bilateral spending by residents. The direct origin coefficient is economically small and statistically indistinguishable from zero ($\beta_2 = 0.0065$).

Column (3) introduces the GE propagation terms. The destination direct effect remains negative (-0.1148), and destination GE propagation is also negative ($\beta_{1\infty} = -0.3236$), indicating that tourism shocks reduce local spending not only in directly exposed tracts but also in destinations connected through the network. At the origin, direct exposure remains close to zero, but propagated exposure is positive and significant ($\beta_{2\infty} = 0.2287$): residents whose neighborhoods are indirectly reached by tourism shocks through the commuting and spending network increase their spending. Column (4), which uses structural GE objects, preserves the same qualitative asymmetry between destination and origin channels.

Shock diffusion. To trace propagation order by order, Table A.4 introduces GE terms from order 0 to the full recursion. The direct destination coefficient remains negative throughout but becomes less negative as higher-order channels are added (from -0.1384 in the 0th-order column to -0.1071 in the full column). At the same time, destination GE coefficients become increasingly negative across orders and are strongly negative in the full specification ($\beta_{1\infty} = -0.4489$).

On the origin side, direct exposure remains near zero, while GE propagation is positive and statistically significant at each order (from 0.5096 at order 0 to 0.2805 in the full specification). Figure A.5 summarizes these patterns.

Overall, these results confirm that tourism shocks generate substantial network spillovers that propagate beyond directly exposed locations, affecting economic activity throughout the city. The spatial structure of these spillovers is highly asymmetric: positive at the origin and negative at the destination, reflecting a combination of income and spending effects on resident consumption.

Table 1. Main Specifications

Dependent Variable:	**Residents Spending Flows** $\log E^R_{nit}$			
Model	OLS	Direct Effect	GE Linearized	GE Structural
Variables				
Origin Spending-Income Effects (n - residence)				
$\log E^T_{nt}$	0.0021			
	(0.0014)			
$\log iv_{nt}$		0.0065	0.0030	-0.0018
		(0.0062)	(0.0062)	(0.0062)
$\log GE^{IV,\infty}_{nt}$			0.2287**	0.0092
			(0.0951)	(0.0936)
$\log \hat{v}^{structural}_{nt}$				0.8166***
				(0.1373)
Destination Spending Effects (i - shopping)				
$\log E^T_{it}$	0.1060***			
	(0.0050)			
$\log iv_{it}$		-0.1538***	-0.1148***	-0.1060***
		(0.0257)	(0.0254)	(0.0249)
$\log GE^{IV,\infty}_{it}$			-0.3236***	-0.5797***
			(0.1253)	(0.1681)
$\log \hat{p}^{structural}_{it}$				0.2926**
				(0.1480)
Fixed Effects				
$n \times i$	Yes	Yes	Yes	Yes
t	Yes	Yes	Yes	Yes
Fit Statistics				
Observations	3,517,327	3,515,125	3,515,125	3,517,125
R^2	0.64984	0.64949	0.64950	0.64951
Within R^2	0.00117	0.00010	0.00012	0.00015

Clustered ($n \times t$ & $i \times t$) standard errors in parentheses.
Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: $\log E^R_{nit}$ denotes bilateral expenditure of residents between origin n and destination i at time t. $\log E^T_{it}$ and $\log E^T_{nt}$ capture tourist expenditure at destination and origin, respectively. $\log \widetilde{IV}_{it}$ and $\log \widetilde{IV}_{nt}$ measure the unexpected component of direct tourism exposure at destination and origin, based on shift–share instruments constructed from pre-determined spending shares and time-varying inflows of tourist groups. All regressions control for the predicted-component vectors $\widehat{\mathbf{v}}_{it}$ and $\widehat{\mathbf{v}}_{nt}$—stacking $\log \widehat{IV}$ and its GE-propagated counterparts at all orders—following Borusyak, Hull, and Jaravel (2022). $\text{GE}^i_k(\log \widetilde{IV}_{it})$ and $\text{GE}^n_k(\log \widetilde{IV}_{nt})$ denote kth-order general-equilibrium exposure terms obtained through network propagation of the unexpected shock (or of tourist spending in the OLS case) across commuting and spending networks. All regressions include origin–destination and various combinations of month and origin–month fixed effects. Standard errors are two-way clustered by origin–month and destination–month. Data are based on bilateral card-transaction records between census tracts in Barcelona, 2018–2019.

Housing prices. In addition to expenditures and incomes, we extend our analysis to housing market outcomes, specifically examining the effects of tourism shocks on local housing prices. Housing prices provide a longer-run indicator of local desirability and capitalized welfare, capturing how tourism-driven economic shocks may influence residential location values across the city.

We adapt the empirical framework developed for expenditures and incomes by focusing on the logarithm of average housing prices, $\log p_{nt}$, in origin census tract n at month t as the dependent variable.

The housing regression specification parallels that for incomes:

$$\log p_{nt} = \theta_2 \log \widetilde{IV}_{nt} + \theta_{2k}\, \text{GE}_k^n(\log \widetilde{IV}_{nt}) + \boldsymbol{\alpha}_n' \widehat{\mathbf{v}}_{nt} + \mu_{\tilde{n}t} + \nu_{nt}, \tag{1}$$

where $\text{GE}_k^n(\log \widetilde{IV}_{nt})$ denotes the k-th order GE exposure to the unexpected tourism shock at the origin tract, and $\widehat{\mathbf{v}}_{nt}$ is the vector of predicted-component controls defined above (stacking $\log \widehat{IV}_{nt}$ and its GE-propagated counterparts). As before, $\mu_{\tilde{n}t}$ are origin fixed effects at the district level[5], absorbing month-specific heterogeneity in housing market fundamentals but allowing for some vicinity spillovers. The error term ν_{nt} captures unobserved residual variation in housing prices.

The specification allows us to assess both the direct and network-propagated effects of tourism shocks on housing prices. Higher-order general equilibrium terms model the progressive diffusion of shocks through commuting and spending networks, reflecting how local economic spillovers may be capitalized into residential real estate markets.

Sectoral Heterogeneity. Figure 7 displays the estimated coefficients of the key instrumented exposure variables across different category groups, along with their 95% confidence intervals. The four coefficients correspond to the direct exposure at destination tracts ($\log \widetilde{IV}_{it}$), the direct exposure at origin tracts ($\log \widetilde{IV}_{nt}$), as well as the full GE propagated exposures at destination and origin tracts ($\text{GE}_\infty^i(\log \widetilde{IV}_{it})$ and $\text{GE}_\infty^n(\log \widetilde{IV}_{nt})$, respectively).

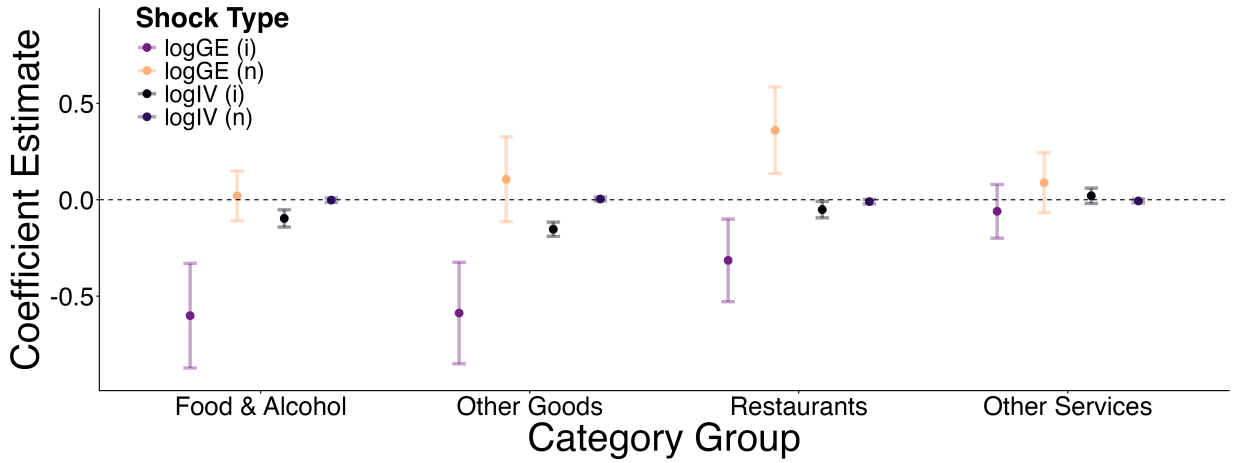
Across most categories, the coefficient estimates for direct exposure at the destination ($\log \widetilde{IV}_{it}$) are negative, indicating that greater exposure to tourism shocks at destinations tends to reduce bilateral expenditures. In contrast, the coefficients for direct exposure at the origin ($\log \widetilde{IV}_{nt}$) are generally close to zero and not statistically distinguishable from zero across most categories. For the GE terms, $\text{GE}_\infty^n(\log \widetilde{IV}_{nt})$ coefficients tend to be positive, reflecting that higher network-based exposure at the origin tract is associated with increased expenditure. Meanwhile, $\text{GE}_\infty^i(\log \widetilde{IV}_{it})$ coefficients are negative across several categories, suggesting that indirect propagation of tourism shocks through the network at the destination tends to reduce local expenditures.

The pattern of coefficients highlights considerable heterogeneity across sectors. In particular, sectors like Food, Hospitality, and Personal Care display larger and more significant negative destination-side coefficients, consistent with greater sensitivity to local tourism activity. Conversely, sectors such as Communications and Education display more variation and wider confidence intervals, reflecting noisier estimates and potentially more complex relationships with tourism-driven demand shifts.

Overall, the figure supports the view that tourism shocks affect local economic activity not only

[5]There are 10 districts in Barcelona city.

Figure 7. Destination Spending Effects by Sector Group



Notes: This figure plots the estimated coefficients and 95% confidence intervals for the key exposure variables across different consumption category groups. **Other Goods**: Clothing and footwear (COICOP 3); furnishings and household equipment (COICOP 5); vehicles and transport equipment (COICOP 7.1 and 7.2); recreational goods (COICOP 9.1 and 9.2). **Other Services**: utilities (COICOP 4); health services (COICOP 6); education services (COICOP 10); information and communication services (COICOP 8); recreational services (COICOP 9.4 and 9.5); insurance and financial services (COICOP 12). The variables shown are the direct destination exposure ($\log \text{IV}_{it}$), direct origin exposure ($\log \text{IV}_{nt}$), fourth-order general equilibrium exposure at destination ($\log \text{GE}_{it}^{IV,\infty}$), and fourth-order general equilibrium exposure at origin ($\log \text{GE}_{nt}^{IV,\infty}$). Coefficients are obtained from separate regressions for each category group, where the dependent variable is the logarithm of bilateral expenditure, and the specification includes origin-destination and year-month fixed effects. Standard errors are clustered at the origin-month and destination-year-month levels.

through direct effects at specific locations but also through spatial propagation across the urban network, with important sectoral differences in the magnitude and significance of these effects.

4.3 Inspecting Identification

We now inspect the identifying variation underlying the estimates reported above. For clarity, we base our discussion on the simplified bilateral specification using only direct shocks:

$$\log E_{nit}^{R} = \beta_1 \log \widetilde{IV}_{it} + \beta_2 \log \widetilde{IV}_{nt} + \alpha_1 \log \widehat{IV}_{it} + \alpha_2 \log \widehat{IV}_{nt} + \gamma_{ni} + \delta_t + \varepsilon_{nit},$$

where identification comes from within-pair, over-time variation in the unexpected component $\widetilde{IV}$, after removing common month shocks and the predicted trends $\widehat{IV}$.

First stage and variance decomposition. The first stage is strong. In destination-time regressions, the coefficient on $\log \widetilde{IV}_{it}$ is 0.829 without predicted-trend controls and 0.269 with those controls. The introduction of predicted-trend controls substantially tightens identifying variation: only 0.30% of the total variance in $\log \widetilde{IV}_{it}$ remains after residualizing by tract effects, month effects, and predicted trends

(0.0137 out of 4.5107 total variance units). In the bilateral specification, destination-side variation drives the main result: $\hat{\beta}_1 = -0.1538$, while $\hat{\beta}_2 = 0.0065$. In destination-only and origin-only designs, the corresponding coefficients are -0.1508 and -0.0061. A decomposition of the fitted identifying signal attributes 98.7% to the destination IV channel and 0.5% to the origin IV channel.

Rotemberg decomposition. We follow Goldsmith-Pinkham, Sorkin and Swift (2018) to decompose the identifying weight across origin countries. Table A.9 in the appendix reports the top-10 Rotemberg weights. The distribution is concentrated but not dominated by a single origin: the largest contributor, USA and Canada, accounts for 22.9% of the total weight, followed by Rest of the World (17.2%) and France (10.4%). The top 3 countries jointly account for 50.6% of identifying variation, the top 5 for 67.3%, and the top 10 for 88.3%.

Predicted vs. unpredicted variation. We run both the predictable trend component $\widehat{IV}_{it}$ and the residual $\widetilde{IV}_{it}$ in the same outcome equation. Three diagnostics stand out. First, the *coefficient comparison*: the estimated coefficient on the predictable component is $\hat{\alpha}_1^{\text{pred}} = +0.076$, while the coefficient on the residual is $\hat{\beta}_1^{\text{resid}} = -0.152$—opposite in sign. The divergence is informative: the predictable component captures secular co-movement between growing tourist areas and rising resident spending, while the residual identifies the causal displacement of residents by unexpected surges. If the predicted trend were merely absorbing collinear variation, the two coefficients would converge; their reversal instead validates that controlling for $\widehat{IV}_{it}$ does substantive work by removing a confound of the opposite sign. In a *variance decomposition*, we find that the majority of the instrument variance is accounted for by the fixed effects and the predictable trend, but paired the strong first-stage confirms that this identifying variation is nonetheless highly informative.

Balance tests. Figure A.7 in the appendix displays balance tests for the identifying component of the instrument. While 9 out of 10 baseline characteristics are significantly correlated with raw IV exposure—reflecting the well-known concentration of tourism in wealthier, denser central tracts—only 6 out of 10 remain significant for the residual after removing fixed effects and predicted trends. Crucially, the signs reverse for income, house prices, and rent: tracts with higher baseline values receive *less* of the unexpected shock, consistent with unexpected tourist surges being less geographically selective once systematic sorting and trend growth are removed.

Pre-trends. Panel placebo tests decompose the instrument into its predicted and residual components: pre-period income and rent are not significantly related to the residual component ($p = 0.69$ and

$p = 0.50$ respectively).

Further Robustness. Table A.7 reports robustness checks using weighted specifications, PPML, trimming, exclusion of August, and aggregation to AEB units. Three patterns emerge. First, the direct origin coefficient remains close to zero in all specifications. Second, GE propagation terms preserve their signs: destination GE effects are negative and origin GE effects are positive in all columns, though precision falls in PPML. Third, the direct destination IV effect remains negative in the baseline, trimmed, no-August, and AEB specifications, but is attenuated and less precise in some weighted and PPML runs. Overall, the spatial propagation pattern is robust even when the direct destination coefficient varies in precision across estimators.

4.4 Empirical Implementation: Price, Income, and Welfare Effects

As in Section 2, small tourism shocks affect welfare through destination-side price responses and origin-side income responses. In the data, the destination-family coefficients map tourism shocks into prices, while the origin-family coefficients map tourism shocks into resident income. For any shock vector z, the implied destination-side price response at block i and the implied origin-side income response at tract n satisfy

$$\widehat{\mathrm{d}\ln p_i} = \frac{\hat{\beta}_1 z_i + \sum_k \hat{\beta}_{1k}\,\mathrm{GE}_k^i(z)}{1-\hat{\sigma}}, \qquad \widehat{\mathrm{d}\ln y_n} = \hat{\beta}_2 z_n + \sum_k \hat{\beta}_{2k}\,\mathrm{GE}_k^n(z).$$

Resident cost of living is then the shopping-share-weighted average of destination price changes,

$$\widehat{\mathrm{d}\ln p_n} = \sum_i s_{ni}\,\widehat{\mathrm{d}\ln p_i},$$

so tract-level welfare satisfies

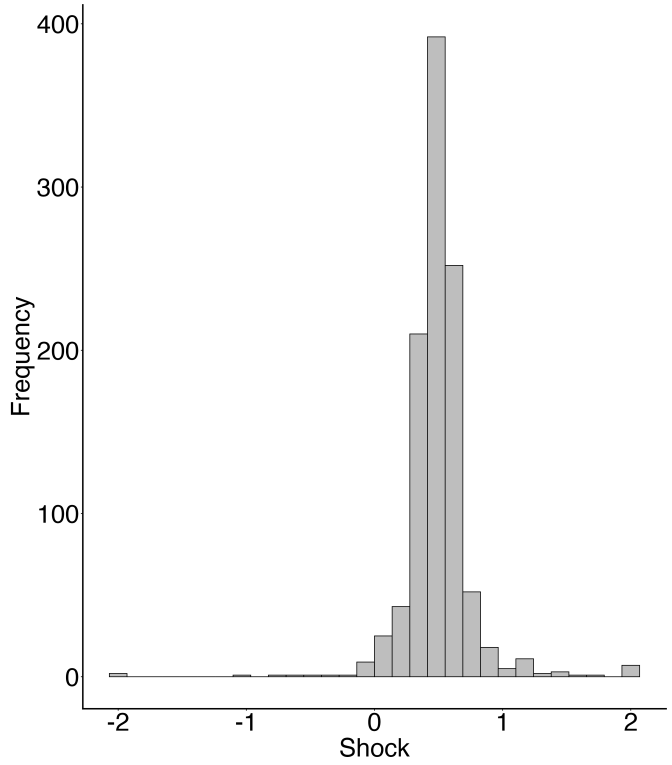
$$\widehat{\mathrm{d}\ln u_n} = \widehat{\mathrm{d}\ln y_n} - \widehat{\mathrm{d}\ln p_n}.$$

Tourism Shock. We define the tourism shock as the log change of our shift-share instrument between January and July 2018. Figure 8 summarizes the distribution of this shock, both cross-sectionally and spatially. In general, there is a larger increase in the tourist center, though some idiosyncratic large changes can be seen in more residential areas where tourist spending is very small in January.

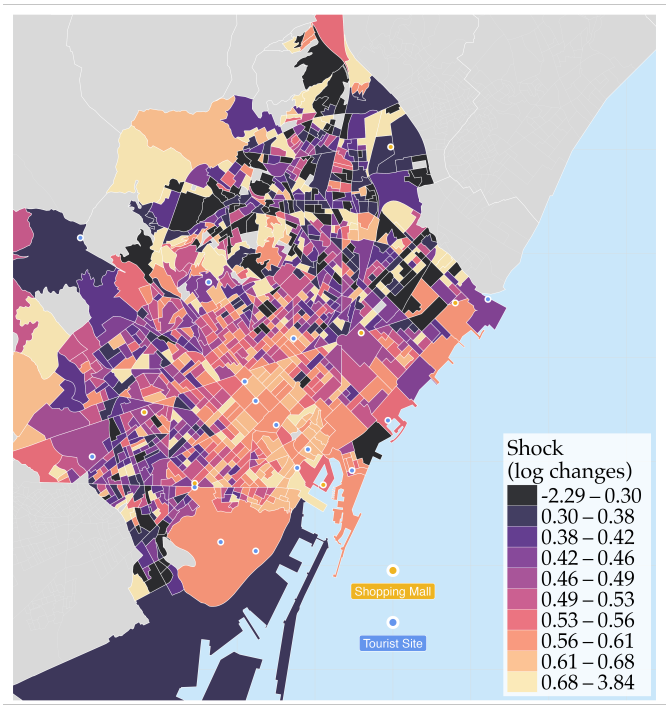
Interpretation of Price Effects. Figure 9 combines the estimates in the spending regressions with the propagation matrix and the shock to trace out the implied price changes. Under a CES preference spec-

Figure 8. Tourist Shock

(a) Cross-Sectional Distribution



(b) Spatial Distribution



Notes: We define the tourism shock as the log change of our shift-share instrument between January and July 2018. **(Panel a)** shows the cross-sectional distribution of the shock. We truncate changes between -2 and 2 into the first and last bins. **(Panel b)** maps this distribution into space.

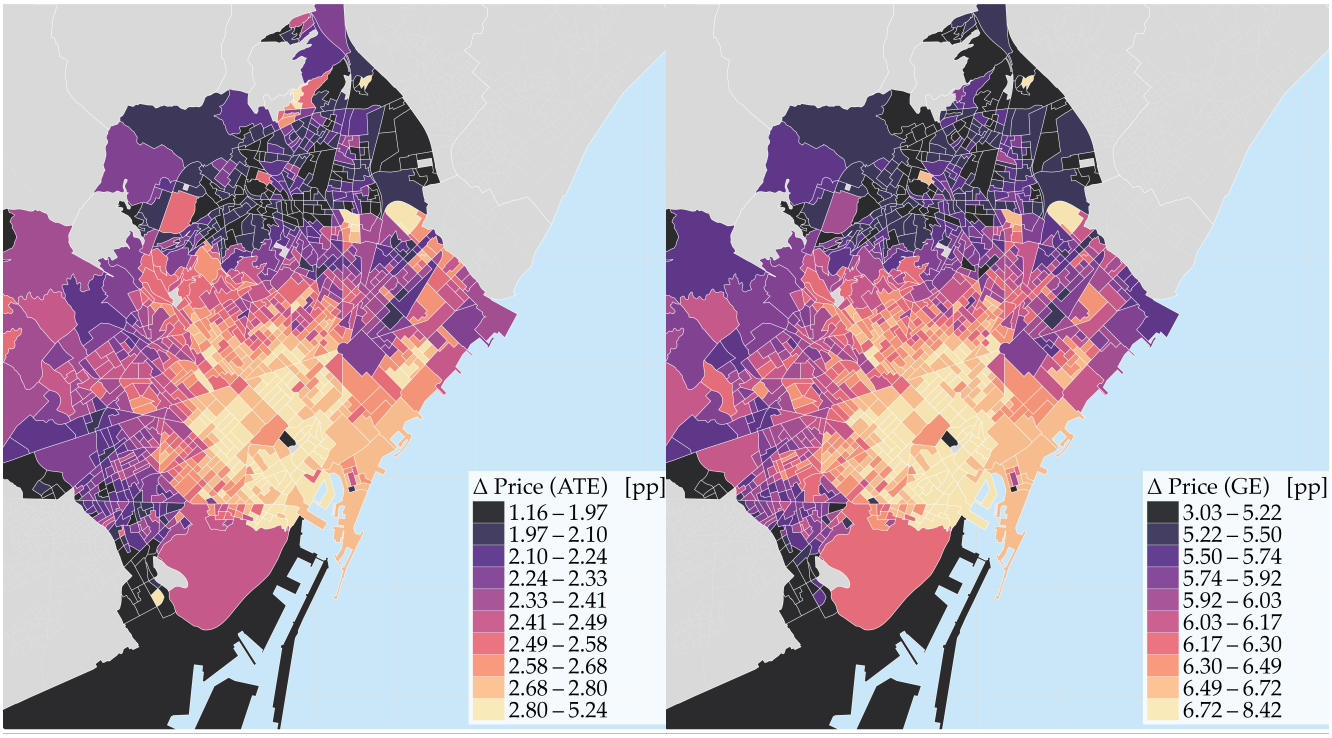
ification, we map spending elasticities into price elasticities as $\frac{\widehat{\beta}_1}{1-\sigma}$ for the direct (ATE) effects and $\frac{\widehat{\beta}_{1\infty}}{1-\sigma}$ for the full GE effects, where $\widehat{\beta}_1$ and $\widehat{\beta}_{1\infty}$ are the estimated coefficients on $\log \widetilde{IV}_{it}$ and $\text{GE}^i_\infty(\log \widetilde{IV}_{it})$ respectively, and σ is the CES trade elasticity from the gravity regression in Table A.2. We use the OLS estimate on the structural distance measure described in Section 3.1.5 as our reference elasticity. We then interact these elasticities with the propagation matrices and the shock summarized in Figure 8.

Overall, the GE specification generates stronger and more spatially diffuse price effects, reflecting the endogenous propagation of shocks through interconnected local markets. The first-order effect is an overall increase in prices, as shocks transmit through income and spending channels. The second-order effect is a widening of the distribution, with more pronounced variation across neighborhoods as feedback effects accumulate. Despite these amplifications, the relative ranking of areas—those experiencing the strongest or weakest price responses—remains largely consistent with the partial-equilibrium (ATE) case.

Interpretation of Income Effects. Figure 10 combines the estimated income elasticities with the spatial propagation matrix and the local tourism shock to trace out the implied income responses across

Figure 9. Price Effects from Tourist Shock

(a) Average Effects **(b)** Effects with GE Propagation



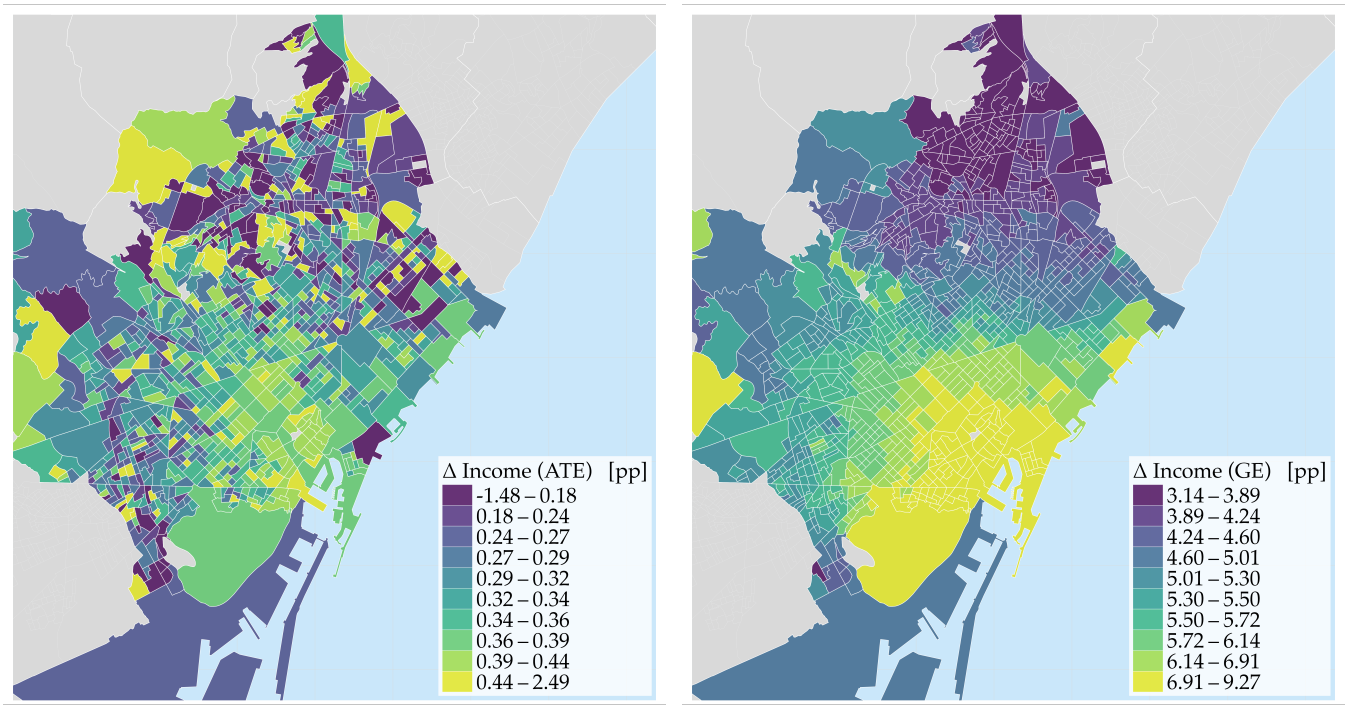
Notes: This figure reports the predicted spatial distribution of price effects induced by a tourism demand shock between the low and high seasons. The shock is defined as the log change in local tourism inflows between January and July 2018. Panel (a) displays the **average treatment effect (ATE)**, which captures the direct local effect of the demand shock on prices at each destination i, estimated from the partial-equilibrium elasticity of local prices with respect to local inflows β_1. Panel (b) reports the corresponding **general equilibrium (GE)** effects, which incorporate spatial spillovers through the full network of trade and commuting links. Price changes are expressed in percentage points. Light colors denote larger price increases.

neighborhoods. Overall, the GE specification produces stronger and more spatially diffuse income effects, reflecting feedback between labor income and local expenditure across interconnected neighborhoods.

As in the price analysis, these effects are derived by multiplying the estimated coefficients $\widehat{\beta}_2$ on $\log \widetilde{IV}_{nt}$ and $\widehat{\beta}_{2\infty}$ on $\text{GE}^n_\infty(\log \widetilde{IV}_{nt})$ by the observed shock, where the latter is further interacted with GE propagation through the network of production and commuting linkages. The resulting maps display the distribution of income changes across locations under partial equilibrium (ATE) and full GE propagation. Incorporating GE effects amplifies both the level and spatial dispersion of income responses, as local increases in demand raise earnings not only in directly affected areas but also in surrounding tracts connected through the commuting network. In this case, incorporating GE linkages has a qualitative effect on some tracts, which had negative income effects in the average case. In this case, the qualitative spatial pattern is different. This is because the propagation to income includes one further iteration from price to income, hence the smooth nature of the GE effects.

Figure 10. Income Effects

(a) Average Effects **(b)** Effects with GE Propagation



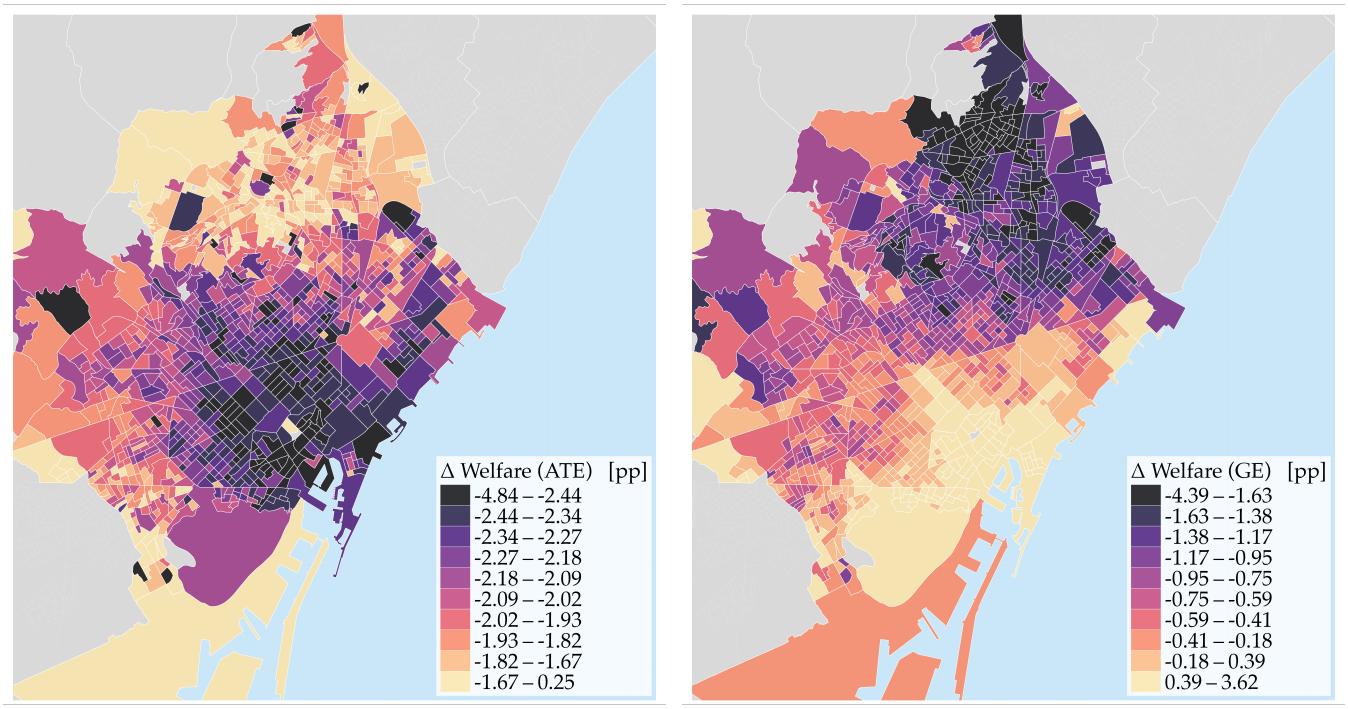
Notes: This figure reports the spatial distribution of predicted income effects induced by the tourism shock between January and July 2018. Panel (a) displays the average treatment effect (ATE), capturing direct local income responses. Panel (b) shows the general equilibrium (GE) effects, which incorporate propagation through the commuting and spending network. Light colors denote larger income gains.

Interpretation of Welfare Effects. 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 (90th percentile). Dissecting this net welfare change into price and income effects reveals that residents in the city center and those near tourist locations bear the largest price increases but also enjoy substantial income gains. In contrast, residents of peripheral neighborhoods suffer lower but still sizable price increases, with income gains varying across outer-city locations: some experience none while others receive moderate income benefits from tourism.

Figure 11 combines the estimated income and price responses to compute the implied change in real welfare across neighborhoods. For each location n, welfare is measured as the difference between income and price effects. Under the partial-equilibrium (ATE) specification, welfare declines modestly across most of the city, with the sharpest losses concentrated in the central and tourist-exposed neighborhoods where prices increase the most. Peripheral areas experience smaller losses, consistent with

Figure 11. Welfare Effects

(a) Average Effects **(b)** Effects with GE Propagation

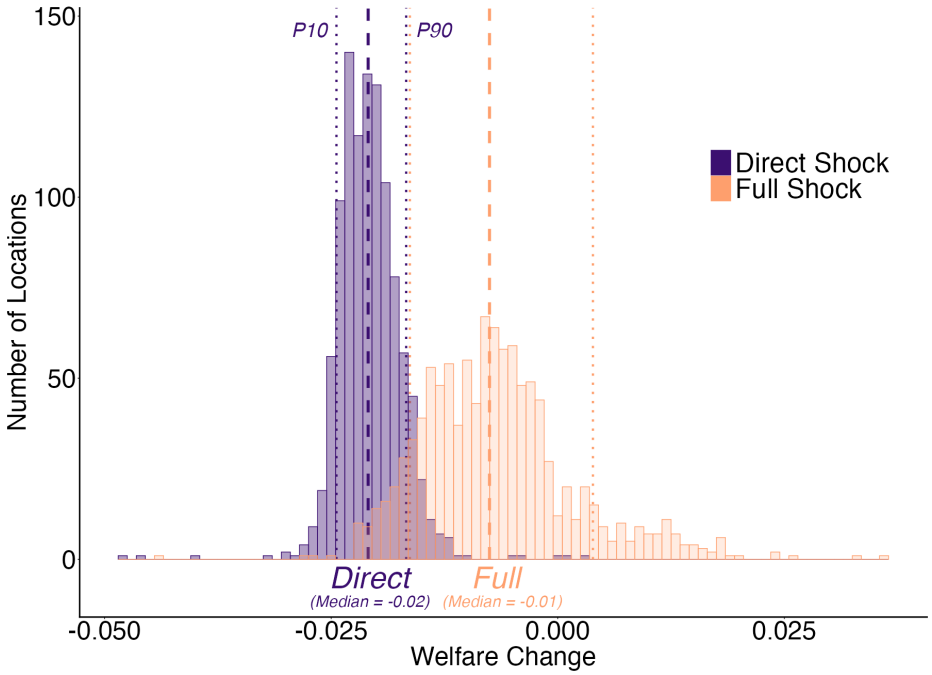


Notes: This figure shows the spatial distribution of predicted welfare effects induced by the tourism shock between January and July 2018. Panel (a) reports the average treatment effect (ATE), which combines local income and price changes under partial equilibrium. Panel (b) displays the general equilibrium (GE) results, which incorporate feedback across neighborhoods through the commuting and spending network. Light colors denote smaller welfare losses (or larger gains).

their lower exposure to the initial tourism shock. When general equilibrium (GE) propagation is introduced, these effects become more positive and more spatially diffuse. GE feedback amplifies the income gains in the most touristy tracts through the commuting network. So while the tourist shock strengthens price pressures, particularly in central tracts, it amplifies the income gains. As a result, the net welfare pattern reflects the balance between these two forces: some areas experience smaller losses or modest gains due to indirect income spillovers, whereas others see larger welfare declines as price effects dominate.

Figure 12 further illustrates the distribution of welfare changes across neighborhoods. Under ATE, the distribution is tightly centered around a median loss of about -0.02, with most locations clustered in a narrow range of small negative changes. By contrast, the GE distribution shifts rightward and becomes wider, with a median loss of roughly -0.01 and a thicker right tail. This indicates that incorporating spatial general equilibrium mechanisms slightly reduces the average welfare loss but mostly increases heterogeneity across neighborhoods. In particular, a subset of locations benefits from strong positive spillovers through the local income channel, while others remain constrained by higher local prices.

Figure 12. Welfare Effects



Notes: This figure shows the distribution of welfare effects across the city of Barcelona. Vertical bands denote 10th and 90th percentiles (dotted) and the median (dashed), for each distribution.

Together, Figures 11 and 12 highlight that the aggregate welfare impact of tourism shocks depends critically on spatial propagation—both in magnitude and in the distribution of gains and losses across the urban landscape.

5 Quantitative Model and Optimal Tourist Taxation

The previous sections used a linear general-equilibrium approximation and reduced-form regressions to map tourism shocks into tract-level income, prices, and welfare. We now turn to a nonlinear city equilibrium that nests those same mechanisms in a unified quantitative model. This model serves two purposes. First, it allows us to evaluate large counterfactual tourism shocks, for which the linear approximation might not be accurate. Second, it provides the policy environment for tourist taxation: the small-tax formulas derived earlier are the baseline derivatives of the same nonlinear hat-algebra system used for the full optimal-tax problem.

5.1 Quantitative model

The quantitative model is a static equilibrium of a city partitioned into blocks and sectors, with residents and tourists interacting through goods and labor markets. The full system is reported in Appendix E; here we summarize the ingredients that matter for the welfare and policy analysis.

On the production side, each block–sector pair (i, s) produces a differentiated variety using Cobb–Douglas technology,

$$Q_{is} = A_{is} L_{is}^{\beta_s} M_{is}^{1-\beta_s}, \qquad 0 < \beta_s < 1,$$

where A_{is} is productivity, L_{is} is labor, and M_{is} is a fixed nonlabor factor. Under perfect competition, producer prices equal unit costs,

$$p_{is} = \frac{w_{is}^{\beta_s} r_{is}^{1-\beta_s}}{A_{is}},$$

so shocks to tourism feed back into prices and wages through the full equilibrium rather than through a fixed partial-equilibrium supply curve.

Residents live in origin tracts n and allocate labor across workplaces subject to commuting frictions. Their nominal income is summarized by a CES wage aggregator, while their expenditure side is a nested CES bundle over local varieties and an outside good. Within sectors, residents substitute across blocks; across sectors, they substitute across sectoral composites. Shopping frictions and amenities shift effective delivered prices but are held fixed across counterfactuals. Tourists are modeled symmetrically on the demand side. They allocate spending across sectors and blocks using a nested CES structure, and total spending on the city responds to the overall tourist price index through an outside-diversion elasticity. A block-specific tourist tax therefore affects both where visitors spend within the city and whether they spend in the city at all.

We solve counterfactual equilibria in hat algebra, writing $\hat{x} \equiv x/x^0$ relative to the observed baseline. For any vector of tourist shocks and taxes, the model jointly determines producer prices, wages, revenues, labor allocations, resident price indices and incomes, and tourist price indices and expenditures. Resident welfare in origin n is measured by the exact nonlinear real-income change

$$\hat{u}_n = \frac{\hat{v}_n}{\hat{G}_n}, \qquad EV_n = v_n^0(\hat{u}_n - 1),$$

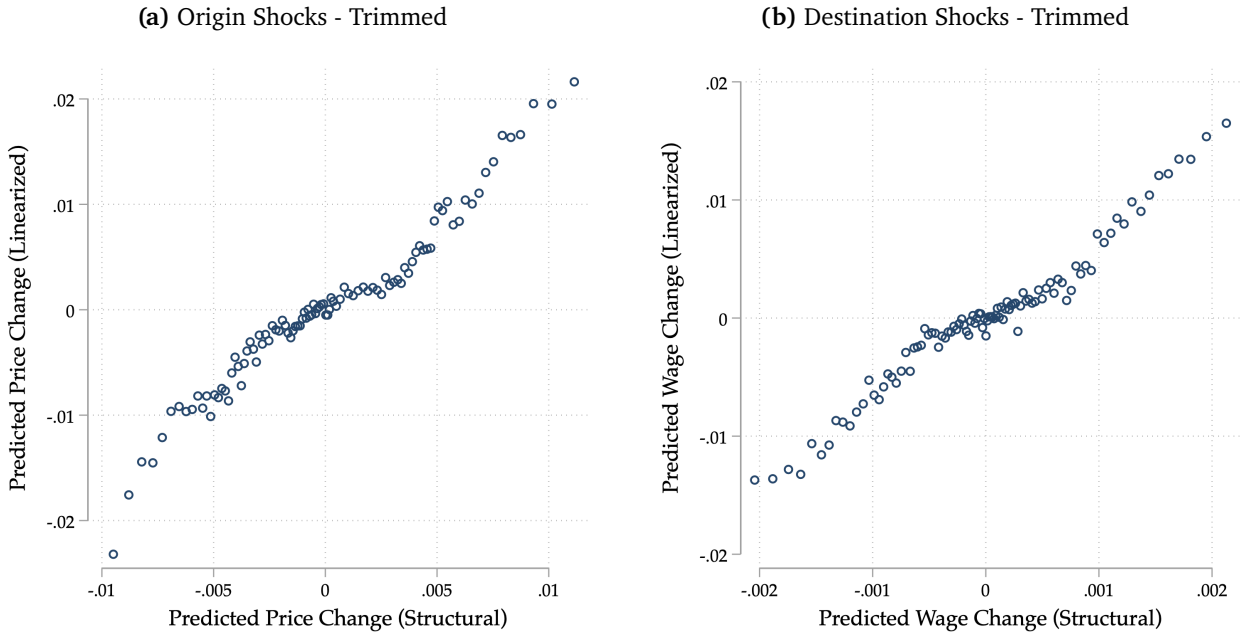
where $\hat{v}_n$ is counterfactual nominal income, $\hat{G}_n$ is the counterfactual cost-of-living index, and EV_n is equivalent variation in baseline income units. This is the nonlinear analogue of the welfare object used in the reduced-form sections.

The model is disciplined using the same inputs that underlie the linear analysis: substitution elas-

ticities from the expenditure-demand estimates, labor shares from national accounts, and commuting elasticities from the relevant commuting literature. Given these parameters and the observed network of commuting and spending flows, the model reproduces the baseline equilibrium exactly. For any counterfactual tourism path or tax vector, we solve the hat system numerically and recover the implied changes in resident incomes, producer prices, tourist spending, and welfare.

This setup also lets us compare the nonlinear model to the earlier linear approximation. Table 1 begins from our baseline reduced-form specification, which finds positive effects on origin income and negative effects on destination prices using the instrumental-variables tourism shocks. Comparing Columns (3) and (4) of, adding the corresponding predicted incomes and prices from the full structural model, $\hat{v}_{nt}^{\text{structural}}$ and $\hat{p}_{it}^{\text{structural}}$ to the linearized GE predictions contributes additional explanatory power, especially through the income terms.

Figure 13. Comparison between Linearized Predicted Effects and Full Structural Model (Raw and Trimmed), residual with n,t FE



The fact that the linear model often performs better in purely within-sample fit is not evidence against the structural approach. The reduced-form linear specification is designed precisely to capture local variation in the observed data, and it can therefore outperform the full equilibrium model when the exercise remains close to the estimation sample. The value of the structural model is different. It imposes equilibrium consistency across prices, wages, and quantities; it smooths sparse cells by borrowing strength across locations; and it remains coherent for out-of-sample counterfactuals, including large tourism shocks and tax experiments, where curvature and nonlinear reallocation matter. This is

the environment in which the structural model has the most to add.

5.2 Optimal tourist taxation

We now use the quantitative model to study optimal tourist taxation. The city chooses a vector of block-specific taxes on purchases made by tourists and rebates the resulting revenue lump-sum to residents. The basic trade-off is intuitive: higher taxes raise revenue from visitors, but they also reduce and reallocate tourist spending across blocks and potentially away from the city altogether. Those spending shifts then feed back into local wages, producer prices, and resident welfare through the same GE channels that govern the model.

For expositional convenience, we parameterize taxes using the log wedge $\theta_i \equiv \ln(1+\tau_i)$, where τ_i is the ad valorem tourist tax rate in block i. It is also useful to distinguish between two spending concepts. Let $E_i^T(\boldsymbol{\theta})$ denote tourist expenditure in block i measured at producer prices. This is the relevant tax base: it is the pre-tax amount on which revenue is collected. Let $X_i^T(\boldsymbol{\theta})$ denote gross tourist payments, including the tax wedge. The two are linked by

$$X_i^T(\boldsymbol{\theta}) = (1+\tau_i)\, E_i^T(\boldsymbol{\theta}) = e^{\theta_i} E_i^T(\boldsymbol{\theta}).$$

This distinction matters because tourist demand is naturally written in terms of what tourists pay, while the government collects revenue on the producer-price base.

The planner chooses the tax vector to maximize resident welfare,

$$W(\boldsymbol{\theta}) = \sum_n \lambda_n \ln u_n(\boldsymbol{\theta}),$$

where $u_n(\boldsymbol{\theta})$ is equilibrium utility in tract n and λ_n is the corresponding population weight. Revenue is

$$R(\boldsymbol{\theta}) = \sum_i \tau_i(\theta_i)\, E_i^T(\boldsymbol{\theta}),$$

and we assume that it is rebated lump-sum to residents. The welfare value of an additional dollar of rebate is summarized by the marginal value of transfers, denoted $m_T(\boldsymbol{\theta})$.

Although the full equilibrium is high-dimensional, the tax problem can be summarized by two local GE objects. The first is the GE tax-base elasticity matrix, denoted $\mathcal{E}(\boldsymbol{\theta})$. Its (i,j) entry measures how a higher tax in block j changes the producer-price tourist tax base in block i, after allowing for within-city substitution, diversion away from the city, and GE feedback through producer prices. The diagonal entries therefore capture how sensitive each block's own tax base is to its own tax, while the off-diagonal

entries capture spatial reallocation across blocks. The second object is the welfare-incidence vector, denoted $\kappa(\boldsymbol{\theta})$. Its ith element measures how a one-percent increase in tourist expenditure in block i affects resident welfare, holding the tax vector fixed. Intuitively, $\kappa(\boldsymbol{\theta})$ summarizes whether marginal tourism in a location is socially valuable or socially costly once wages, prices, and network spillovers are taken into account.

Using these two objects, the first-order condition for the optimal tax can be written as

$$\operatorname{diag}\big(\mathbf{E}^T(\boldsymbol{\theta}^\star)\big)\big(\mathbf{1}+\boldsymbol{\tau}^\star\big) = \mathcal{E}(\boldsymbol{\theta}^\star)'\big(\boldsymbol{\tau}^\star \odot \mathbf{E}^T(\boldsymbol{\theta}^\star)\big) + \frac{\Lambda}{m_T(\boldsymbol{\theta}^\star)}\mathcal{E}(\boldsymbol{\theta}^\star)'\kappa(\boldsymbol{\theta}^\star), \tag{2}$$

where $\Lambda = \sum_n \lambda_n$ and $\odot$ denotes elementwise multiplication. This is the exact nonlinear Ramsey condition implied by the quantitative model. The left-hand side is the direct marginal revenue gain from raising the tax in each block. The first term on the right-hand side captures the erosion of the tax base when tourists reduce spending or substitute across locations. The second term captures the welfare incidence of those spending shifts through wages, prices, and the rest of the equilibrium system.

A useful benchmark shuts down the incidence channel by setting $\kappa(\boldsymbol{\theta}) = \mathbf{0}$. In that case, the planner cares only about revenue and the problem reduces to a revenue-maximizing, or "optimal tariff," benchmark. The full Ramsey problem departs from that benchmark whenever tourism has welfare effects on residents beyond the revenue it generates. If marginal tourism in a location raises resident welfare, the optimal tax there is lower than the revenue-only benchmark. If marginal tourism lowers resident welfare, the optimal tax is higher.

This formulation also clarifies the connection to the earlier linear analysis. Evaluated at the baseline, the nonlinear objects reduce to their local counterparts, and the Ramsey condition becomes a first-order welfare gradient. In the linear model, that gradient is governed by three familiar ingredients: the matrix H, which summarizes how tourist payments respond to tourist-facing prices; the matrix Π, which maps tourism shocks into wage changes through the local network; and the vector ψ, which aggregates resident exposure to wage gains and price increases across origins. These objects recover the same cross-sectional ranking of locations that emerged in the small-tax analysis.

The key limitation of the linear object is that it provides a ranking, not a full tax schedule. It tells us which blocks should face relatively higher taxes at the margin, but it does not pin down scale. Scale requires curvature. As taxes become large, tourist diversion strengthens, pass-through changes, and the welfare incidence of tourism does not need to remain constant. Those are precisely the margins captured by the nonlinear model. In practice, we solve the tax problem by nesting the policy choice inside the hat-algebra equilibrium, updating the tax vector until the nonlinear Ramsey condition in equation (2) is satisfied. For numerical purposes, we initialize the procedure with a frozen-elasticity approximation

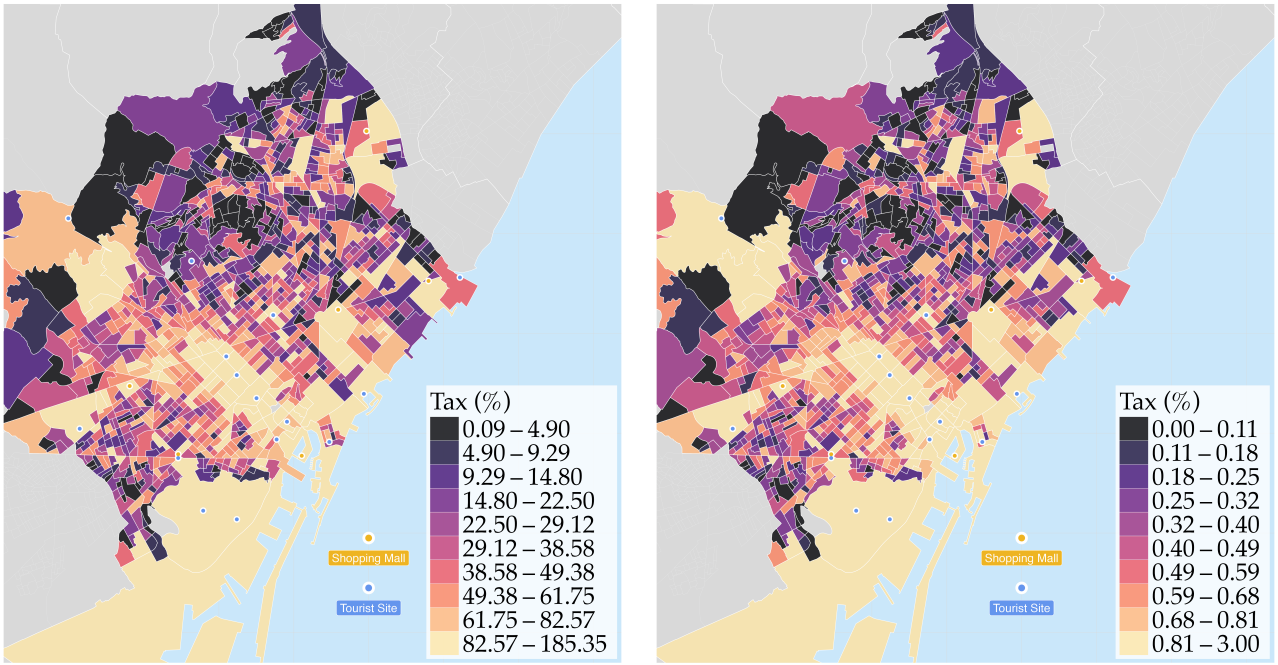
that holds the local GE objects fixed at a reference point, typically the baseline equilibrium.

To implement the tax exercises, we require two demand margins beyond those already estimated in the local expenditure system. The first is an aggregate tourist elasticity, which determines diversion away from the city as taxes raise the overall cost of visiting. The second is a spatial tourist elasticity, which drives substitution across locations within the city. We calibrate the aggregate elasticity using the external evidence in Faber and Gaubert (2019) and the within-city substitution margin using our local consumption estimates.

5.3 Comparing linear and nonlinear tax predictions

Figure 14. Optimal Taxes

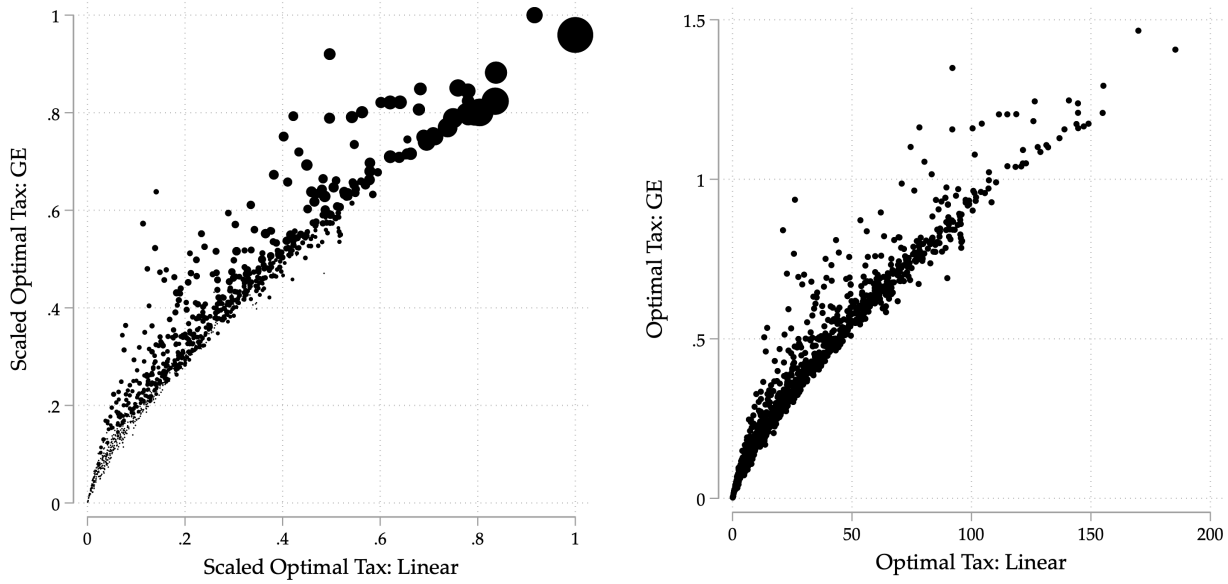
(a) Linearized GE Optimal Tax **(b)** Structural GE Optimal Tax



Notes: On the left, we plot the distribution of the optimal tourist tax from the Linearized GE model (see equation XX). Of the right, we plot the distribution of the optimal tourist tax from the full Structural GE model.

Figure 14 compares the spatial pattern of taxes implied by the linearized GE approximation and the full nonlinear structural model. The correspondence is strong: locations that the linear model identifies as relatively desirable places to tax more heavily are, in general, the same locations that receive higher taxes in the nonlinear solution. Figure 15 makes the same point in cross section. Whether we look at the unweighted comparison or weight locations by their share of tourist expenditure, the two objects are highly correlated.

Figure 15. Scaling of Linearized Tax and Full Optimal Structural Tax



Notes: On the RHS we weight the observations with the share of tourist expenditures at the location (note, we could formally put the 'GE' weight there)

At the same time, the comparison highlights two systematic differences between the linear and nonlinear approaches: scale and curvature. The linearized GE object is primarily a within-city ranking. It tells us which locations should be taxed more than others, but it does not by itself determine the overall level of taxation. In particular, it does not incorporate the nonlinear dampening of percolation and substitution effects as taxes become large. The structural model does. Once taxes move far enough from the baseline, second-order effects become quantitatively important, and they determine the scale of the optimal tax schedule.

The second difference is curvature across locations. In the linear approximation, the mapping from local exposure to optimal taxation is essentially proportional, so departures from the trend are driven mostly by estimated local elasticities. In the nonlinear model, that relationship bends as general-equilibrium feedback, diversion, and reallocation intensify. Some of that curvature reflects genuine economics: heavily taxed locations become less attractive at a non-constant rate, and the burden of taxation shifts toward nearby substitutes and toward the outside option. Some may also reflect the fact that the structural model smooths across sparse cells, echoing the concerns in Dingel and Tintelnot (2020). Either way, these are margins that are absent by construction from the linearized approximation.

This comparison also clarifies where the structural model adds value. For within-sample prediction, the linearized GE approach can be hard to beat because it is closely tied to the variation used in estimation. But the policy problem is inherently out of sample. Designing tourist taxes requires extrapolating

beyond the observed equilibrium and asking how the city responds when wedges become large enough to reallocate spending across blocks and away from the city. That is exactly the setting in which the nonlinear model becomes useful. It preserves the cross-sectional ranking emphasized by the linear object, while adding the scale, curvature, and incidence adjustments needed to turn that ranking into a coherent equilibrium tax schedule.

6 Conclusion

This paper studies how tourism shocks affect resident welfare within the city. We show that tourist shocks operate through two distinct channels: tourism raises price pressure in the places where spending takes place, but it can also increase resident income in connected locations through the urban network. Once residents differ in where they live, work, and shop, the welfare incidence of tourism becomes inherently heterogeneous across neighborhoods. Measuring that heterogeneity therefore requires going beyond direct exposure and tracing how shocks propagate through commuting and spending linkages.

Using bilateral expenditure data for Barcelona, we develop an empirical strategy that is explicitly designed for this environment. The analysis separates destination price effects from origin income effects and shows that both are quantitatively important. Tourism shocks do not remain confined to the initially exposed tracts. Instead, they diffuse through the city, generating negative consumption effects in some locations and positive income effects in others. This network propagation is central for welfare measurement: the same tourism shock can reduce real income for residents facing higher local prices while benefiting residents whose incomes rise through indirect exposure.

More broadly, the paper shows how spatial expenditure data can be combined with an urban-network framework to recover welfare-relevant objects directly from the data. This provides a way to move from reduced-form estimates to tract-level welfare incidence without imposing that price and income responses coincide mechanically.

A natural next step is to embed these empirical welfare objects in a fuller nonlinear quantitative model, allowing the analysis of larger shocks and stronger equilibrium feedback. Another is to study dynamic adjustment, including housing, labor supply, and location responses over longer horizons. More generally, the framework could be used to evaluate local fiscal policies by tracing how their effects are transmitted through urban networks and distributed across residents.

References

Adão, Rodrigo, Costas Arkolakis, and Sharat Ganapati. 2020. "From heterogeneous firms to heterogeneous trade elasticities: The aggregate implications of firm export decisions." National Bureau of Economic Research.

Agarwal, Sumit, J. Bradford Jensen, and Ferdinando Monte. 2017. "Consumer Mobility and the Local Structure of Consumption Industries." National Bureau of Economic Research, Inc NBER Working Papers 23616.

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

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

Almagro, Milena, and Tomás Domínguez-Iino. 2019. "Location Sorting and Endogenous Amenities: Evidence from Amsterdam." Working Paper.

Athey, Susan, David Blei, Robert Donnelly, Francisco Ruiz, and Tobias Schmidt. 2018. "Estimating Heterogeneous Consumer Preferences for Restaurants and Travel Time Using Mobile Location Data." *AEA Papers and Proceedings*, 108: 64–67.

Atkin, David, Benjamin Faber, and Marco Gonzalez-Navarro. 2018. "Retail Globalization and Household Welfare: Evidence from Mexico." *Journal of Political Economy*, 126(1): 1–73.

Baqaee, David, and Ariel Burstein. 2021. "Welfare and Output with Income Effects and Demand Instability." National Bureau of Economic Research Working Paper 28754.

Baqaee, David Rezza, and Emmanuel Farhi. 2019. "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem." *Econometrica*, 87(4): 1155–1203.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." *Review of Economic Studies*, 89(1): 181–213.

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

Couture, Victor, Jonathan Dingel, Allison Green, and Jessie Handbury. 2020. "Quantifying Social Interactions Using Smartphone Data." *Working paper*.

Davis, Donald R., Jonathan I. Dingel, Joan Monras, and Eduardo Morales. 2019. "How Segregated Is Urban Consumption?" *Journal of Political Economy*, 127(4): 1684–1738.

Deaton, Angus. 1989. "Rice Prices and Income Distribution in Thailand: A Non-parametric Analysis." *Economic Journal*, 99(395): 1–37.

Dingel, Jonathan I, and Felix Tintelnot. 2020. "Spatial economics for granular settings." National Bureau of Economic Research.

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

Faber, Benjamin, and Cecile Gaubert. 2019. "Tourism and Economic Development: Evidence from Mexico's Coastline." *American Economic Review*, 109(6): 2245–93.

García-López, Miquel-Àngel, Jordi Jofre-Monseny, Rodrigo Martínez Mazza, and Mariona Segú. 2019. "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona." Working Papers 2019/05.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2018. "Bartik Instruments: What, When, Why, and How." National Bureau of Economic Research, Inc NBER Working Papers 24408.

Houthakker, H. S. 1952. "Compensated Changes in Quantities and Qualities Consumed." *The Review of Economic Studies*, 19(3): 155–164.

Kim, Ryan, and Jonathan Vogel. 2020. "Trade and Welfare (Across Local Labor Markets)." National Bureau of Economic Research, Inc NBER Working Papers 27133.

Kleinman, Benny, Ernest Liu, and Stephen J Redding. 2020. "International Friends and Enemies." National Bureau of Economic Research Working Paper 27587.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review*, 108(12): 3855–90.

Porto, Guido G. 2006. "Using survey data to assess the distributional effects of trade policy." *Journal of International Economics*, 70(1): 140–160.

Appendix: Notation and Roadmap

This appendix collects additional derivations and details for the theoretical and quantitative framework in the main text. We use the following global notation throughout.

Indices and sets.

- $i \in \{1, \ldots, N\}$: production block (destination / workplace).
- $n \in \{1, \ldots, N\}$: residential block (origin).
- $s \in \{1, \ldots, S\}$: sector.
- $g \in \{1, \ldots, G\}$: tourist-origin group.
- t: time (e.g. month-year) index in the empirical sections.

Prices, wages, and quantities.

- w_i: wage in block i (one-sector case); w_{is} in the multi-sector model.
- p_i: producer price in block i (one-sector); p_{is} in the multi-sector model.
- $p_{ni} = \tau_{ni} p_i$: delivered price to resident n for good i, with $\tau_{ni} \geq 1$ shopping frictions and $\tau_{nn} = 1$.
- $P_i^T = (1 + \tau_i) p_i$: tourist-facing price in block i under tourist tax τ_i.
- q_{ni}: quantity of i consumed by resident n.
- Q_i or Q_{is}: physical output of good (i, s).

Incomes and expenditures.

- T_n: time endowment of resident n.
- $H_n(\{\ell_{ni}\})$: commuting / time-usage aggregator.
- $J(\mathbf{w})$, J_n: wage aggregator(s) dual to H_n.
- $v_n = T_n J(\mathbf{w})$: disposable income of resident n in the one-sector linear model.
- v_n: more generally, disposable income of resident n in the structural GE / hat-algebra model.
- $y_i = p_i q_i$: nominal sales (revenue) of block i.
- E_i^T: tourist expenditure in block i; $\mathbf{E}^T = (E_1^T, \ldots, E_N^T)'$.
- E_C^T: total tourist expenditure in the city.

Preferences and price indices.

- $G(\mathbf{p}_n)$: homothetic price index for resident n; $u_n = v_n / G(\mathbf{p}_n)$.
- P_{ns}: within-sector CES price index for residents in origin n, sector s.
- P_{gs}^T: within-sector CES price index for tourists from group g, sector s.
- G_n: across-sector CES price index for residents in n.
- G_g^T: across-sector CES price index for tourists from g.

Shares and exposure weights.

- c_{ni}: income share of resident n earned in block i.
- s_{ni}: expenditure share of resident n spent in block i.
- $\mathbf{c}_n$: row vector $(c_{n1}, \ldots, c_{nN})$.
- $\mathbf{s}_n$: row vector $(s_{n1}, \ldots, s_{nN})$.
- s_{nis}, s^T_{gis}: within-sector shares for residents and tourists in the structural model.
- b_{ns}, b^T_{gs}: across-sector budget shares for residents and tourists.

Shock notation and GE operators.

- $\mathbf{g} = d\ln \mathbf{E}^T = (d\ln E^T_1, \ldots, d\ln E^T_N)'$: tourism-expenditure shock vector.
- S: matrix of baseline expenditure shares, $S_{in} = s_{ni}$.
- C: matrix of baseline income shares, $C_{nj} = c_{nj}$.
- $D_y = \mathrm{diag}(y_1, \ldots, y_N)$: diagonal matrix of baseline revenues.
- $D_v = \mathrm{diag}(v_1, \ldots, v_N)$: diagonal matrix of baseline incomes.
- $\mathbf{D}_T = \mathrm{diag}(E^T_1/y_1, \ldots, E^T_N/y_N)$: tourism-intensity diagonal.
- $\mathbf{M} = D_y^{-1} S D_v C$: one-round income–spending propagation matrix.
- $\mathbf{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T$: percolation operator mapping shocks into wages (and, via β, prices).
- H: nested-CES semi-elasticity matrix for tourist demand (Appendix D).

Parameters.

- $\beta \in (0, 1)$: labor share in Cobb–Douglas production; also wage-to-price pass-through parameter in the linear model.
- σ, σ_s: within-sector CES elasticities for residents.
- σ_T, σ^T_s: within-city CES elasticities for tourists.
- η, η^T: across-sector (residents) and across-city (tourists) elasticities.
- $\epsilon > 1$: commuting elasticity.

Roadmap.

- Appendix A (A) derives the envelope theorem for resident welfare and shows how income and expenditure shares enter the welfare differential.
- Appendix B (B) introduces production, market clearing, and the linear GE mapping from tourism shocks to wages, prices, and welfare, defining the percolation operator $\mathbf{\Pi}$.
- Appendix C (C) studies causal inference in this GE setting, showing how naive shift–share exposure violates SUTVA and defining the model-consistent percolated exposure.

- Appendix D (D) introduces nested CES tourist demand and derives the linear small-tax optimal tourist tax rule in terms of H, $\mathbf{\Pi}$, and incidence.
- Appendix E (E) records the full nonlinear structural GE system in hat algebra (counterfactual-to-baseline ratios) with nested CES demand, CES commuting, and Cobb–Douglas production.
- Appendix F (F) embeds block-specific tourist taxes into the hat system, derives the exact welfare-based first-order conditions, and shows that their linearization reproduces the small-tax rule in Appendix D.

A Appendix A: Setting and Envelope Theorem

This appendix formalizes the primitives of the household problem and derives the welfare envelope used in the main text. The derivation proceeds in four steps: (i) we introduce preferences and show that homotheticity implies an indirect-utility representation of the form $u_n = v_n/G(\mathbf{p}_n)$; (ii) we characterize the income-maximization problem with commuting and define the wage aggregator $J(\mathbf{w})$; (iii) we use the envelope theorem and Euler's theorem to obtain labor-income shares c_{ni}, and Roy's identity to obtain expenditure shares s_{ni}; (iv) we combine these pieces to derive the log-differential of indirect utility, which corresponds to equation (T.1) in the main text.

Throughout, there are $N \in \mathbb{N}$ production blocks indexed by

$$i \in \{1, \dots, N\}.$$

A.1 Preferences and Indirect Utility

Producer prices and wages are stacked as

$$\mathbf{p} = (p_1, \dots, p_N)' \in \mathbb{R}^N_{++}, \qquad \mathbf{w} = (w_1, \dots, w_N)' \in \mathbb{R}^N_{++}.$$

A resident of block $n \in \{1, \dots, N\}$ faces delivered prices

$$\mathbf{p}_n = (p_{n1}, \dots, p_{nN})' \in \mathbb{R}^N_{++}, \qquad p_{ni} = \tau_{ni} p_i,$$

where $\tau_{ni} \geq 1$ captures shopping and access frictions, with $\tau_{nn} = 1$ by normalization.

Preferences over the N varieties are homothetic and admit an expenditure function of the form

$$C(\mathbf{p}_n, u_n) = u_n \, G(\mathbf{p}_n),$$

where u_n is utility and

$$G : \mathbb{R}^N_{++} \to \mathbb{R}_{++}$$

is a continuous, differentiable, strictly increasing componentwise, and homogeneous-of-degree-one price index (e.g., a CES price index). Setting the expenditure equal to disposable income v_n and solving for utility yields the indirect utility representation

$$u_n \;=\; \frac{v_n}{G(\mathbf{p}_n)}. \tag{A.1}$$

By Roy's identity applied to $u_n(v_n, \mathbf{p}_n) = v_n/G(\mathbf{p}_n)$, the Marshallian demand for variety i is

$$q_{ni}(v_n, \mathbf{p}_n) = -\frac{\partial u_n/\partial p_{ni}}{\partial u_n/\partial v_n} = \frac{v_n}{G(\mathbf{p}_n)} \frac{\partial G(\mathbf{p}_n)}{\partial p_{ni}}.$$

Defining the budget share

$$s_{ni} \;\equiv\; \frac{p_{ni} q_{ni}}{v_n},$$

we obtain

$$s_{ni} \;=\; \frac{p_{ni}}{G(\mathbf{p}_n)}\,\frac{\partial G(\mathbf{p}_n)}{\partial p_{ni}}. \tag{A.2}$$

Because G is homogeneous of degree one and increasing, Euler's theorem implies

$$\sum_{i=1}^{N} s_{ni} = \sum_{i=1}^{N} \frac{p_{ni}}{G(\mathbf{p}_n)}\,\frac{\partial G(\mathbf{p}_n)}{\partial p_{ni}} = 1,$$

and monotonicity implies $s_{ni} \geq 0$ for all i.

As a CES sanity check, consider

$$G(\mathbf{p}_n) = \left(\sum_{i=1}^{N} \tau_{ni}^{1-\sigma} p_i^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \qquad \sigma > 1.$$

Then

$$s_{ni} = \frac{\tau_{ni}^{1-\sigma} p_i^{1-\sigma}}{\sum_{j=1}^{N} \tau_{nj}^{1-\sigma} p_j^{1-\sigma}},$$

which is the standard CES budget share.

A.2 Commuting, Income Maximization, and the Wage Aggregator

Residents allocate time across blocks according to a commuting technology that is continuously differentiable, strictly increasing componentwise, convex, and homogeneous of degree one:

$$H_n : \mathbb{R}_+^N \to \mathbb{R}_+, \qquad H_n(\boldsymbol{\ell}_n) = T_n,$$

where

$$\boldsymbol{\ell}_n = (\ell_{n1}, \ldots, \ell_{nN})' \in \mathbb{R}_+^N, \qquad T_n > 0$$

denote labor supplied from residence n to each block i and the total time endowment, respectively.

Given wages $\mathbf{w}$, the resident of block n solves:

$$y_n(\mathbf{w}, T_n) \;=\; \max_{\boldsymbol{\ell}_n \in \mathbb{R}_+^N} \big\{ \mathbf{w}'\boldsymbol{\ell}_n :\; H_n(\boldsymbol{\ell}_n) = T_n \big\}. \tag{A.3}$$

With $T_n > 0$ and H_n continuous and increasing, Slater's condition is satisfied. Strict convexity of H_n ensures a unique optimizer $\boldsymbol{\ell}_n^*(\mathbf{w}, T_n)$ and multiplier $\lambda^*(\mathbf{w}, T_n) > 0$.

Define the Lagrangian

$$\mathcal{L}(\boldsymbol{\ell}_n, \lambda) = \mathbf{w}'\boldsymbol{\ell}_n - \lambda\big(H_n(\boldsymbol{\ell}_n) - T_n\big).$$

Stationarity and feasibility yield the KKT conditions

$$w_i - \lambda\,\frac{\partial H_n(\boldsymbol{\ell}_n)}{\partial \ell_{ni}} = 0 \;\; \forall i, \qquad H_n(\boldsymbol{\ell}_n) = T_n. \tag{A.4}$$

Using degree-one homogeneity of H_n, every feasible $\boldsymbol{\ell}_n$ can be written as

$$\boldsymbol{\ell}_n = T_n \mathbf{x}, \qquad H_n(\mathbf{x}) = 1.$$

Substituting into (A.3) shows that optimal income is linear in T_n:

$$y_n(\mathbf{w}, T_n) = T_n \max_{\mathbf{x} \in \mathbb{R}_+^N} \big\{ \mathbf{w}'\mathbf{x} :\; H_n(\mathbf{x}) = 1 \big\}. \tag{A.5}$$

This motivates the definition of the *wage (commuting) aggregator*

$$J(\mathbf{w}) \;:=\; \max_{\mathbf{x}\in\mathbb{R}^N_+} \big\{\mathbf{w}'\mathbf{x} :\; H_n(\mathbf{x}) = 1\big\}, \qquad y_n(\mathbf{w}, T_n) = T_n\, J(\mathbf{w}). \tag{A.6}$$

On $\mathbb{R}^N_{++}$, J is convex, nondecreasing, and homogeneous of degree one in wages.

We identify optimized income with disposable income:

$$v_n := y_n(\mathbf{w}, T_n) = T_n J(\mathbf{w}).$$

A.3 Income Shares and Labor Supply

By the envelope theorem, keeping T_n fixed, the derivative of y_n with respect to wage w_i equals the optimal labor supplied in location i:

$$\frac{\partial y_n(\mathbf{w}, T_n)}{\partial w_i} \;=\; \ell^*_{ni}(\mathbf{w}, T_n). \tag{A.7}$$

Differentiating (A.6) with respect to w_i and using differentiability of J gives

$$\ell^*_{ni}(\mathbf{w}, T_n) \;=\; T_n\, \frac{\partial J(\mathbf{w})}{\partial w_i}. \tag{A.8}$$

Combining this with $v_n = T_n J(\mathbf{w})$ yields a Roy-type representation of optimal labor supply:

$$\ell^*_{ni}(\mathbf{w}, T_n) \;=\; v_n\, \frac{1}{J(\mathbf{w})}\, \frac{\partial J(\mathbf{w})}{\partial w_i}. \tag{A.9}$$

Define the fraction of income that resident n earns in location i as

$$c_{ni} := \frac{w_i\, \ell^*_{ni}}{v_n}.$$

Substituting (A.9) produces

$$c_{ni} \;=\; \frac{w_i}{J(\mathbf{w})}\, \frac{\partial J(\mathbf{w})}{\partial w_i}. \tag{A.10}$$

Since J is homogeneous of degree one, Euler's theorem implies

$$J(\mathbf{w}) = \sum_{i=1}^{N} w_i\, \frac{\partial J(\mathbf{w})}{\partial w_i}.$$

Dividing both sides by $J(\mathbf{w})$ and using (A.10) yields

$$\sum_{i=1}^{N} c_{ni} = 1, \qquad c_{ni} \geq 0 \;\text{ for all } i,$$

where nonnegativity follows from the fact that J is increasing in each component of $\mathbf{w}$. Thus $(c_{n1}, \ldots, c_{nN})$ is a well-defined vector of labor-income shares that sum to one.

A.4 Welfare Envelope

We now combine the income and price components to obtain the first-order welfare response to small changes in wages and prices. Differentiating $\ln u_n$ from (A.1) along any path that holds frictions $\boldsymbol{\tau}_n$ fixed gives

$$d\ln u_n = d\ln v_n - d\ln G(\mathbf{p}_n). \tag{A.11}$$

Income component. Because $v_n = T_n J(\mathbf{w})$ and T_n is fixed,

$$dv_n = T_n \sum_{i=1}^{N} \frac{\partial J(\mathbf{w})}{\partial w_i}\, dw_i.$$

Dividing by $v_n = T_n J(\mathbf{w})$ and rewriting in log-differential form,

$$d\ln v_n = \sum_{i=1}^{N} \frac{w_i}{J(\mathbf{w})} \frac{\partial J(\mathbf{w})}{\partial w_i}\, d\ln w_i = \sum_{i=1}^{N} c_{ni}\, d\ln w_i,$$

where the last equality uses (A.10).

Price component. For the price index, the chain rule and homogeneity of G imply

$$d\ln G(\mathbf{p}_n) = \sum_{i=1}^{N} \frac{\partial \ln G(\mathbf{p}_n)}{\partial p_{ni}}\, dp_{ni} = \sum_{i=1}^{N} \frac{p_{ni}}{G(\mathbf{p}_n)} \frac{\partial G(\mathbf{p}_n)}{\partial p_{ni}}\, d\ln p_{ni} = \sum_{i=1}^{N} s_{ni}\, d\ln p_{ni},$$

where the last equality invokes (A.2).

Combining income and prices. Substituting these expressions into (A.11) yields

$$d\ln u_n = \left[\sum_{i=1}^{N} c_{ni}\, d\ln w_i\right] - \left[\sum_{i=1}^{N} s_{ni}\, d\ln p_{ni}\right].$$

When delivered prices satisfy $p_{ni} = \tau_{ni} p_i$ with frictions $\boldsymbol{\tau}_n$ held fixed, we have $d\ln p_{ni} = d\ln p_i$ for all i, and the envelope simplifies to

$$d\ln u_n \;=\; \sum_{i=1}^{N} c_{ni}\, d\ln w_i \;-\; \sum_{i=1}^{N} s_{ni}\, d\ln p_i. \tag{A.12}$$

This is exactly equation (T.1) in the main text.

As a numeraire sanity check, if all wages and all producer prices change by the same proportion κ, then

$$d\ln w_i = \kappa, \quad d\ln p_i = \kappa \;\; \forall i,$$

so using $\sum_i c_{ni} = 1 = \sum_i s_{ni}$, we obtain

$$d\ln u_n = \kappa - \kappa = 0,$$

consistent with homotheticity: a common proportional scaling of all income and prices leaves real utility unchanged.

Summary and regularity conditions. Equations (A.1)–(A.12) hold under the following regularity assumptions:

- G is continuously differentiable, strictly increasing, and homogeneous of degree one on $\mathbb{R}^N_{++}$.
- H_n is continuously differentiable, strictly increasing, convex, and homogeneous of degree one on $\mathbb{R}^N_{+}$.
- Wages and prices are strictly positive, $\mathbf{w} \in \mathbb{R}^N_{++}$, $\mathbf{p}_n \in \mathbb{R}^N_{++}$, and the time endowment satisfies $T_n > 0$.

These conditions ensure existence and uniqueness of the optimal labor allocation $\boldsymbol{\ell}^*_n$, differentiability of J, validity of the envelope theorem and Euler's theorem, and well-defined, nonnegative budget and income shares that each sum to one. Under these conditions, (A.12) provides the foundation for the welfare analysis in Sections 2.2 and **??** of the main text.

B Appendix B (for Section 2.3): Production, Market Clearing, and GE Transmission

This appendix derives the general–equilibrium mapping from a vector of tourism–expenditure shocks to wages, prices, and (via Appendix A) welfare. The argument proceeds in five steps: (i) we solve the unit–cost problem under Cobb–Douglas technology and derive the short–run price–to–wage pass–through; (ii) we log–linearize the goods–market clearing condition with tourism demand; (iii) we impose labor–market clearing and obtain the linear system linking wages to tourism shocks; (iv) we stack this system in matrix form, establish invertibility via a spectral–radius bound, and define the percolation operator; (v) we substitute the implied wage and price responses into the welfare envelope to obtain the closed–form GE welfare mapping used in the main text.

Throughout this appendix we adopt the notation from Sections 2–?? and Appendix A. In particular,

$$s_{ni} \quad \text{and} \quad c_{ni}$$

denote resident n's expenditure and income shares on block i as derived in Appendix A, and v_n denotes resident n's disposable income.

B.1 Technology and Unit Cost

Each production block $i \in \{1, \dots, N\}$ produces a differentiated good using Cobb–Douglas technology with labor share $\beta \in (0, 1)$:

$$Q_i = A_i L_i^{\beta} K_i^{1-\beta},$$

where $A_i > 0$ is productivity, L_i is labor input, and K_i is (fixed) capital. Let w_i be the wage in block i and $r_i > 0$ the rental rate of capital and other nonlabor inputs. Under perfect competition, the producer price equals unit cost:

$$p_i \;=\; \min_{L_i, K_i} \left\{ w_i L_i + r_i K_i : \; A_i L_i^{\beta} K_i^{1-\beta} \geq 1 \right\}. \tag{B.1}$$

The solution to (B.1) is the familiar Cobb–Douglas unit cost:

$$p_i = \frac{1}{A_i} \big(\beta w_i\big)^{\beta} \big((1-\beta) r_i\big)^{1-\beta}. \tag{B.2}$$

Taking logs and totally differentiating (B.2) along a short–run path with $d \ln A_i = 0$ and $d \ln r_i = 0$ yields

$$d \ln p_i \;=\; \beta \, d \ln w_i. \tag{B.3}$$

This is the reduced–form short–run pass–through from wages to producer prices used in the main text. Heterogeneous labor shares β_i can be accommodated at the cost of heavier notation; we maintain a common β for clarity.

B.2 Goods–Market Clearing with Tourism Demand

Let

$$y_i := p_i Q_i$$

denote nominal sales (revenue) of good i, and let $E_i^T \geq 0$ denote tourist expenditure on i. As in the main text, goods–market clearing in revenue terms equates total sales to residents' spending plus tourist spending:

$$y_i \;=\; \sum_{n=1}^{N} s_{ni} \, v_n \;+\; E_i^T, \tag{B.4}$$

where

$$s_{ni} = \frac{p_{ni} q_{ni}}{v_n}, \qquad v_n = T_n J(\mathbf{w}),$$

and $p_{ni} = \tau_{ni} p_i$ is the delivered price faced by resident n for good i (Appendix A).

We treat the expenditure shares s_{ni} as fixed at their baseline values when we consider small shocks to tourism; the time endowment T_n and commuting technology are also held fixed, so the baseline v_n's respond only to wages.

Totally differentiate (B.4) and use that s_{ni} is fixed:

$$dy_i = \sum_{n=1}^{N} s_{ni}\, dv_n + dE_i^T.$$

Divide through by $y_i > 0$ to obtain the log–differential:

$$d\ln y_i = \sum_{n=1}^{N} \left(\frac{s_{ni} v_n}{y_i} \right) d\ln v_n \;+\; \left(\frac{E_i^T}{y_i} \right) d\ln E_i^T. \tag{B.5}$$

From Appendix A, optimized income satisfies

$$v_n = T_n J(\mathbf{w}),$$

and, holding T_n fixed,

$$d\ln v_n = \sum_{j=1}^{N} c_{nj}\, d\ln w_j,$$

where c_{nj} is the income share earned by resident n in block j, with $\sum_j c_{nj} = 1$. Substituting this into (B.5) gives

$$d\ln y_i = \sum_{j=1}^{N} \underbrace{\left[\sum_{n=1}^{N} \left(\frac{s_{ni} v_n}{y_i} \right) c_{nj} \right]}_{=:M_{ij}} d\ln w_j \;+\; \underbrace{\left(\frac{E_i^T}{y_i} \right)}_{=:D_{T,i}} d\ln E_i^T. \tag{B.6}$$

B.3 Labor–Market Clearing and the Wage System

Let $L_i = \sum_n \ell_{ni}$ denote total labor used in block i. Under Cobb–Douglas technology with labor share β and perfect competition, labor's share of revenue is constant:

$$w_i L_i = \beta\, y_i. \tag{B.7}$$

We adopt a short–run closure in which L_i is fixed at its baseline level (commuting patterns across blocks are treated as predetermined at the moment of the shock). With L_i and β constant, log–differentiating (B.7) yields

$$d\ln w_i = d\ln y_i.$$

Combining this identity with (B.6), we obtain for each block i:

$$d\ln w_i = \sum_{j=1}^{N} M_{ij}\, d\ln w_j \;+\; D_{T,i}\, d\ln E_i^T. \tag{B.8}$$

Stacking (B.8) over all i gives the vector equation

$$d\ln \mathbf{w} = \mathbf{M}\, d\ln \mathbf{w} + \mathbf{D}_T\, d\ln \mathbf{E}^T, \tag{B.9}$$

where

$$d\ln \mathbf{w} = (d\ln w_1, \ldots, d\ln w_N)' \in \mathbb{R}^N, \quad d\ln \mathbf{E}^T = (d\ln E_1^T, \ldots, d\ln E_N^T)' \in \mathbb{R}^N,$$

$\mathbf{M} = [M_{ij}] \in \mathbb{R}^{N\times N}$ with M_{ij} defined in (B.6), and

$$\mathbf{D}_T := \operatorname{diag}(D_{T,1}, \ldots, D_{T,N}) = \operatorname{diag}\big(E_1^T/y_1, \ldots, E_N^T/y_N\big) \in \mathbb{R}^{N\times N}.$$

It is useful to express $\mathbf{M}$ in compact matrix notation. Let

- $S \in \mathbb{R}^{N\times N}$ collect expenditure shares by (i,n) such that $S_{in} := s_{ni}$;
- $C \in \mathbb{R}^{N\times N}$ collect income shares by (n,j) such that $C_{nj} := c_{nj}$;
- $D_y := \operatorname{diag}(y_1, \ldots, y_N)$ and $D_v := \operatorname{diag}(v_1, \ldots, v_N)$.

Then (B.6) implies

$$\mathbf{M} = D_y^{-1}\, S\, D_v\, C, \qquad \mathbf{M} \in \mathbb{R}^{N\times N}_{\geq 0}. \tag{B.10}$$

The element M_{ij} is thus the share–weighted effect of wage changes in block j on sales in block i through the income–spending network.

Rearranging (B.9) gives

$$(I - \mathbf{M})\, d\ln \mathbf{w} = \mathbf{D}_T\, d\ln \mathbf{E}^T.$$

To solve for $d\ln \mathbf{w}$ we must ensure that $(I - \mathbf{M})$ is invertible. This is guaranteed by a spectral–radius condition.

B.4 Invertibility and the Percolation Operator

Since $\mathbf{M} \geq 0$ elementwise, Perron–Frobenius implies that its spectral radius satisfies

$$\rho(\mathbf{M}) \leq \|\mathbf{M}\|_\infty,$$

where $\|\mathbf{M}\|_\infty$ is the maximum row sum. From the definition of M_{ij} in (B.6),

$$\sum_{j=1}^N M_{ij} = \sum_{j=1}^N \sum_{n=1}^N \Big(\frac{s_{ni}v_n}{y_i}\Big) c_{nj} = \sum_{n=1}^N \Big(\frac{s_{ni}v_n}{y_i}\Big) \underbrace{\sum_{j=1}^N c_{nj}}_{=1} = \frac{\sum_{n=1}^N s_{ni}v_n}{y_i}.$$

Using goods–market clearing (B.4), $\sum_n s_{ni}v_n = y_i - E_i^T$, so

$$\sum_{j=1}^N M_{ij} = \frac{y_i - E_i^T}{y_i} = 1 - \frac{E_i^T}{y_i} = 1 - D_{T,i}.$$

Therefore

$$\|\mathbf{M}\|_\infty = \max_i \sum_j M_{ij} = \max_i (1 - D_{T,i}) = 1 - \min_i D_{T,i}$$

and

$$\rho(\mathbf{M}) \;\leq\; 1 - \min_i D_{T,i}. \tag{B.11}$$

If every block has a strictly positive tourism share of revenue, $D_{T,i} > 0$ for all i, then $\min_i D_{T,i} > 0$ and hence $\rho(\mathbf{M}) < 1$. More generally, the condition $\rho(\mathbf{M}) < 1$ can hold even if some blocks have $D_{T,i} = 0$, provided the row sums remain strictly below one.

Whenever $\rho(\mathbf{M}) < 1$, $(I - \mathbf{M})$ is nonsingular and its inverse admits the convergent Neumann series

$$(I - \mathbf{M})^{-1} = \sum_{k=0}^\infty \mathbf{M}^k.$$

Thus the solution to the wage system (B.9) is

$$d\ln \mathbf{w} \;=\; (I-\mathbf{M})^{-1}\,\mathbf{D}_T\,d\ln \mathbf{E}^T. \tag{B.12}$$

It is convenient to define the *percolation operator*

$$\mathbf{\Pi} \;\equiv\; (I-\mathbf{M})^{-1}\,\mathbf{D}_T = \sum_{k=0}^{\infty}\mathbf{M}^k\,\mathbf{D}_T. \tag{B.13}$$

The diagonal matrix $\mathbf{D}_T$ injects a marginal tourism shock where it is spent; the matrix $\mathbf{M}$ maps this shock through one round of income–spending feedback; and the powers $\mathbf{M}^k$ with $k \geq 1$ capture higher–order rounds of propagation.

Using the pass–through relation (B.3), the corresponding price response is

$$d\ln \mathbf{p} = \beta\,d\ln \mathbf{w} = \beta\,\mathbf{\Pi}\,d\ln \mathbf{E}^T. \tag{B.14}$$

B.5 GE Welfare Mapping

We now combine the wage and price responses with the welfare envelope derived in Appendix A. Let

$$\mathbf{c}_n = (c_{n1},\ldots,c_{nN}) \in \mathbb{R}^{1\times N}, \qquad \mathbf{s}_n = (s_{n1},\ldots,s_{nN}) \in \mathbb{R}^{1\times N}$$

denote the income and expenditure share rows associated with resident n. As shown in (**??**), along any path holding frictions $\boldsymbol{\tau}_n$ fixed,

$$d\ln u_n = \sum_{i=1}^{N} c_{ni}\,d\ln w_i \;-\; \sum_{i=1}^{N} s_{ni}\,d\ln p_i. \tag{B.15}$$

Using $p_{ni} = \tau_{ni}p_i$ with fixed τ_{ni}, we have $d\ln p_{ni} = d\ln p_i$. Substituting (B.12)–(B.14) into (B.15), and writing $d\ln \mathbf{E}^T = \mathbf{g}$, gives

$$d\ln u_n = \mathbf{c}_n\,d\ln \mathbf{w} - \mathbf{s}_n\,d\ln \mathbf{p} = \mathbf{c}_n\,\mathbf{\Pi}\,\mathbf{g} - \beta\,\mathbf{s}_n\,\mathbf{\Pi}\,\mathbf{g} = \big[\mathbf{c}_n - \beta\,\mathbf{s}_n\big]\,\mathbf{\Pi}\,\mathbf{g}.$$

That is,

$$d\ln u_n \;=\; \big[\mathbf{c}_n - \beta\,\mathbf{s}_n\big]\,(I-\mathbf{M})^{-1}\,\mathbf{D}_T\,\mathbf{g}, \qquad \mathbf{g} = d\ln \mathbf{E}^T. \tag{B.16}$$

Equation (B.16) is the closed–form GE mapping from proportional tourism shocks to first–order welfare changes for residents of block n. It corresponds to equation (T.4) in the main text.

Sanity checks. Two simple cases illustrate the structure of (B.16).

- *No spillovers.* If $\mathbf{M} = \mathbf{0}$ (no income–spending feedback), then $\mathbf{\Pi} = \mathbf{D}_T$ and the GE system collapses to partial equilibrium: wage and price responses depend only on local tourism shares, and welfare changes are

  $$d\ln u_n = (\mathbf{c}_n - \beta\mathbf{s}_n)\,\mathbf{D}_T\,\mathbf{g},$$

  with no cross–block spillovers.

- *Common tourism shock.* If the proportional tourism shock is common across blocks, $g_i \equiv \kappa$, then

  $$d\ln \mathbf{p} = \beta\,\kappa\,\mathbf{\Pi}\mathbf{1}, \qquad d\ln u_n = \kappa\,(\mathbf{c}_n - \beta\mathbf{s}_n)\,\mathbf{\Pi}\mathbf{1}.$$

  Residents value the percolated common index $\mathbf{\Pi}\mathbf{1}$ with their net incidence vector $(\mathbf{c}_n - \beta\mathbf{s}_n)$, as emphasized in Section **??**.

Relation to the structural hat–algebra model. Appendix E shows that linearizing the nonlinear hat–algebra system around the baseline equilibrium ($\hat{x} \equiv 1$) recovers exactly the GE system (B.12)–(B.16): the matrix $\mathbf{M}$ collects the one–round income–spending feedback implied by baseline flows, $\mathbf{D}_T$ collects baseline tourism intensities, and $\mathbf{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T$ is the structural percolation operator that appears both in the reduced–form theory and in the full quantitative model.

C Appendix C: Causal Inference in Urban Networks

This appendix formalizes the causal-exposure logic summarized in Section 4.1. We show that: (i) in our general-equilibrium (GE) setting, naive shift–share exposure violates the Stable Unit Treatment Value Assumption (SUTVA); (ii) the correct model-consistent exposure is the *percolated exposure* that uses the percolation operator $\mathbf{\Pi}$ defined in Appendix B; (iii) when shocks are as-good-as-random conditional on predetermined weights, the percolated exposure satisfies the orthogonality conditions underlying shock-based identification, including the recentering logic of **?**.

C.1 GE Structure and Shock Transmission

We begin by summarizing the linear GE transmission derived in Appendix B. Let

$$\mathbf{g} \equiv \mathrm{d}\ln \mathbf{E}^T = (\mathrm{d}\ln E_1^T, \ldots, \mathrm{d}\ln E_N^T)' \in \mathbb{R}^N$$

denote the vector of proportional tourism-expenditure shocks across blocks. Under the assumptions of Appendix B (short-run Cobb–Douglas technology, fixed shares, and $\rho(\mathbf{M}) < 1$), the wage and price responses satisfy

$$\mathrm{d}\ln \mathbf{w} = \mathbf{\Pi}\,\mathbf{g}, \qquad \mathrm{d}\ln \mathbf{p} = \beta\,\mathbf{\Pi}\,\mathbf{g}, \tag{C.1}$$

where

$$\mathbf{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T \in \mathbb{R}^{N\times N}$$

is the percolation operator defined in (B.21), with $\mathbf{M} = D_y^{-1} S D_v C$ the one-round income–spending propagation matrix and $\mathbf{D}_T = \mathrm{diag}(E_i^T/y_i)$ the tourism-intensity diagonal.

From Appendix A, the welfare of resident n reacts to wage and price changes according to

$$\mathrm{d}\ln u_n = \sum_{i=1}^{N} c_{ni}\,\mathrm{d}\ln w_i - \sum_{i=1}^{N} s_{ni}\,\mathrm{d}\ln p_i, \tag{C.2}$$

where $\mathbf{c}_n = (c_{n1}, \ldots, c_{nN})$ and $\mathbf{s}_n = (s_{n1}, \ldots, s_{nN})$ are baseline income and expenditure share rows satisfying $\sum_i c_{ni} = \sum_i s_{ni} = 1$. Substituting (C.1) into (C.2) gives

$$\mathrm{d}\ln u_n = \big[\mathbf{c}_n - \beta\,\mathbf{s}_n\big]\mathbf{\Pi}\,\mathbf{g} + \varepsilon_n, \tag{C.3}$$

where ε_n collects approximation errors, measurement error, and any other small residual terms. Equation (C.3) is the structural GE outcome equation for a welfare-like outcome at origin n.

C.2 Naive Shift–Share Exposure and SUTVA Violation

A standard shift–share exposure at origin n is

$$z_n = \sum_{i=1}^{N} s_{ni}\, g_i = \mathbf{s}_n\,\mathbf{g}, \tag{C.4}$$

which weights shocks by baseline expenditure shares. Implicitly, this formula assumes a SUTVA environment in which the outcome at n depends only on its own scalar exposure index z_n, i.e.

$$Y_n(\mathbf{g}) = Y_n(z_n),$$

where $Y_n(\cdot)$ is the potential outcome at n.

In our GE setting, however, outcomes depend on the full shock vector $\mathbf{g}$ through the percolation operator:

$$\mathrm{d}\ln u_n = \big[\mathbf{c}_n - \beta\,\mathbf{s}_n\big]\mathbf{\Pi}\,\mathbf{g} + \varepsilon_n.$$

Except in knife-edge cases where the loading vector $(\mathbf{c}_n - \beta\,\mathbf{s}_n)\mathbf{\Pi}$ happens to be collinear with $\mathbf{s}_n$, the outcome cannot be written as a function of z_n alone. Hence SUTVA fails generically for the naive exposure z_n.

To see how this affects linear regression, suppose we project the outcome on z_n in a large-sample OLS regression

$$y_n = \alpha + \gamma^{\mathrm{OLS}} z_n + \eta_n,$$

where y_n is a realized outcome (e.g., observed log welfare change, or a price or income response) and η_n is a residual. For simplicity, condition on all predetermined weights

$$\mathcal{W} \equiv \{\mathbf{M}, \mathbf{D}_T, (\mathbf{c}_n, \mathbf{s}_n)_n\},$$

and assume that shocks have zero mean conditional on weights,

$$\mathbb{E}[\mathbf{g} \mid \mathcal{W}] = \mathbf{0},$$

and are uncorrelated with the residual conditional on weights,

$$\mathbb{E}[\mathbf{g}\,\eta_n \mid \mathcal{W}] = \mathbf{0}.$$

Let $\mathbf{\Sigma}_g \equiv \mathbb{E}[\mathbf{g}\mathbf{g}' \mid \mathcal{W}]$ denote the conditional covariance matrix of shocks. Plugging the structural form (C.3) into the covariance formulas, the OLS slope converges in probability to

$$\gamma_n^{\mathrm{OLS}} = \frac{\mathrm{Cov}(z_n, y_n \mid \mathcal{W})}{\mathrm{Var}(z_n \mid \mathcal{W})} = \frac{\mathbf{s}_n\,\mathbf{\Sigma}_g\,\mathbf{\Pi}'(\mathbf{c}_n' - \beta\,\mathbf{s}_n')}{\mathbf{s}_n\,\mathbf{\Sigma}_g\,\mathbf{s}_n'}. \tag{C.5}$$

In general, this is *not* a structural elasticity of interest (e.g., it is not simply a component of $(\mathbf{c}_n - \beta\,\mathbf{s}_n)\mathbf{\Pi}$), because the numerator aggregates spillovers across all blocks according to $\mathbf{\Pi}'(\mathbf{c}_n' - \beta\,\mathbf{s}_n')$ but the denominator uses only the naive shares $\mathbf{s}_n$.

Similarly, if we attempt to use z_n as an instrument for a component of the shock vector (say g_n) in a scalar IV regression $y_n = \alpha + \beta^{\mathrm{IV}} x_n + \eta_n$ with $x_n = g_n$, the 2SLS probability limit becomes

$$\beta_n^{\mathrm{IV,naive}} = \frac{\mathrm{Cov}(z_n, y_n \mid \mathcal{W})}{\mathrm{Cov}(z_n, x_n \mid \mathcal{W})} = \frac{\mathbf{s}_n\,\mathbf{\Sigma}_g\,\mathbf{\Pi}'(\mathbf{c}_n' - \beta\,\mathbf{s}_n')}{\mathbf{s}_n\,\mathbf{\Sigma}_g\,\mathbf{e}_n}, \tag{C.6}$$

with $\mathbf{e}_n$ the nth standard basis vector. Again, the numerator mixes the full GE spillover pattern into a single scalar, so $\beta_n^{\mathrm{IV,naive}}$ is generally a biased object.

In short, naive shift–share exposure is not structural in this environment: the GE mapping makes the outcome depend on the full vector of shocks, and the scalar z_n does not summarize this dependence in a way that aligns with SUTVA.

C.3 Percolated Exposure and Identification

The GE solution (C.1) suggests the correct exposure for price-related outcomes: the log price at block i responds as

$$\mathrm{d}\ln p_i = \beta\,[\mathbf{\Pi g}]_i,$$

so the relevant "shock" entering a price equation for p_i is $[\mathbf{\Pi g}]_i$. Analogously, for tract-level outcomes indexed by residence n and driven by prices and wages, the outcome depends on

$$\mathbf{\Pi g} \quad \text{through} \quad (\mathbf{c}_n - \beta \mathbf{s}_n)\mathbf{\Pi g}.$$

This motivates the *percolated exposure* at origin n:

$$\widetilde{z}_n \equiv \mathbf{s}_n\,\mathbf{\Pi}\,\mathbf{g} = \sum_{i=1}^{N}\sum_{j=1}^{N} s_{ni}\,\Pi_{ij}\,g_j. \tag{C.7}$$

Intuitively, $\mathbf{\Pi g}$ is the vector of GE effects of the shock on (log) wages or prices; $\mathbf{s}_n$ averages those effects across the places where n shops. The mapping $(\mathbf{s}_n, \mathbf{\Pi}, \mathbf{g}) \mapsto \widetilde{z}_n$ is thus structurally aligned with the underlying GE model.

Orthogonality with shocks. Assume now that shocks are mean-zero and uncorrelated with the structural residual conditional on the weights:

$$\mathbb{E}[\mathbf{g} \mid \mathcal{W}] = \mathbf{0}, \qquad \mathbb{E}[\mathbf{g}\,\varepsilon_n \mid \mathcal{W}] = \mathbf{0}. \tag{C.8}$$

Using the law of iterated expectations,

$$\mathbb{E}[\widetilde{z}_n\,\varepsilon_n \mid \mathcal{W}] = \mathbf{s}_n\mathbf{\Pi}\,\mathbb{E}[\mathbf{g}\,\varepsilon_n \mid \mathcal{W}] = 0,$$

so $\widetilde{z}_n$ is orthogonal to the structural error term ε_n under (C.8). This orthogonality is exactly what is needed for shock-based identification in the spirit of **?** and **?**: exposure is a linear functional of shocks, and shocks are exogenous given weights.

Price equations. Equation (C.1) implies a directly estimable relationship between prices and percolated shocks:

$$\ln p_{it} = \delta_i + \delta_t + \beta\,[\mathbf{\Pi g}_t]_i + u_{it}, \tag{C.9}$$

where $\mathbf{g}_t$ is the vector of shocks at time t, δ_i and δ_t are location and time fixed effects, and u_{it} is a residual. If $\mathbf{\Pi}$ were known, β would be identified by OLS under $\mathbb{E}([\mathbf{\Pi g}_t]_i u_{it} \mid \mathcal{W}) = 0$. In practice, we estimate $\widehat{\mathbf{M}}$ and $\widehat{\mathbf{D}}_T$ from data and construct

$$\widehat{\mathbf{\Pi}} = (I - \widehat{\mathbf{M}})^{-1}\widehat{\mathbf{D}}_T,$$

treating $\widehat{\mathbf{\Pi}}\mathbf{g}_t$ as a regressor (or instrumenting it with leave-out or simulated shocks). Under standard regularity conditions and consistency of $\widehat{\mathbf{\Pi}}$, $\widehat{\beta}$ is a consistent estimator of the pass-through parameter.

Welfare. For welfare, identification proceeds by combining an estimate of β with the percolation operator and the household shares:

$$\widehat{\mathrm{d}\ln u_n} = \left[\mathbf{c}_n - \widehat{\beta}\,\mathbf{s}_n\right]\widehat{\mathbf{\Pi}}\,\mathbf{g}. \tag{C.10}$$

This expression embeds all GE spillovers via $(I - \mathbf{M})^{-1}$ and delivers the welfare mapping reported in Section 4.4. It is the empirical analogue of (C.3).

C.4 Truncation, Approximation, and Error Bounds

In finite samples and large cities, it may be convenient to approximate $(I - \mathbf{M})^{-1}$ with a truncated Neumann series. For any matrix norm $\|\cdot\|$ that is submultiplicative and subordinate, and with $\|\mathbf{M}\| < 1$, the exact percolation operator satisfies

$$\mathbf{\Pi} = \sum_{k=0}^{\infty} \mathbf{M}^k \mathbf{D}_T,$$

and we can approximate it by

$$\mathbf{\Pi}^{(K)} \equiv \sum_{k=0}^{K} \mathbf{M}^k \mathbf{D}_T.$$

A standard geometric-series bound gives

$$\left\|\mathbf{\Pi} - \mathbf{\Pi}^{(K)}\right\| \leq \frac{\|\mathbf{M}\|^{K+1}}{1 - \|\mathbf{M}\|} \|\mathbf{D}_T\|. \tag{C.11}$$

An analogous bound holds for $\widehat{\mathbf{\Pi}}$ and its truncated version. This bound directly translates into approximation error bounds for percolated exposure measures and for any regressors or instruments constructed from them.

C.5 Recentered Exposure for Observational Shocks

The discussion so far assumed that shocks have zero mean conditional on weights, $\mathbb{E}[\mathbf{g} \mid \mathcal{W}] = \mathbf{0}$. When shocks are observational rather than randomized, this assumption may fail. Following **?**, we can restore orthogonality by recentering the exposure.

Define the expected percolated exposure

$$\widetilde{\mu}_n \equiv \mathbb{E}[\widetilde{z}_n \mid \mathcal{W}] = \mathbf{s}_n \mathbf{\Pi} \mathbb{E}[\mathbf{g} \mid \mathcal{W}]. \tag{C.12}$$

We can compute $\widetilde{\mu}_n$ from the assignment mechanism for shocks, for example by simulating random permutations of tourist-group inflows and averaging their percolated exposures.

The recentered exposure is then

$$\widetilde{z}_n^{\mathrm{rc}} \equiv \widetilde{z}_n - \widetilde{\mu}_n. \tag{C.13}$$

By construction,

$$\mathbb{E}[\widetilde{z}_n^{\mathrm{rc}} \mid \mathcal{W}] = 0,$$

and, under the maintained exogeneity condition

$$\mathbb{E}[\mathbf{g}\, \varepsilon_n \mid \mathcal{W}] = \mathbf{0},$$

we have

$$\mathbb{E}[\widetilde{z}_n^{\mathrm{rc}} \varepsilon_n \mid \mathcal{W}] = \mathbb{E}[\widetilde{z}_n \varepsilon_n \mid \mathcal{W}] - \widetilde{\mu}_n \mathbb{E}[\varepsilon_n \mid \mathcal{W}] = 0,$$

so the recentered percolated exposure is orthogonal to the structural error term. This is the GE analogue of the recentering argument in **?**, applied to percolated exposure in our urban network setting.

C.6 Sanity Checks and Special Cases

Two special cases help clarify the relationship between naive and percolated exposure:

- **No GE spillovers.** If $\mathbf{M} = \mathbf{0}$ (no income–spending feedback), then $\mathbf{\Pi} = \mathbf{D}_T$ and

$$\widetilde{z}_n = \mathbf{s}_n \mathbf{D}_T \mathbf{g} = \sum_{i=1}^{N} s_{ni} (E_i^T / y_i)\, g_i.$$

  In this case, the outcome depends only on a diagonal weighting of shocks, and naive exposure reduces to a special case of percolated exposure with appropriate rescaling.

- **Common shock.** If $g_i \equiv \kappa$ for all i, then

$$\mathrm{d} \ln \mathbf{p} = \beta\, \kappa\, \mathbf{\Pi 1}, \qquad \mathrm{d} \ln u_n = \kappa\, (\mathbf{c}_n - \beta \mathbf{s}_n)\, \mathbf{\Pi 1},$$

  so all heterogeneity in outcomes reflects heterogeneity in net incidence vectors $(\mathbf{c}_n - \beta \mathbf{s}_n)$ and in the percolated common index $\mathbf{\Pi 1}$. In this knife-edge case, naive exposure $z_n = \kappa \sum_i s_{ni}$ is constant across n, so it carries no cross-sectional variation, whereas percolated exposure $\widetilde{z}_n = \kappa\, \mathbf{s}_n \mathbf{\Pi 1}$ captures heterogeneity induced by network percolation.

Overall, the percolation operator $\mathbf{\Pi}$ plays a dual role: it maps shocks into wages and prices (Appendix B), and it defines the correct causal exposure in this non-SUTVA environment. The empirical implementation in Section **??** proceeds by estimating $\widehat{\mathbf{M}}$ and $\widehat{\mathbf{D}}_T$, forming $\widehat{\mathbf{\Pi}}$, and constructing percolated exposures $\widehat{\widetilde{z}}_n = \mathbf{s}_n \widehat{\mathbf{\Pi}}\, \mathbf{g}$ (and their recentered versions) as instruments and regressors.

D Appendix D: Tourist demand, GE elasticities, and optimal tourist taxation

This appendix derives the tax objects used in Section 5.2. The logic follows the same sequence as in the main text. We begin from the welfare mapping developed in Sections 2.1–2.2, then endogenize tourist expenditure responses to block-specific taxes, and finally derive the nonlinear Ramsey condition together with its linearization around the baseline.

D.1 Welfare and GE pass-through

The starting point is the short-run mapping from tourist expenditure shocks into wages, prices, and welfare. Let $\mathbf{E}^T = (E_1^T, \ldots, E_N^T)'$ denote tourist expenditure across blocks, measured at producer prices. From the general-equilibrium system in the main text,

$$d \ln \mathbf{w} = \mathbf{\Pi}\, d \ln \mathbf{E}^T, \qquad d \ln \mathbf{p} = \beta\, \mathbf{\Pi}\, d \ln \mathbf{E}^T, \tag{D.1}$$

where $\mathbf{\Pi} = (I - \mathbf{M})^{-1} \mathbf{D}_T$ is the percolation operator, $\mathbf{M}$ is the one-round income–spending propagation matrix, and $\mathbf{D}_T = \mathrm{diag}(E_i^T / y_i)$ is the tourism-intensity matrix.

For resident n, the welfare envelope from equation (T.1) in the main text implies

$$d \ln u_n = (\mathbf{c}_n - \beta\, \mathbf{s}_n)\, \mathbf{\Pi}\, d \ln \mathbf{E}^T, \tag{D.2}$$

where $\mathbf{c}_n$ is the vector of labor-income shares across workplaces and $\mathbf{s}_n$ is the vector of expenditure shares across shopping destinations. Aggregating with population weights λ_n, define

$$\boldsymbol{\psi} \equiv \sum_n \lambda_n (\mathbf{c}_n - \beta\, \mathbf{s}_n), \tag{D.3}$$

so that the incidence component of aggregate welfare satisfies

$$dW^{\mathrm{inc}} = \boldsymbol{\psi}\, \mathbf{\Pi}\, d \ln \mathbf{E}^T. \tag{D.4}$$

The row vector $\boldsymbol{\psi}$ is the citywide net-incidence vector: it summarizes which locations matter for welfare once income exposure and price exposure are combined.

D.2 Tourist taxes, gross payments, and the tax base

We now introduce taxes on tourist purchases. Let $\tau_i \geq 0$ denote the ad valorem tax levied on tourist purchases in block i, and define the associated log tax wedge

$$\theta_i \equiv \ln(1+\tau_i), \qquad \boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)'. \tag{D.5}$$

Tourists therefore face the tax-inclusive price

$$P_i^T(\boldsymbol{\theta}) = e^{\theta_i} p_i(\boldsymbol{\theta}). \tag{D.6}$$

It is useful to distinguish between two spending concepts. Let $E_i^T(\boldsymbol{\theta})$ denote tourist expenditure in block i measured at producer prices. This is the relevant tax base. Let $X_i^T(\boldsymbol{\theta})$ denote gross tourist payments in block i, inclusive of the tax wedge. These objects satisfy

$$X_i^T(\boldsymbol{\theta}) = (1+\tau_i) E_i^T(\boldsymbol{\theta}) = e^{\theta_i} E_i^T(\boldsymbol{\theta}). \tag{D.7}$$

Equivalently,

$$\ln \mathbf{E}^T(\boldsymbol{\theta}) = \ln \mathbf{X}^T(\boldsymbol{\theta}) - \boldsymbol{\theta}. \tag{D.8}$$

This accounting identity is important. Tourist demand is naturally written in terms of the gross payments $\mathbf{X}^T$, while revenue is collected on the producer-price tax base $\mathbf{E}^T$.

D.3 Tourist demand and the GE elasticity of the tax base

Tourist demand follows a nested CES system. At the block level, the relevant local demand object is the matrix

$$\mathbf{H}(\boldsymbol{\theta}) \equiv -\frac{\partial \ln \mathbf{X}^T(\boldsymbol{\theta})}{\partial \ln \mathbf{P}^T(\boldsymbol{\theta})'}. \tag{D.9}$$

The (i,j) entry of $\mathbf{H}(\boldsymbol{\theta})$ measures how a change in the tourist-facing price in block j changes gross tourist payments in block i. In the one-sector nested CES case used for intuition in the text, this matrix has the closed form

$$H_{ii}(\boldsymbol{\theta}) = (\sigma_T - 1)(1 - \alpha_i(\boldsymbol{\theta})) + \eta^{\text{out}} \alpha_i(\boldsymbol{\theta}), \tag{D.10}$$

$$H_{ij}(\boldsymbol{\theta}) = -(\sigma_T - 1 - \eta^{\text{out}}) \alpha_j(\boldsymbol{\theta}), \qquad j \neq i, \tag{D.11}$$

where $\alpha_i(\boldsymbol{\theta}) = X_i^T(\boldsymbol{\theta}) / \sum_k X_k^T(\boldsymbol{\theta})$ is the current tourist payment share of block i, σ_T is the within-city elasticity of substitution, and η^{out} is the elasticity governing diversion away from the city. The own terms combine within-city substitution and exit from the city, while the off-diagonal terms capture spatial reallocation across blocks.

To translate this demand system into tax-base elasticities, we also need the GE pass-through from the tax base to producer prices. Define

$$\mathbf{B}(\boldsymbol{\theta}) \equiv \frac{\partial \ln \mathbf{p}(\boldsymbol{\theta})}{\partial \ln \mathbf{E}^T(\boldsymbol{\theta})'}. \tag{D.12}$$

Under the short-run linear GE closure developed in the main text, this reduces at the baseline to

$$\mathbf{B}(\mathbf{0}) = \beta \, \boldsymbol{\Pi}. \tag{D.13}$$

We can now derive the GE elasticity of the tax base. From (D.8),

$$d \ln \mathbf{E}^T = d \ln \mathbf{X}^T - d\boldsymbol{\theta}.$$

From the tourist demand system,

$$d \ln \mathbf{X}^T = -\mathbf{H}(\boldsymbol{\theta}) \, d \ln \mathbf{P}^T.$$

And because tourist-facing prices equal the tax wedge times producer prices,

$$d\ln \mathbf{P}^T = d\boldsymbol{\theta} + d\ln \mathbf{p}.$$

Using $d\ln \mathbf{p} = \mathbf{B}(\boldsymbol{\theta})\, d\ln \mathbf{E}^T$, we obtain

$$d\ln \mathbf{E}^T = -\,\mathbf{H}(\boldsymbol{\theta})\big(d\boldsymbol{\theta} + \mathbf{B}(\boldsymbol{\theta})\, d\ln \mathbf{E}^T\big) - d\boldsymbol{\theta}.$$

Rearranging yields

$$\big(I + \mathbf{H}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})\big)\, d\ln \mathbf{E}^T = -\big(\mathbf{H}(\boldsymbol{\theta}) + I\big)\, d\boldsymbol{\theta}.$$

This motivates the central tax object in the main text:

$$\mathcal{E}(\boldsymbol{\theta}) \equiv -\,\frac{\partial \ln \mathbf{E}^T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \big(I + \mathbf{H}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})\big)^{-1}\big(\mathbf{H}(\boldsymbol{\theta}) + I\big). \tag{D.14}$$

The extra $+I$ term is mechanical. It reflects the fact that the tax base is measured at producer prices, so that a tax increase reduces the base not only through tourist demand but also through the accounting wedge between gross payments and pre-tax expenditure.

Evaluated at the baseline and using $\mathbf{B}(\mathbf{0}) = \beta\, \boldsymbol{\Pi}$, this becomes

$$\mathcal{E}(\mathbf{0}) = \big(I + \mathbf{H}(\mathbf{0})\beta\, \boldsymbol{\Pi}\big)^{-1}\big(\mathbf{H}(\mathbf{0}) + I\big). \tag{D.15}$$

D.4 Revenue, transfers, and the nonlinear Ramsey condition

Revenue from tourist taxes is

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{N} \tau_i(\theta_i)\, E_i^T(\boldsymbol{\theta}), \qquad \tau_i(\theta_i) = e^{\theta_i} - 1. \tag{D.16}$$

This revenue is rebated lump-sum to residents. Let $\Lambda = \sum_n \lambda_n$ denote the total population weight, and let

$$T(\boldsymbol{\theta}) = \frac{R(\boldsymbol{\theta})}{\Lambda} \tag{D.17}$$

denote the per-capita transfer. The marginal social value of one additional unit of transfer is

$$m_T(\boldsymbol{\theta}) \equiv \sum_n \lambda_n \frac{1}{v_n(\boldsymbol{\theta}) + T(\boldsymbol{\theta})}. \tag{D.18}$$

To separate the welfare effect of changing tourism from the welfare effect of rebating revenue, define the incidence vector $\kappa(\boldsymbol{\theta})$ by

$$\kappa(\boldsymbol{\theta})' \equiv \frac{\partial W^{\text{inc}}(\boldsymbol{\theta})}{\partial \ln \mathbf{E}^T(\boldsymbol{\theta})'}. \tag{D.19}$$

The ith element of $\kappa(\boldsymbol{\theta})$ measures how a one-percent increase in tourist expenditure in block i affects resident welfare, holding the tax vector fixed. At the baseline under the linear GE closure,

$$\kappa(\mathbf{0})' = \boldsymbol{\psi}\, \boldsymbol{\Pi}. \tag{D.20}$$

Differentiating revenue and using the definition of $\mathcal{E}(\boldsymbol{\theta})$ gives

$$\nabla_{\boldsymbol{\theta}} R(\boldsymbol{\theta}) = (\mathbf{1} + \boldsymbol{\tau}) \odot \mathbf{E}^T(\boldsymbol{\theta}) - \mathcal{E}(\boldsymbol{\theta})'\big(\boldsymbol{\tau} \odot \mathbf{E}^T(\boldsymbol{\theta})\big). \tag{D.21}$$

The first term is the direct gain from raising the tax wedge. The second term is the erosion of the tax base due to tourist reallocation and diversion.

Combining incidence and transfers, the welfare gradient is

$$\nabla_{\boldsymbol{\theta}} W(\boldsymbol{\theta}) = -\,\mathcal{E}(\boldsymbol{\theta})'\kappa(\boldsymbol{\theta}) + \frac{m_T(\boldsymbol{\theta})}{\Lambda}\left[(\mathbf{1}+\boldsymbol{\tau})\odot \mathbf{E}^T(\boldsymbol{\theta}) - \mathcal{E}(\boldsymbol{\theta})'\big(\boldsymbol{\tau}\odot \mathbf{E}^T(\boldsymbol{\theta})\big)\right]. \tag{D.22}$$

An interior optimum therefore satisfies the nonlinear Ramsey condition

$$\operatorname{diag}\big(\mathbf{E}^T(\boldsymbol{\theta}^\star)\big)\big(\mathbf{1}+\boldsymbol{\tau}^\star\big) = \mathcal{E}(\boldsymbol{\theta}^\star)'\big(\boldsymbol{\tau}^\star\odot \mathbf{E}^T(\boldsymbol{\theta}^\star)\big) + \frac{\Lambda}{m_T(\boldsymbol{\theta}^\star)}\mathcal{E}(\boldsymbol{\theta}^\star)'\kappa(\boldsymbol{\theta}^\star). \tag{D.23}$$

This is the exact nonlinear first-order condition used in the quantitative section. The left-hand side is the direct marginal revenue effect, the first term on the right-hand side is the loss of tax base, and the second term is the welfare incidence of shifting tourism across locations.

A useful benchmark sets $\kappa(\boldsymbol{\theta}) = \mathbf{0}$. In that case the planner cares only about revenue, and (D.23) collapses to a revenue-maximizing, or "optimal tariff," condition.

D.5 Linearization and the frozen-elasticity approximation

Evaluating the welfare gradient at the baseline $\boldsymbol{\theta} = \mathbf{0}$ gives

$$\left.\frac{\partial W}{\partial \boldsymbol{\theta}'}\right|_{\mathbf{0}} = -\,\kappa(\mathbf{0})'\mathcal{E}(\mathbf{0}) + \frac{m_T(\mathbf{0})}{\Lambda}\,\mathbf{E}^{T,0\prime}. \tag{D.24}$$

Using (D.15) and (D.20), the linear GE closure implies

$$\left.\frac{\partial W}{\partial \boldsymbol{\theta}'}\right|_{\mathbf{0}} = -\,\boldsymbol{\psi}\,\boldsymbol{\Pi}\big(I + \mathbf{H}(\mathbf{0})\beta\,\boldsymbol{\Pi}\big)^{-1}\big(\mathbf{H}(\mathbf{0}) + I\big) + \frac{\mathbf{E}^{T,0\prime}}{\bar{v}}, \tag{D.25}$$

where the last term uses the baseline approximation $m_T(\mathbf{0})/\Lambda \approx 1/\bar{v}$. This is the small-tax object reported in the main text. It is a marginal ranking of locations: it tells us which blocks should face relatively higher taxes at the margin, but it does not by itself pin down tax levels.

To obtain explicit tax levels, we use a frozen-elasticity approximation. Fix a reference point $\bar{\boldsymbol{\theta}}$, typically the baseline, and treat

$$\mathbf{E}^T(\bar{\boldsymbol{\theta}}), \quad \mathcal{E}(\bar{\boldsymbol{\theta}}), \quad \kappa(\bar{\boldsymbol{\theta}}), \quad m_T(\bar{\boldsymbol{\theta}})$$

as constant. Define the dollar tax-base vector

$$x \equiv \boldsymbol{\tau}\odot \mathbf{E}^T(\bar{\boldsymbol{\theta}}).$$

Then the Ramsey condition (D.23) becomes the linear system

$$\big(I - \mathcal{E}(\bar{\boldsymbol{\theta}})'\big)x = -\mathbf{E}^T(\bar{\boldsymbol{\theta}}) + \frac{\Lambda}{m_T(\bar{\boldsymbol{\theta}})}\mathcal{E}(\bar{\boldsymbol{\theta}})'\kappa(\bar{\boldsymbol{\theta}}), \tag{D.26}$$

with implied tax vector

$$\boldsymbol{\tau}^{\,\text{frozen}} = x \oslash \mathbf{E}^T(\bar{\boldsymbol{\theta}}). \tag{D.27}$$

This approximation is the natural bridge between the linear ranking and the full nonlinear solution: it converts the local GE objects into a one-shot level approximation and provides a convenient initialization for the nonlinear solver.

D.6 Interpretation

The appendix introduces four objects that organize the tax problem. The matrix $\mathbf{H}(\boldsymbol{\theta})$ summarizes how tourists reallocate spending when tourist-facing prices change. The matrix $\mathbf{B}(\boldsymbol{\theta})$ summarizes how changes in tourist expenditure pass through to producer prices. Combining these two gives $\mathcal{E}(\boldsymbol{\theta})$, the GE elasticity of the producer-price tax base with respect to taxes. Finally, the vector $\kappa(\boldsymbol{\theta})$ summarizes whether marginal tourism in each location is socially valuable or socially costly once wages, prices, and network spillovers are taken into account.

Together, these objects yield a unified characterization of tourist taxation. The linear model evaluates them at the baseline and delivers a marginal ranking of blocks. The nonlinear quantitative model recomputes them at each candidate tax vector and therefore determines the equilibrium scale and curvature of the optimal tax schedule.

E Appendix E: Structural General Equilibrium in Hat Algebra

This appendix records the counterfactual-to-baseline (hat–algebra) system for the full static general equilibrium with nested CES demand for residents and tourists, CES commuting, and Cobb–Douglas production. A hat denotes a counterfactual-to-baseline ratio,

$$\hat{x} \equiv \frac{x}{x^0}.$$

Baseline objects carry a superscript 0 and are treated as data; hats are the unknowns in a counterfactual.

Frictions (commuting μ_{ni}, shopping τ_{ni} and τ_{gi}) and amenities (b_{ni}, b_{gi}), fixed factors M_{is}, and productivities A_{is} are kept fixed unless explicitly shocked.

Indices:

- blocks $i, n \in \{1, \ldots, N\}$,
- sectors $s \in \{1, \ldots, S\}$,
- tourist origin groups $g \in \{1, \ldots, G\}$.

Technology uses sectoral labor shares $0 < \beta_s < 1$. Within–sector CES elasticities are $\sigma_s > 1$ for residents and $\sigma_s^T > 1$ for tourists. Across–sector (outer) CES elasticities are $\eta > 0$ for residents and $\eta^T > 0$ for tourists. The commuting elasticity is $\epsilon > 1$.

E.1 Delivered Prices and Price Indices

Let p_{is} be the producer price in block i, sector s. Residents in origin n face delivered prices

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

where $\tau_{ni} \geq 1$ are shopping frictions and b_{ni} are local amenities. Tourists in group g face delivered prices

$$\tilde{p}_{gis} \equiv \tau_{gi}\, b_{gi}\, p_{is},$$

possibly shifted by a tourist tax (introduced explicitly in Appendix **??**). In this appendix we treat frictions and amenities as fixed, so

$$\widehat{\tilde{p}}_{nis} = \frac{\tilde{p}_{nis}}{\tilde{p}^0_{nis}} = \hat{p}_{is}, \qquad \widehat{\tilde{p}}_{gis} = \hat{p}_{is}. \tag{E.1}$$

Within–sector CES price indices. Residents' within–sector price index in origin n and sector s is

$$P_{ns} = \left[\sum_{i=1}^{N} \left(\tilde{p}_{nis} \right)^{1-\sigma_s} \right]^{\frac{1}{1-\sigma_s}}.$$

Define baseline CES weights

$$\omega_{nis}^{P,0} \equiv \frac{(\tilde{p}_{nis}^{0})^{1-\sigma_s}}{\sum_j (\tilde{p}_{njs}^{0})^{1-\sigma_s}}, \qquad \sum_i \omega_{nis}^{P,0} = 1,$$

so the hat of P_{ns} is

$$\hat{P}_{ns} = \left[\sum_{i=1}^{N} \omega_{nis}^{P,0} \, \hat{p}_{is}^{1-\sigma_s} \right]^{\frac{1}{1-\sigma_s}}. \tag{E.2}$$

Similarly, tourists from group g face sectoral price index

$$P_{gs}^{T} = \left[\sum_{i=1}^{N} \left(\tilde{p}_{gis} \right)^{1-\sigma_s^T} \right]^{\frac{1}{1-\sigma_s^T}},$$

with baseline weights

$$\omega_{gis}^{P,T,0} \equiv \frac{(\tilde{p}_{gis}^{0})^{1-\sigma_s^T}}{\sum_j (\tilde{p}_{gjs}^{0})^{1-\sigma_s^T}}, \qquad \sum_i \omega_{gis}^{P,T,0} = 1,$$

and

$$\hat{P}_{gs}^{T} = \left[\sum_{i=1}^{N} \omega_{gis}^{P,T,0} \, \hat{p}_{is}^{1-\sigma_s^T} \right]^{\frac{1}{1-\sigma_s^T}}. \tag{E.3}$$

Across–sector CES price indices. Residents aggregate sectoral prices into a consumption price index

$$G_n = \left[\sum_{s=1}^{S} \alpha_s \, P_{ns}^{1-\eta} \right]^{\frac{1}{1-\eta}},$$

with sector weights $\alpha_s > 0$, $\sum_s \alpha_s = 1$. Define baseline outer weights

$$\omega_{ns}^{G,0} \equiv \frac{\alpha_s (P_{ns}^{0})^{1-\eta}}{\sum_t \alpha_t (P_{nt}^{0})^{1-\eta}}, \qquad \sum_s \omega_{ns}^{G,0} = 1,$$

so the consumption-price hat is

$$\hat{G}_n = \left[\sum_{s=1}^{S} \omega_{ns}^{G,0} \, \hat{P}_{ns}^{1-\eta} \right]^{\frac{1}{1-\eta}}. \tag{E.4}$$

Tourists from group g have city-level price index

$$G_g^{T} = \left[\sum_{s=1}^{S} \alpha_s^{T} (P_{gs}^{T})^{1-\eta^T} \right]^{\frac{1}{1-\eta^T}},$$

with weights α_s^T, and baseline outer weights

$$\omega_{gs}^{G,T,0} \equiv \frac{\alpha_s^T (P_{gs}^{T0})^{\,1-\eta^T}}{\sum_t \alpha_t^T (P_{gt}^{T0})^{\,1-\eta^T}}, \qquad \sum_s \omega_{gs}^{G,T,0} = 1,$$

so

$$\widehat{G_g^T} = \left[\sum_{s=1}^{S} \omega_{gs}^{G,T,0} \, (\hat{P}_{gs}^{T})^{\,1-\eta^T} \right]^{\frac{1}{1-\eta^T}}. \tag{E.5}$$

E.2 Demand Shares in Hats

Across blocks within a sector (residents). Residents' within–sector expenditure shares on (n, i, s) are

$$s_{nis} = \frac{\tilde{p}_{nis}^{1-\sigma_s}}{\sum_j \tilde{p}_{njs}^{1-\sigma_s}}.$$

Using (E.1), their hat satisfies

$$\frac{s_{nis}}{s_{nis}^0} = \frac{\hat{p}_{is}^{1-\sigma_s}}{\sum_j s_{njs}^0 \hat{p}_{js}^{1-\sigma_s}}. \tag{E.6}$$

Across blocks within a sector (tourists). Tourists' within–sector shares are

$$s_{gis}^T = \frac{\tilde{p}_{gis}^{1-\sigma_s^T}}{\sum_j \tilde{p}_{gjs}^{1-\sigma_s^T}},$$

so

$$\frac{s_{gis}^T}{s_{gis}^{T,0}} = \frac{\hat{p}_{is}^{1-\sigma_s^T}}{\sum_j s_{gjs}^{T,0} \hat{p}_{js}^{1-\sigma_s^T}}. \tag{E.7}$$

Across sectors (residents). Residents' across-sector budget shares b_{ns} are

$$b_{ns} = \frac{\alpha_s P_{ns}^{1-\eta}}{G_n^{1-\eta}},$$

so the share hat is

$$\hat{b}_{ns} = \frac{b_{ns}}{b_{ns}^0} = \frac{\hat{P}_{ns}^{1-\eta}}{\sum_t \omega_{nt}^{G,0} \hat{P}_{nt}^{1-\eta}}. \tag{E.8}$$

Across sectors (tourists). Tourists' sectoral budget shares b_{gs}^T satisfy

$$b_{gs}^T = \frac{\alpha_s^T (P_{gs}^T)^{1-\eta^T}}{(G_g^T)^{1-\eta^T}},$$

and the hat is

$$\hat{b}_{gs}^T = \frac{(P_{gs}^T)^{1-\eta^T}/(G_g^T)^{1-\eta^T}}{(P_{gs}^{T0})^{1-\eta^T}/(G_g^{T0})^{1-\eta^T}} = \frac{(\hat{P}_{gs}^T)^{1-\eta^T}}{\sum_t \omega_{gt}^{G,T,0} (\hat{P}_{gt}^T)^{1-\eta^T}}. \tag{E.9}$$

E.3 Commuting, Incomes, and Labor Allocation

Residents in origin n supply labor to blocks according to a CES commuting technology

$$H_n(\{\ell_{ni}\}) = \left[\sum_i \ell_{ni}^{(\epsilon-1)/\epsilon}\right]^{\frac{\epsilon}{\epsilon-1}} = T_n, \qquad \epsilon > 1.$$

The dual income aggregator is

$$J_n = \left[\sum_i \left(\widetilde{w}_{ni}\right)^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, \qquad \widetilde{w}_{ni} \equiv \frac{1}{\mu_{ni}} \sum_s w_{is},$$

where $\mu_{ni} \geq 1$ are commuting frictions and w_{is} are sectoral wages in block i. Disposable income is

$$v_n = T_n J_n.$$

Define baseline commuting weights

$$\omega_{ni}^{J,0} \equiv \frac{\left(\widetilde{w}_{ni}^0\right)^{1-\epsilon}}{\sum_j (\widetilde{w}_{nj}^0)^{1-\epsilon}}, \qquad \sum_i \omega_{ni}^{J,0} = 1,$$

and a block-level effective wage hat

$$\hat{W}_{ni} \equiv \frac{\widetilde{w}_{ni}}{\widetilde{w}_{ni}^0} = \frac{\sum_s w_{is}/\mu_{ni}}{\sum_s w_{is}^0/\mu_{ni}}.$$

Then the income hat is

$$\hat{v}_n = \left[\sum_{i=1}^{N} \omega_{ni}^{J,0}\, \hat{W}_{ni}^{\,1-\epsilon}\right]^{\frac{1}{1-\epsilon}}. \tag{E.10}$$

In many applications sectoral wage movements within a block are similar, so $\hat{W}_{ni} \approx \hat{w}_i$ is a good approximation.

Labor allocation across blocks and sectors. Let L_{is} denote labor used in (i, s). Total labor in (i, s) aggregates residents' CES commuting and sectoral choices. It is convenient to link L_{is} to revenues via the Cobb–Douglas cost share:

$$w_{is} L_{is} = \beta_s\, y_{is},$$

with y_{is} revenue (defined below). The labor hat then follows from

$$\hat{L}_{is} = \frac{\hat{y}_{is}}{\hat{w}_{is}}, \tag{E.11}$$

since β_s is fixed.

E.4 Goods-Market Clearing and Revenues

Output in (i, s) is produced with Cobb–Douglas technology

$$Q_{is} = A_{is}\, L_{is}^{\beta_s}\, M_{is}^{1-\beta_s}, \qquad p_{is} = \frac{w_{is}^{\beta_s} r_{is}^{1-\beta_s}}{A_{is}},$$

under perfect competition, with fixed nonlabor factors M_{is} and rental rates r_{is} in the short run.

Revenue is $y_{is} = p_{is} Q_{is}$. Goods-market clearing in value terms equates revenues to spending by residents and tourists. Let e_{nis}^0 and e_{gis}^0 be baseline expenditures by residents and tourists on (i, s). Define baseline revenue shares

$$\omega_{nis}^{R,0} \equiv \frac{e_{nis}^0}{y_{is}^0}, \qquad \omega_{gis}^{T,0} \equiv \frac{e_{gis}^0}{y_{is}^0}, \qquad \sum_n \omega_{nis}^{R,0} + \sum_g \omega_{gis}^{T,0} = 1.$$

Tourist group g has exogenous budget shock $\hat{E}_g^T$ (e.g. a change in its total spending on the city). Goods-market clearing in hats is

$$\hat{y}_{is} = \sum_n \omega_{nis}^{R,0}\, \hat{v}_n\, \hat{b}_{ns}\, \frac{\hat{p}_{is}^{\,1-\sigma_s}}{\sum_j s_{njs}^0\, \hat{p}_{js}^{\,1-\sigma_s}} + \sum_g \omega_{gis}^{T,0}\, \hat{E}_g^T\, \hat{b}_{gs}^T\, \frac{\hat{p}_{is}^{\,1-\sigma_s^T}}{\sum_j s_{gjs}^{T,0}\, \hat{p}_{js}^{\,1-\sigma_s^T}}. \tag{E.12}$$

E.5 Technology and Labor-Market Identities

Cobb–Douglas unit cost implies a simple link between producer prices and wages:

$$\hat{p}_{is} = \hat{w}_{is}^{\beta_s}. \tag{E.13}$$

Labor payments equal labor's share of revenue both at baseline and in the counterfactual, yielding the convenient identity (E.11),

$$\hat{L}_{is} = \frac{\hat{y}_{is}}{\hat{w}_{is}}.$$

E.6 Equilibrium, Normalization, and Welfare

Unknowns and equilibrium system. The unknown hats are

$$\left\{\hat{p}_{is}, \hat{w}_{is}, \hat{y}_{is}, \hat{L}_{is}\right\}_{i,s}, \qquad \left\{\hat{P}_{ns}, \hat{G}_n, \hat{v}_n\right\}_{n,s},$$

plus tourist-sector indices $\hat{P}^T_{gs}$ and $\widehat{G^T_g}$, together with corresponding share hats. An equilibrium is a collection of hats that simultaneously satisfies:

- delivered-price relations (E.1);
- price index definitions (E.2)–(E.5);
- share equations (E.6)–(E.9);
- income hats (E.10);
- goods-market clearing (E.12);
- technology and labor-market identities (E.11)–(E.13).

Normalization. A monetary numeraire is fixed by choosing one price or price index. A convenient choice is a population-weighted consumption price index:

$$\sum_n \lambda_n \ln \hat{G}_n = 0.$$

Welfare. Resident n's indirect utility is $u_n = v_n/G_n$, so the welfare hat is

$$\hat{u}_n = \frac{\hat{v}_n}{\hat{G}_n}. \tag{E.14}$$

The equivalent variation (EV) in units of baseline income is

$$EV_n = v_n^0(\hat{u}_n - 1).$$

E.7 Small-Shock Linearization and Connection to Appendices B–D

Linearizing the hat system around the baseline $\hat{x} \equiv 1$ yields the small-shock GE mappings used in the main text and Appendices B–C. In particular:

- the log-differential of wages satisfies $d\ln \mathbf{w} = (I - \mathbf{M})^{-1}\mathbf{D}_T\,\mathbf{g}$,
- producer prices satisfy $d\ln \mathbf{p} = \beta\, d\ln \mathbf{w}$,
- resident welfare satisfies $d\ln u_n = [\mathbf{c}_n - \beta\,\mathbf{s}_n](I - \mathbf{M})^{-1}\mathbf{D}_T\,\mathbf{g}$,

where $\mathbf{M}$, $\mathbf{D}_T$, and $\mathbf{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T$ are as defined in Appendix B, and $\mathbf{g}$ collects small tourism shocks. This confirms that the linear analysis in the main text is the first-order approximation to the full nonlinear hat-algebra equilibrium described above.

F Appendix F: Optimal Tourist Taxation in Hat Algebra

This appendix embeds block-specific tourist taxes into the nonlinear hat–algebra equilibrium of Appendix E and characterizes the welfare-based first-order conditions. We then show how linearizing this system around the baseline reproduces the small-tax rule derived in Appendix D.

F.1 Taxes in the Hat–Algebra System

Recall from Appendix E that tourists from group g face delivered prices

$$\tilde{p}_{gis} = \tau_{gi}\, b_{gi}\, (1+\tau_i)\, p_{is},$$

so in hats,

$$\widehat{\tilde{p}}_{gis} = (1+\tau_i)\, \hat{p}_{is}.$$

Residents do not pay the tourist tax, so $\widehat{\tilde{p}}_{nis} = \hat{p}_{is}$.

The only modification relative to the no-tax hat system in Appendix E is that $(1+\tau_i)$ enters:

- tourist within-sector price indices $\hat{P}^T_{gs}$ [equation (E.3)],
- tourist within-sector expenditure shares s^T_{gis} [equation (E.6)],
- tourist across-sector budget shares $\hat{b}^T_{gs}$ [equation (E.8)],
- and tourist contributions to goods-market clearing [equation (E.11)].

The producer-side relations (Cobb–Douglas unit cost, factor payments) remain:

$$\hat{p}_{is} = \hat{w}_{is}^{\beta_s}, \qquad \hat{L}_{is} = \frac{\hat{y}_{is}}{\hat{w}_{is}}.$$

Collect all hats into a vector

$$X = \Big(\{\hat{p}_{is}, \hat{w}_{is}, \hat{y}_{is}, \hat{L}_{is}\}_{i,s}, \{\hat{P}_{ns}, \hat{G}_n, \hat{v}_n\}_{n,s}, \{\hat{P}^T_{gs}, \widehat{G^T_g}\}_{g,s}, \text{share hats}\Big),$$

and write the equilibrium conditions as

$$\mathcal{F}_X\big(X, \boldsymbol{\tau}\big) = \mathbf{0}, \qquad \boldsymbol{\tau} = (\tau_1, \ldots, \tau_N)'. \tag{F.1}$$

For each $\boldsymbol{\tau}$ in a neighborhood of the baseline, there is a solution $X(\boldsymbol{\tau})$.

F.2 Tax Revenue and Transfers

Let E^{T0}_{gis} be baseline tourist spending of group g on variety (i,s). In a counterfactual with taxes $\boldsymbol{\tau}$, tourist spending on (i,s) is

$$E^T_{gis}(\boldsymbol{\tau}) = E^{T0}_g\, \hat{E}^T_g(\boldsymbol{\tau})\, \hat{b}^T_{gs}(\boldsymbol{\tau})\, \frac{[(1+\tau_i)\hat{p}_{is}(\boldsymbol{\tau})]^{1-\sigma^T_s}}{\sum_j s^{T,0}_{gjs}\, [(1+\tau_j)\hat{p}_{js}(\boldsymbol{\tau})]^{1-\sigma^T_s}}.$$

Total tourist spending at block i (across sectors and groups) is

$$E^T_i(\boldsymbol{\tau}) = \sum_{s,g} E^T_{gis}(\boldsymbol{\tau}),$$

so tax revenue at block i is $R_i(\boldsymbol{\tau}) = \tau_i E_i^T(\boldsymbol{\tau})$. Total tax revenue is

$$R(\boldsymbol{\tau}) = \sum_{i=1}^{N} R_i(\boldsymbol{\tau}) = \sum_{i,s,g} \tau_i \, E_{gis}^T(\boldsymbol{\tau}). \tag{F.2}$$

Assume revenue is rebated lump-sum and equally across residents. With population weights $\{\lambda_n\}$, the per-capita transfer is

$$T(\boldsymbol{\tau}) = \frac{R(\boldsymbol{\tau})}{\sum_n \lambda_n}. \tag{F.3}$$

Total income for resident n is $v_n^{\text{tot}} = v_n^{\text{earn}} + T$, with $v_n^{\text{earn}} = v_n^0 \, \hat{v}_n$. A convenient multiplicative representation is

$$\hat{t}_n(\boldsymbol{\tau}) \equiv \frac{v_n^{\text{tot}}}{v_n^{\text{earn}}} = 1 + \frac{T(\boldsymbol{\tau})}{v_n^0 \, \hat{v}_n(\boldsymbol{\tau})}, \tag{F.4}$$

so that exact welfare is

$$\hat{u}_n(\boldsymbol{\tau}) = \frac{\hat{v}_n(\boldsymbol{\tau}) \, \hat{t}_n(\boldsymbol{\tau})}{\hat{G}_n(\boldsymbol{\tau})}. \tag{F.5}$$

F.3 Welfare Objective and Gradient

Define the population-weighted welfare index

$$W(\boldsymbol{\tau}) \equiv \sum_n \lambda_n \ln \hat{u}_n(\boldsymbol{\tau}). \tag{F.6}$$

Using (F.5),

$$\ln \hat{u}_n = \ln \hat{v}_n - \ln \hat{G}_n + \ln \hat{t}_n.$$

By the chain rule and the equilibrium system (F.1), we can write the derivative of $X(\boldsymbol{\tau})$ with respect to taxes using the implicit-function theorem. Let

$$J_X \equiv \frac{\partial \mathcal{F}_X}{\partial X}, \qquad J_\tau \equiv \frac{\partial \mathcal{F}_X}{\partial \boldsymbol{\tau}'},$$

evaluated at $(X(\boldsymbol{\tau}), \boldsymbol{\tau})$. Under regularity (local uniqueness and non-singularity of J_X), the implicit-function theorem implies

$$\frac{\partial X}{\partial \boldsymbol{\tau}'} = -\, J_X^{-1} J_\tau. \tag{F.7}$$

Differentiating $\ln \hat{u}_n$ gives

$$\frac{\partial \ln \hat{u}_n}{\partial \boldsymbol{\tau}'} = \left(\frac{\partial \ln \hat{v}_n}{\partial X} - \frac{\partial \ln \hat{G}_n}{\partial X} + \frac{\partial \ln \hat{t}_n}{\partial X} \right) \frac{\partial X}{\partial \boldsymbol{\tau}'} + \frac{\partial \ln \hat{t}_n}{\partial \boldsymbol{\tau}'}.$$

Aggregating across n leads to the welfare gradient

$$\frac{\partial W}{\partial \boldsymbol{\tau}'} = \underbrace{\sum_n \lambda_n \left[\frac{\partial \ln \hat{v}_n}{\partial X} - \frac{\partial \ln \hat{G}_n}{\partial X} + \frac{\partial \ln \hat{t}_n}{\partial X} \right]}_{\boldsymbol{\Psi}(X)} \cdot (-J_X^{-1} J_\tau) + \sum_n \lambda_n \frac{\partial \ln \hat{t}_n}{\partial \boldsymbol{\tau}'}. \tag{F.8}$$

Here $\boldsymbol{\Psi}(X)$ is a row vector of GE incidence weights evaluated at the equilibrium $X(\boldsymbol{\tau})$.

An interior optimum $\boldsymbol{\tau}^\star$ satisfies the first-order condition

$$\frac{\partial W}{\partial \boldsymbol{\tau}'}\Big|_{\boldsymbol{\tau}^\star} = \mathbf{0}'. \tag{F.9}$$

Equation (F.9) is the exact (nonlinear) Ramsey condition in the structural model.

F.4 Local Approximation and Connection to Appendix D

To relate (F.9) to the small-tax rule in Appendix D, linearize the system around the baseline $\boldsymbol{\tau} = \mathbf{0}$. At $\boldsymbol{\tau} = \mathbf{0}$ we have

$$\hat{v}_n = \hat{G}_n = \hat{t}_n = 1, \quad T(\mathbf{0}) = 0, \quad X(\mathbf{0}) = X^0,$$

and the GE structure reduces to the small-shock model of Appendices B–D. In particular:

- The tax–to–demand mapping can be written as

$$\mathbf{g} \equiv d\ln \mathbf{E}^T = \mathbf{L}\,\boldsymbol{\tau}, \qquad \mathbf{L} = -\big(I + H\beta\boldsymbol{\Pi}\big)^{-1} H,$$

  where H is the nested-CES semi-elasticity matrix from Appendix D, and $\boldsymbol{\Pi} = (I - \mathbf{M})^{-1}\mathbf{D}_T$ is the percolation operator from Appendix B.

- Wages and prices respond as

$$d\ln \mathbf{w} = \boldsymbol{\Pi}\,\mathbf{g}, \qquad d\ln \mathbf{p} = \beta\,\boldsymbol{\Pi}\,\mathbf{g}.$$

- The transfer effect on welfare is captured by

$$\sum_n \lambda_n \frac{\partial \ln \hat{t}_n}{\partial \boldsymbol{\tau}'}\Big|_{\boldsymbol{\tau}=\mathbf{0}} = \frac{(\mathbf{E}^T)'}{\bar{v}}, \qquad \bar{v} \equiv \frac{\sum_n \lambda_n v_n}{\sum_n \lambda_n},$$

  where $\mathbf{E}^T$ collects baseline tourist expenditures across blocks.

At $\boldsymbol{\tau} = \mathbf{0}$, $\boldsymbol{\Psi}(X^0)$ reduces to the citywide net-incidence row vector

$$\boldsymbol{\psi} \equiv \sum_n \lambda_n (\mathbf{c}_n - \beta\,\mathbf{s}_n),$$

with $\mathbf{c}_n$ and $\mathbf{s}_n$ the income and expenditure share rows from Appendix A. Substituting into (F.8) and using the linear structure above yields the local welfare gradient

$$\frac{\partial W}{\partial \boldsymbol{\tau}'}\Big|_{\boldsymbol{\tau}=\mathbf{0}} = \boldsymbol{\psi}\,\boldsymbol{\Pi}\,\mathbf{L} + \frac{(\mathbf{E}^T)'}{\bar{v}} = -\,\boldsymbol{\psi}\,\boldsymbol{\Pi}\,(I + H\beta\boldsymbol{\Pi})^{-1} H + \frac{(\mathbf{E}^T)'}{\bar{v}}.$$

Setting this to zero gives

$$\frac{E_i^T}{\bar{v}} = \Big[\boldsymbol{\psi}\,\boldsymbol{\Pi}\,(I + H\beta\boldsymbol{\Pi})^{-1} H\Big]_i \quad \text{for each } i,$$

and rearranging yields the small-tax Ramsey rule

$$\tau_i^\star = \frac{E_i^T/\bar{v}}{\Big[\boldsymbol{\psi}\,\boldsymbol{\Pi}\,(I + H\beta\boldsymbol{\Pi})^{-1} H\Big]_i},$$

which is exactly equation (D.22) in Appendix D (and equation (T.7) in the main text).

Thus, Appendix D provides the linear (small-tax) approximation to the fully nonlinear optimal-tax problem characterized here; all higher-order GE feedbacks are captured in the Jacobians J_X and J_τ of the structural hat–algebra system.

G Appendix: Empirical

G.1 Summary Statistics

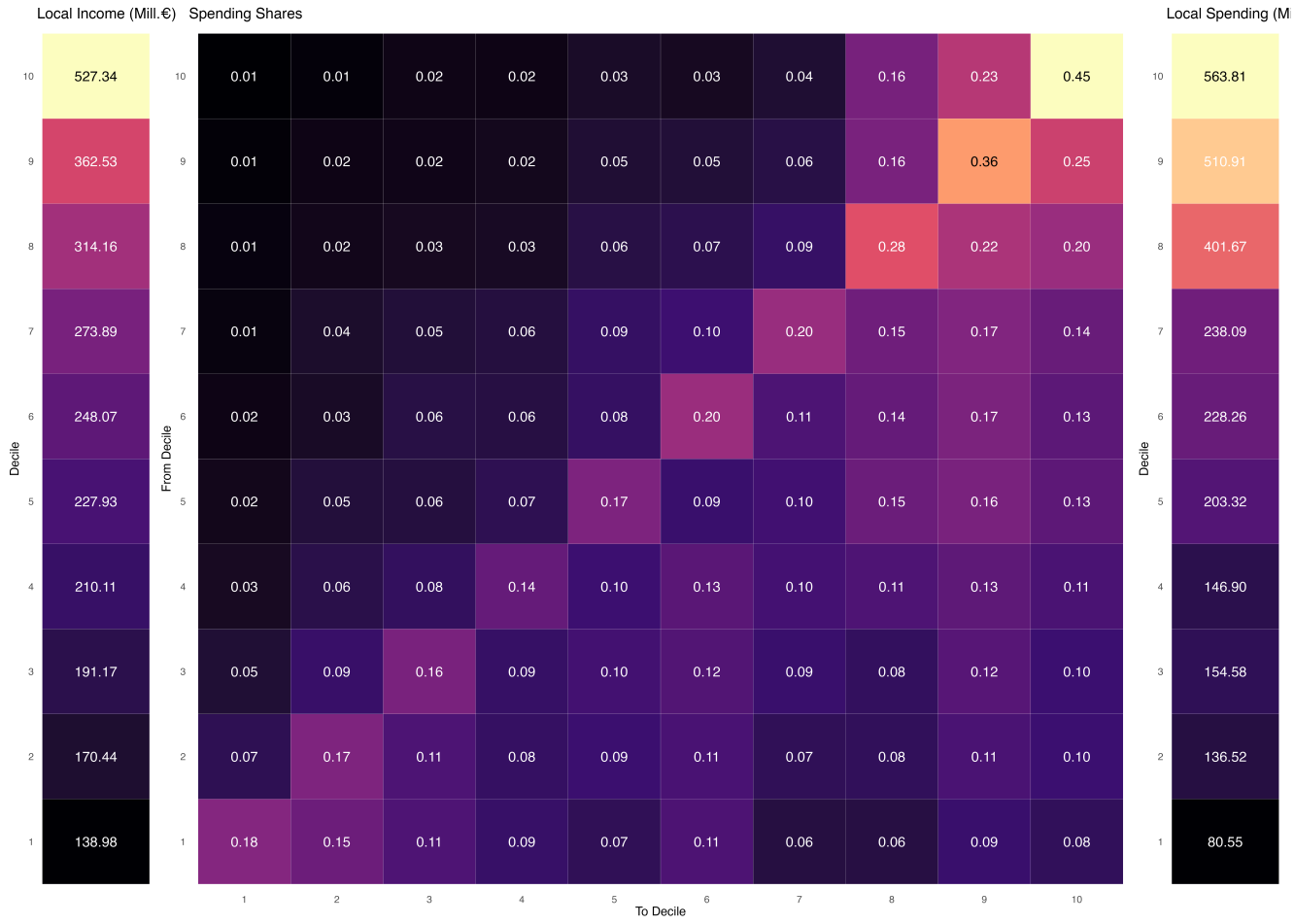
Table A.1. Final Sample Descriptives

	Mean	p10	p50	p90	std
Local income (mil EUR)	30.18	17.87	26.79	45.17	18.01
Spending by residents (mil EUR)	22.66	1.28	10.09	51.35	48.39
Spending by tourists (mil EUR)	3.54	0.06	0.31	5.04	17.80
Shares by local income at residence					
Share spent in Network	0.75	0.71	0.75	0.79	0.04
Share spent out of Network	0.25	0.21	0.25	0.29	0.04
Share in own location	0.06	0.01	0.03	0.17	0.07
Shares by spending in location					
Share of spending by tourists	0.09	0.01	0.04	0.28	0.14

G.1.1 Income and Spending Flows

To illustrate the full spatial structure of local income and spending flows, we group locations into ten deciles based on their average monthly income. The left panel of Figure A.1 shows the geographic distribution of total income. Total income in the top decile exceeds that of the bottom decile by nearly a factor of four. The middle panel quantifies the spending interactions across deciles. The diagonal pattern highlights that much of local spending is concentrated within the same income group, especially for the richest areas, where 45% of income is spent in other top-decile neighborhoods. This pattern of spatially stratified spending reinforces and amplifies the income distribution, leading to a consumption distribution that closely mirrors the underlying disparities in local income, as seen in the right panel.

Figure A.1. From income to spending

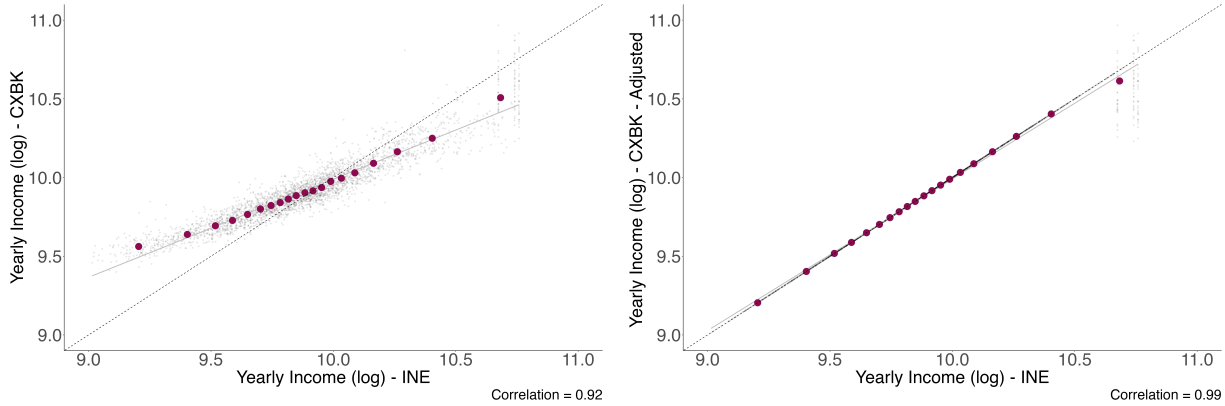


Notes: This figure summarizes the spending network linking income and expenditure across local deciles. The left color bar reports average local income by decile (in million euros), while the right color bar shows local spending. The central heatmap displays the matrix of spending shares across deciles, where each cell represents the share of total spending by residents in income decile i (rows) directed toward businesses in decile j (columns). Darker colors correspond to smaller spending shares, while lighter colors indicate stronger local demand linkages. High values along the diagonal illustrate within-decile spending concentration, whereas off-diagonal entries capture cross-decile propagation of income and spending.

G.2 Data Representativeness

G.2.1 Income Data

Figure A.2. Income: CXBK vs. INE



Notes: Comparison of CXBK and INE Annual Income Estimates by Census Tract. The left panel plots log yearly income from CXBK against INE statistics using raw CXBK data. The right panel shows the same comparison after applying the adjustment procedure described in Section 3.1.2.

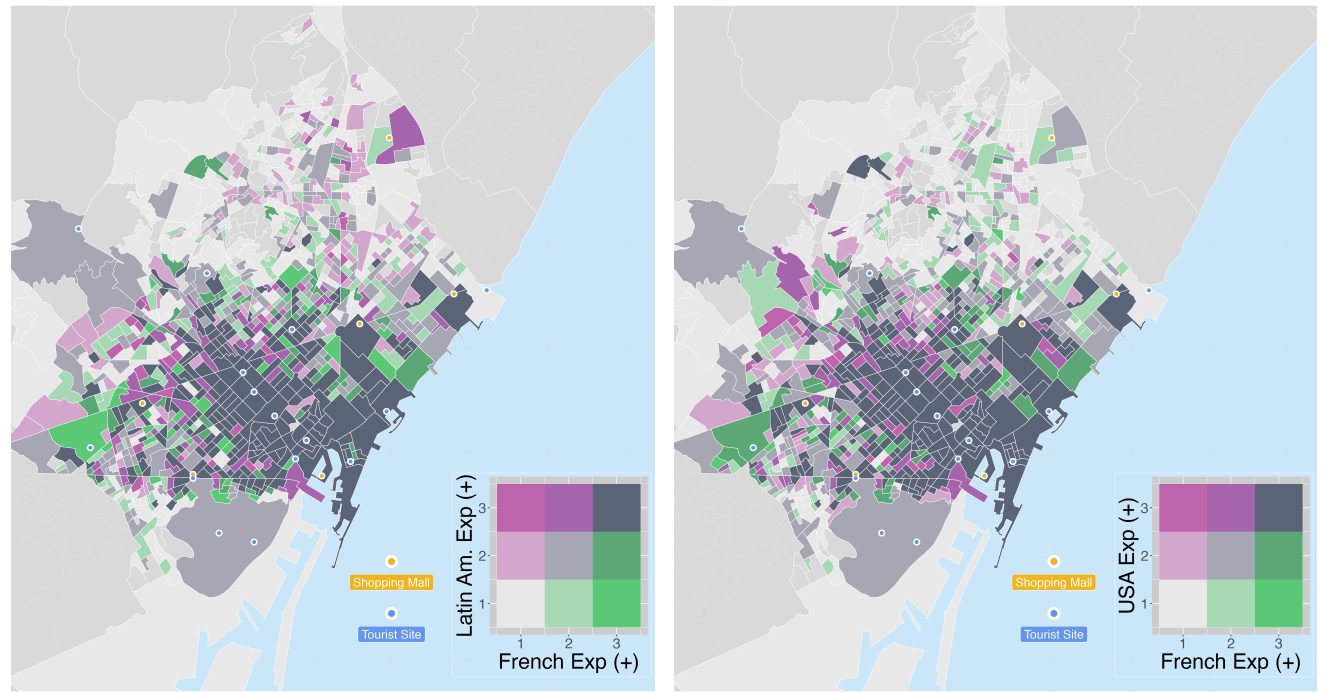
G.3 Supplementary Empirical Patterns

G.3.1 Tourist Spending Across Space for Different Groups

Figure A.3.

Figure A.3. Different tourists have different preferences within the city and year

(a) Tourists' Spending Across Space **(b) Tourists' Spending Across Seasons**



Notes: **(Panel a)** This figure compares the percentage change in Latin American and French tourist expenditure between the tourist high season (July) 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 increase for both groups. **(Panel b)** Same for US and Canada compared to French tourists.

G.4 Income and Price Effects Estimation

G.4.1 Gravity

Table A.2. Gravity Regressions

	Residents Spending Flows				Residents Commuting Flows			
Model:	(1) OLS	(2) OLS	(3) PPML	(4) PPML	(1) OLS	(2) OLS	(3) PPML	(4) PPML
Variables								
Distance (seconds)	-1.708***		-1.883***		-2.644***		-2.105***	
	(0.0056)		(0.0061)		(0.0020)		(0.0042)	
Distance (structural)		-2.456***		-3.666***		-3.945***		-4.597***
		(0.0088)		(0.0174)		(0.0059)		(0.0118)
Fixed-effects								
$n \times t$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$i \times t$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	3,690,214	3,690,214	3,690,214	3,690,214	3,690,214	3,690,214	3,690,214	3,690,214
Pseudo R^2	0.1091	0.1018	0.5764	0.5723	0.7403	0.6869	0.9025	0.9114

Clustered (i,t) and (n,t) standard errors in parentheses.
*Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.*

Notes: All variables are expressed in logs in the OLS models and in levels in the PPML models. Spending and commuting flows are defined in 3.1.3. Distance in seconds is calculated using scraped data from the HERE API, averaging reported durations for both public transit and driving. The structural measure of distance combines travel time, local income, and the cost of consumption trips. See 3.1.5 for details on distance data.

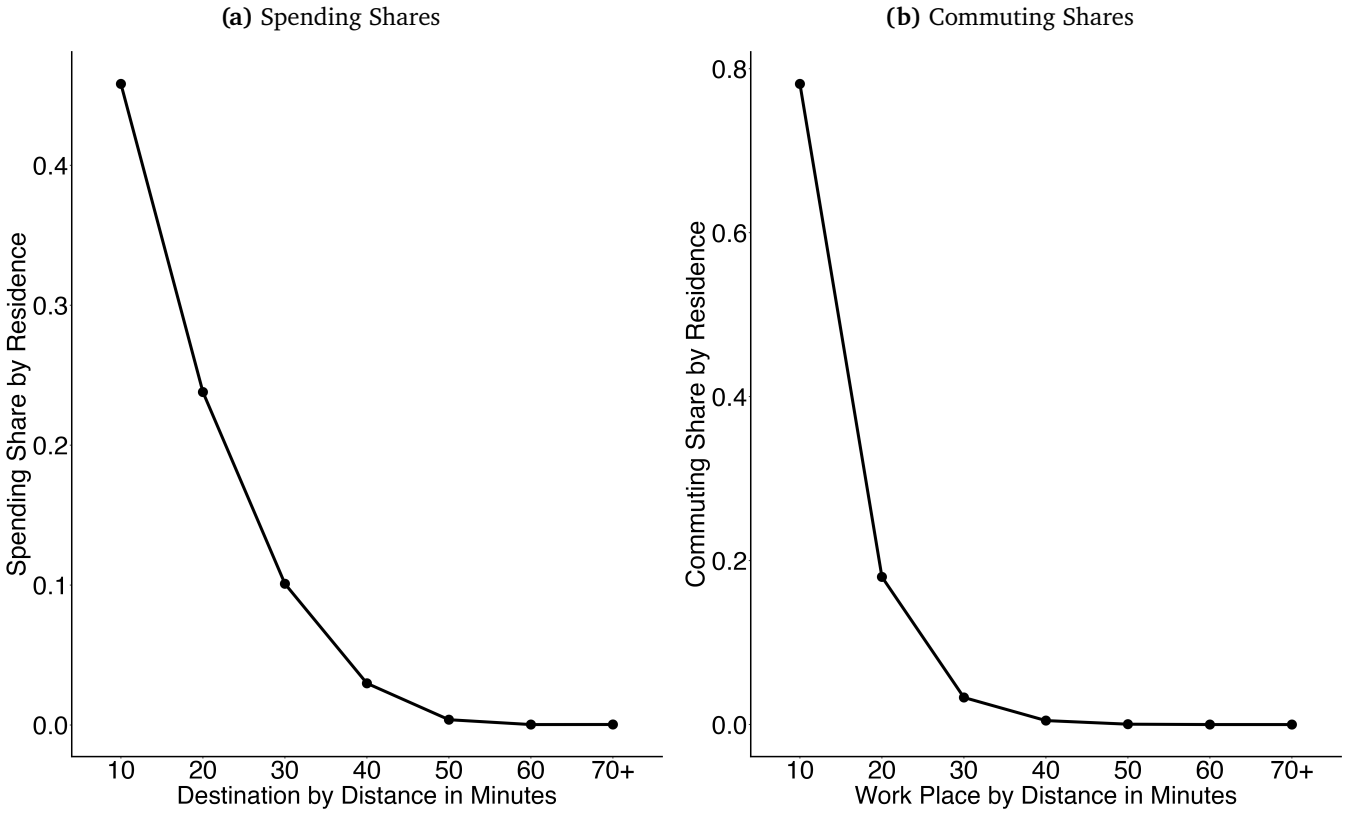
Table A.3. Origin Income Effects

Dependent Variable:	**Residents Income** $\log y_{nt}$			
Model:	(1)	(2)	(3)	(4)
Variables				
$\log E^T_{nt}$	0.0026***			
	(0.0009)			
$\log iv_{nt}$		0.0019	0.0023	
		(0.0032)	(0.0032)	
$\log GE^{IV,\infty}_{nt}$			-0.0265	-0.0390
			(0.0585)	(0.0571)
Fixed-effects				
n	Yes	Yes	Yes	Yes
t	Yes	Yes	Yes	Yes
Fit statistics				
Observations	25,006	24,980	24,980	25,006
R^2	0.95129	0.95122	0.95122	0.95127
Within R^2	0.00025	0.00017	0.00019	1.97×10^{-5}

Clustered (n,t)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: $\log y_{nt}$ denotes at gross labor income (payroll) at residential location n. $\log E^T_{nt}$ captures actual tourist expenditure at residential location n, while $\log \text{IV}_{nt}$ captures direct exposure to tourism shocks using shift-share instruments based on pre-determined spending shares and time-varying inflows of tourist groups. $\log \text{GE}^{IV,k}_{nt}$ denotes the kth-order general equilibrium exposure terms computed via network propagation of shocks across commuting and spending networks. The regressions include origin and year-month fixed effects. Standard errors are clustered two-way at the origin-year-month level.

Figure A.4. Spending and Income are Spatial (APX)



Notes: Each panel plots how the share of spending (Panel a) or commuting (Panel b) from a given residential tract declines with distance from home. Destinations (for spending) and workplaces (for commuting) are ranked by deciles of distance from the origin tract. Lines correspond to quintiles of tourist intensity in the residential area, defined by average tourist spending per square meter in 2017. Higher quintiles represent neighborhoods more exposed to tourism and typically closer to the city center.

G.4.2 Origin Total Income Effects

To complement the expenditure-based analysis, we extend the empirical framework to study the effects of tourism shocks on local resident income. While expenditure regressions capture the immediate flows of economic activity across space, income regressions provide a broader measure of local welfare impacts at the resident level.

We maintain the same network structure and exposure measures developed above but shift the focus to the income at the origin census tract n. Specifically, we use the logarithm of total monthly income, $\log y_{nt}$, as the dependent variable. The income regressions exploit the origin-side exposure to tourism shocks and their propagation through the spatial network of commuting and spending connections.

The baseline regression specification is given by:

$$\log y_{nt} = \beta_2 \log IV_{nt} + \sum_{k=0}^{4} \beta_{2k} \log \text{GE}_k^n(\text{IV}_{nt}) + \mu_n + \delta_t + \eta_{nt}, \tag{3}$$

where $\log \text{GE}_k^n(\text{IV}_{nt})$ represents the k-th order general equilibrium exposure of residents in origin tract n to tourism shocks, constructed iteratively as described in the previous section. The specification includes origin tract fixed effects μ_n to absorb time-invariant characteristics of each tract, and month-year fixed effects δ_t to control for aggregate macroeconomic shocks and seasonal fluctuations common across the city. The error term η_{nt} captures unobserved residual variation.

We estimate the model sequentially for each propagation order $k = 0, 1, 2, 3, 4$, allowing us to trace how the inclusion of progressively higher-order network spillovers affects the estimates. Standard errors are clustered at

the origin-month level to account for potential serial correlation and heteroskedasticity across census tracts over time.

Implementation Details The income data are sourced from high-frequency administrative records aggregated to the census tract level. We adjust the income measures to account for possible top-coding and compositional shifts using auxiliary information from the national statistics office.

The spatial propagation matrices are the same as those used for the expenditure regressions, relying on pre-period bilateral spending and commuting patterns. The shift-share instrument IV_{nt} is constructed analogously to the destination-side instrument, but using the exposure of residents to tourism shocks mediated through their economic linkages across the city.

The estimation proceeds by first regressing $\log y_{nt}$ on the direct exposure term $\log \text{IV}_{nt}$ alone, and then sequentially adding the higher-order GE terms $\log \text{GE}_k^n(\text{IV}_{nt})$. This approach parallels the strategy used for expenditure regressions, ensuring a consistent and comparable assessment of the direct and indirect effects of tourism shocks on different dimensions of urban economic outcomes.

G.4.3 Higher-Order Propagation

Table A.4. Shock Propagation

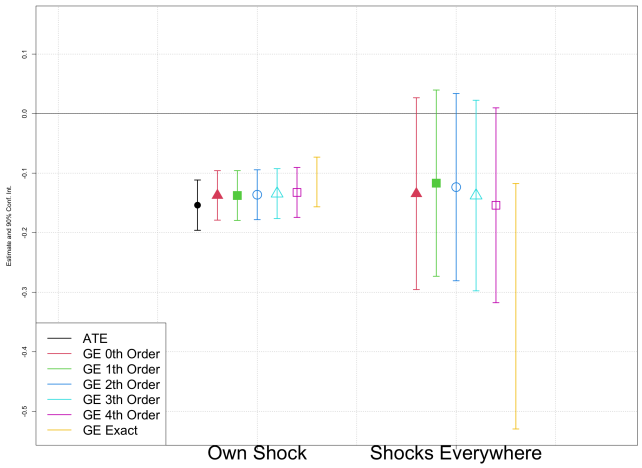
Dependent Variable:	**Residents Spending Flows** $\log E^R_{nit}$					
Model	0th Order	1st	2nd	3rd	4th	Full
Variables						
Origin Spending Effects (n)						
$\log iv_{nt}$	0.0004	0.0012	0.0016	0.0018	0.0020	0.0039
	(0.0063)	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0062)
$\log GE0iv_{nt}$	0.5096***					
	(0.1176)					
$\log GE1iv_{nt}$		0.3972***				
		(0.1026)				
$\log GE2iv_{nt}$			0.3656***			
			(0.1005)			
$\log GE3iv_{nt}$				0.3560***		
				(0.1014)		
$\log GE4iv_{nt}$					0.3547***	
					(0.1033)	
$\log GEiv_{nt}$						0.2805**
						(0.1276)
Destination Spending Effects (i - shopping)						
$\log iv_{it}$	-0.1384***	-0.1376***	-0.1358***	-0.1337***	-0.1315***	-0.1071***
	(0.0253)	(0.0254)	(0.0254)	(0.0255)	(0.0255)	(0.0256)
$\log GE0iv_{it}$	-0.1200					
	(0.0980)					
$\log GE1iv_{it}$		-0.1177				
		(0.0964)				
$\log GE2iv_{it}$			-0.1326			
			(0.0979)			
$\log GE3iv_{it}$				-0.1525		
				(0.1006)		
$\log GE4iv_{it}$					-0.1741*	
					(0.1038)	
$\log GEiv_{it}$						-0.4489***
						(0.1498)
Fixed Effects						
$n \times i$	Yes	Yes	Yes	Yes	Yes	Yes
t	Yes	Yes	Yes	Yes	Yes	Yes
Fit Statistics						
Observations	3,513,738	3,513,738	3,513,738	3,513,738	3,513,738	3,513,738
R^2	0.6495	0.6495	0.6495	0.6495	0.6495	0.6495
Within R^2	0.00013	0.00012	0.00012	0.00012	0.00012	0.00013

Clustered ($n \times t$ & $i \times t$) standard errors in parentheses.
Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

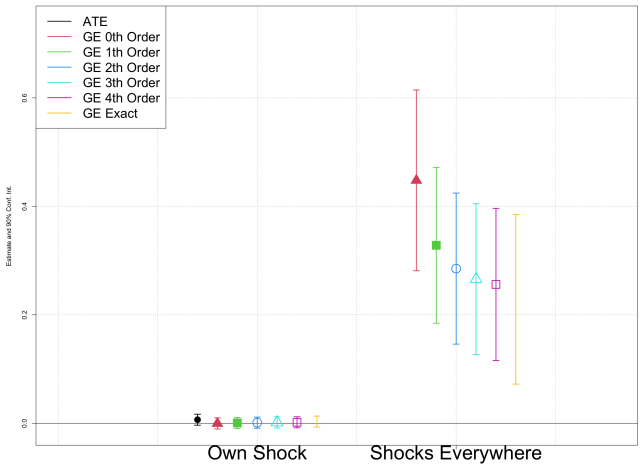
Notes: All variables are expressed in logarithms. $\log E^R_{nit}$ denotes bilateral expenditure of residents between origin n and destination i at time t. $\log E^T_{it}$ and $\log E^T_{nt}$ capture tourist expenditure at destination and origin, respectively. $\log \text{IV}_{it}$ and $\log \text{IV}_{nt}$ measure direct exposure to tourism shocks at destination and origin, based on shift–share instruments constructed from pre-determined spending shares and time-varying inflows of tourist groups. $\log \text{GE}^{IV,k}_{it}$ and $\log \text{GE}^{IV,k}_{nt}$ denote kth-order general-equilibrium exposure terms obtained through network propagation of shocks (or of tourist spending in the OLS case) across commuting and spending networks. All regressions include origin–destination and various combinations of month and origin–month fixed effects. Standard errors are two-way clustered by origin–month and destination–month. Data are based on bilateral card-transaction records between census tracts in Barcelona, 2018–2019.

Figure A.5. Shock Propagation Effects

(a) Destination Spending Effects



(b) Origin Spending Effects



Notes: These figures plot the implied elasticity of local spending with respect to tourism shocks combining the estimates in Table A.4, combining the estimates from the destination spending regressions with the spatial propagation matrices described in Section **??**. Panel (a) shows effects at destination (shopping location) while Panel (b) shows effects at origin (residential location). Each point represents the estimated elasticity under a different degree of general-equilibrium (GE) feedback, ranging from the direct local effect (ATE) to successive orders of propagation through the network of destinations. The first cluster labeled *Own Shock* reflects the effect of the local tourism shock in isolation, while the second cluster labeled *Shocks Everywhere* captures the cumulative response when shocks occur across all neighborhoods simultaneously. Confidence intervals are shown at the 90% level.

G.4.4 Housing Regressions

The housing price data come from administrative transaction records aggregated at the census tract level and standardized to ensure consistency across locations and over time. We deflate housing prices to account for city-wide inflation and seasonality effects.

The construction of the shift-share instrument IV_{nt} and the associated general equilibrium exposure terms $\log \text{GE}_k^n(\text{IV}_{nt})$ follows the approach described earlier, ensuring that all exposure measures are predetermined and exogenous to contemporaneous price movements.

We estimate the regression sequentially across orders $k = 0, 1, 2, 3, 4$, progressively incorporating higher-order network spillovers. Standard errors are clustered at the origin-month level to allow for arbitrary correlation patterns within tracts over time.

Table A.5. Housing Effects

Dependent Variables:	log **Housing Prices**$_{nt}$		log **Rent**$_{nt}$	
Model:	(1)	(2)	(3)	(4)
Variables				
log iv_{nt}	0.0371***	0.0099***	0.0166***	0.0062***
	(0.0007)	(0.0006)	(0.0004)	(0.0004)
$\log GE_{nt}^{IV,\infty}$		0.2402***		0.1178***
		(0.0134)		(0.0113)
Fixed-effects				
District × t	Yes	Yes	Yes	Yes
Fit statistics				
Observations	24,957	24,957	24,937	24,937
R^2	0.73340	0.83185	0.68246	0.73071
Within R^2	0.14386	0.46002	0.09245	0.23037

Clustered (n,t)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Housing prices and rent are in euros per square meter.

Table A.6. Housing Effects

Dependent Variables:	logHp.nt		logRent.nt	
Model:	(1)	(2)	(3)	(4)
Variables				
logiv.nt	0.0103***	0.0119***	-0.0016	2.9×10^{-5}
	(0.0031)	(0.0031)	(0.0027)	(0.0027)
logGEinfiv.nt		-0.1014***		-0.1015***
		(0.0372)		(0.0315)
Fixed-effects				
n	Yes	Yes	Yes	Yes
t	Yes	Yes	Yes	Yes
Fit statistics				
Observations	24,957	24,957	24,937	24,937
R^2	0.98108	0.98110	0.91072	0.91082
Within R^2	0.00112	0.00237	0.00047	0.00158

Clustered (n,t)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Housing prices and rent are in euros per square meter.

G.4.5 Robustness

Table A.7. Main Specifications: Robustness

Dependent Variable:	$\log E^R_{nit}$			E^R_{nit}	$\log E^R_{nit}$		
Model:	Baseline	Weighted (Trans.)	Weighted (Cards)	PPML	Trimmed (top/bot 1%)	Exclude August	Aggregate to AEB
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
$\log iv_{nt}$	0.0036	0.0006	0.0006	0.0004	0.0038	0.0012	-0.0066
	(0.0062)	(0.0060)	(0.0061)	(0.0085)	(0.0075)	(0.0065)	(0.0314)
$\log iv_{it}$	-0.1182***	-0.0508	-0.0629*	0.0056	-0.0922***	-0.1120***	-0.3736***
	(0.0263)	(0.0333)	(0.0348)	(0.0571)	(0.0286)	(0.0281)	(0.0672)
$\log GE^{IV,\infty}_{nt}$	0.2213**	0.5257***	0.4966***	0.1925	0.2637**	0.3872***	0.5256*
	(0.0969)	(0.1364)	(0.1325)	(0.2110)	(0.1112)	(0.1051)	(0.2703)
$\log GE^{IV,\infty}_{it}$	-0.3510***	-0.5500***	-0.4957***	-0.3241	-0.4850***	-0.3967***	-0.9834**
	(0.1308)	(0.1631)	(0.1617)	(0.2041)	(0.1533)	(0.1350)	(0.4281)
Fixed Effects							
$n \times i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
t	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit Statistics							
Observations	3,515,125	3,515,125	3,515,125	3,515,125	3,061,104	3,260,707	827,514
R^2	0.64950	0.63009	0.64254	0.88110	0.65612	0.65501	0.73548
Within R^2	0.00013	0.00028	0.00023	—	0.00013	0.00013	0.00049

Clustered (i,t) standard errors in parentheses
Significance levels: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Notes: $\log E^R_{nit}$ denotes the bilateral expenditure of residents between origin n and destination i at time t. $\log iv_{nt}$ and $\log iv_{it}$ capture direct exposure to tourism shocks at the origin and destination, respectively, using shift-share instruments based on pre-determined spending shares and time-varying inflows of tourist groups. $\log GE^{IV,\infty}_{nt}$ and $\log GE^{IV,\infty}_{it}$ denote general equilibrium exposure terms computed via network propagation of shocks across commuting and spending networks. All regressions include origin-destination ($n \times i$) and month (t) fixed effects. Standard errors are clustered two-way at the (i,t) level. Robustness specifications consist of: (1) the baseline OLS specification; (2–3) OLS regressions weighted by the number of transactions and number of cards, respectively; (4) PPML estimates conducted at the census tract level; (5) a sample that trims the top and bottom 1 percent of the instrumental variable (iv) distribution to reduce the influence of outliers; (6) a specification excluding the month of August to remove peak-season tourism dynamics; and (7) bilateral flows aggregated to the Area Estadistica Basica (AEB) level.

Table A.8. Robustness to specific locations

Dependent Variable:	**Residents Spending Flows** $\log E^R_{nit}$		
Exercise	All locations	Drop one origin at a time	Drop one destination at a time
	(1)	(2)	(3)
Direct Effects			
$\log iv_{nt}$	0.0030	0.0030 (0.00020)	0.0030 (0.00019)
$\log iv_{it}$	-0.1148	-0.1148 (0.00040)	-0.1149 (0.00136)
General Equilibrium Effects			
$\log GE^{IV,\infty}_{nt}$	0.2287	0.2287 (0.00327)	0.2287 (0.00352)
$\log GE^{IV,\infty}_{it}$	-0.3236	-0.3236 (0.00173)	-0.3236 (0.00634)

Notes: Column (1) reports the baseline full-sample estimates for the GE effects. Columns (2) and (3) summarize leave-one-out jackknife exercises in which one importer (origin) or exporter (destination) is excluded at a time. Entries in these columns report the median and standard deviation of the resulting coefficient estimates across jackknife samples. The reported standard deviations capture sensitivity of the GE estimates to individual locations and should not be interpreted as measures of sampling uncertainty.

G.5 Identification Diagnostics

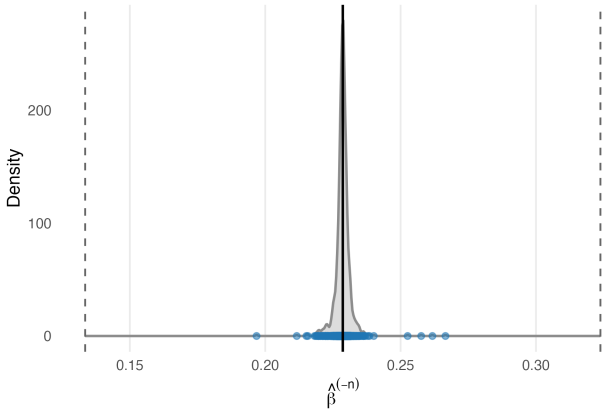
Rotemberg Decomposition

Table A.9 decomposes the identifying weight of the shift-share instrument across origin-country groups following Goldsmith-Pinkham, Sorkin and Swift (2018). Each weight $\hat{\alpha}_g$ captures the share of the total identifying variation attributable to shocks from country group g, normalized to sum to one. All weights are positive, confirming that every country group contributes in the same direction to the aggregate estimate. The distribution is concentrated: the top-3 groups (USA and Canada, Rest of the World, and France) jointly account for 50.6% of the weight, and the top 10 for 88.3%. While identifying variation is not spread uniformly across all source countries, it is not dominated by a single origin either—the largest weight (USA and Canada) is 22.9%. The robustness checks reported in the main text confirm that the main coefficients are not materially altered when any single high-weight origin is excluded.

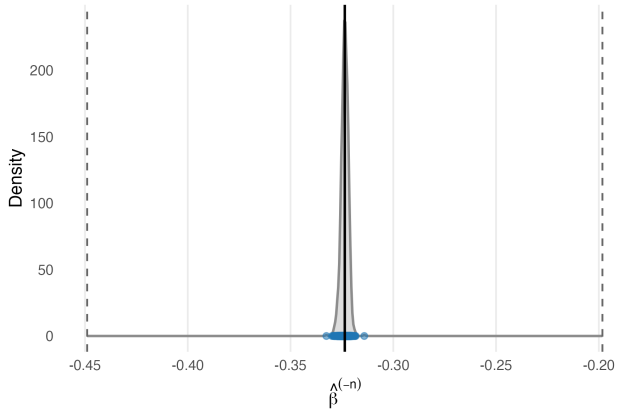
Balance Tests

Figure A.7 reports cross-sectional balance tests for the identifying component of the instrument. The left panel regresses each 2017 baseline characteristic on the raw IV exposure IV_{it}. As expected for a shift-share instrument, raw exposure is strongly correlated with pre-existing tract characteristics: tracts with higher baseline income, housing prices, rent, and tourist spending all receive substantially higher raw IV exposure, reflecting the initial concentration of tourism in wealthier central areas. The right panel instead uses the *identifying* variation—the standard deviation of $\widetilde{IV}_{it}$ after removing tract fixed effects, month fixed effects, and the predicted-trend component $\widehat{IV}_{it}$. Two features are notable. First, the number of significant correlations falls from 9/10 to 6/10 at the 5% level. Second, and more importantly, the sign of the remaining correlations reverses for income, housing prices, and rent: tracts with higher baseline economic status receive *less* of the unexpected instrument variation, the opposite of the raw-IV pattern. This reversal is consistent with unexpected surges in tourist-origin inflows being less geographically selective once systematic spatial sorting and predictable trends are removed. Panel placebo tests at the bottom of Figure A.7 show that pre-period income and rent are not significantly related to

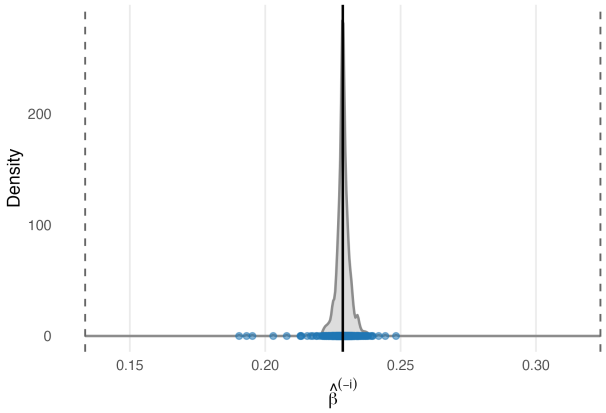
Figure A.6. The Influence of Specific Locations: Leave-one-out Robustness



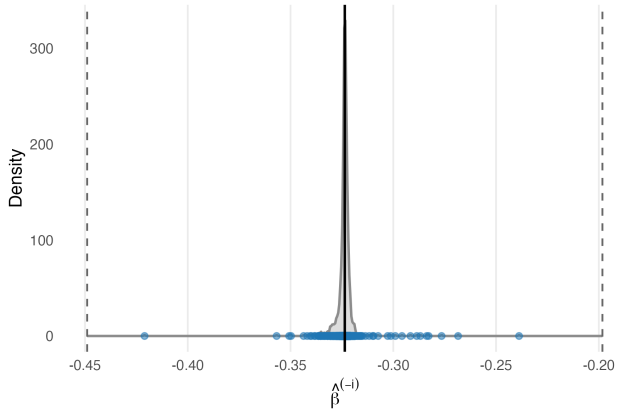
(a) Drop one origin: origin GE coefficient



(b) Drop one origin: destination GE coefficient



(c) Drop one destination: origin GE coefficient



(d) Drop one destination: destination GE coefficient

Notes: Leave-one-location-out robustness of general equilibrium (GE) estimates. Each panel reports the distribution of leave-one-out estimates obtained by excluding one origin or destination at a time. Shaded areas show kernel densities and points denote individual leave-one-out estimates. Solid vertical lines indicate the full-sample estimates, while dashed lines mark $\pm$ one standard deviation of the full-sample distribution. The tight concentration around the baseline indicates that the GE effects are not driven by any single location.

Table A.9. Rotemberg Weights: Top-10 Origin Country Groups

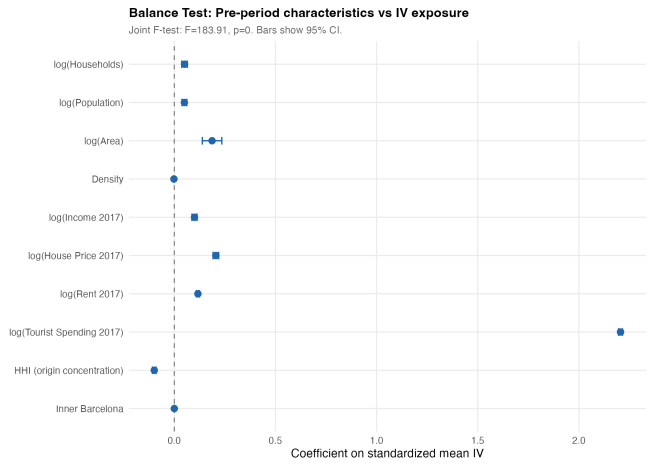
Country group	$\hat{\alpha}_g$	Cumulative
USA and Canada	22.9%	22.9%
Rest of the World	17.2%	40.1%
France	10.4%	50.6%
United Kingdom	9.1%	59.7%
Northern EU Countries	7.6%	67.3%
Rest of America	7.1%	74.3%
Germany	4.0%	78.3%
Russia	3.5%	81.9%
Rest of Europe	3.4%	85.3%
Netherlands	3.0%	88.3%

Notes: Rotemberg weights $\hat{\alpha}_g$ are computed following Goldsmith-Pinkham, Sorkin and Swift (2018) using the main bilateral specification under the baseline fixed-effects structure ($n \times i$ and t fixed effects). The weights sum to 1 and are all positive. The cumulative column reports the running share of identifying variation accounted for by the top-k country groups, ordered by $\hat{\alpha}_g$.

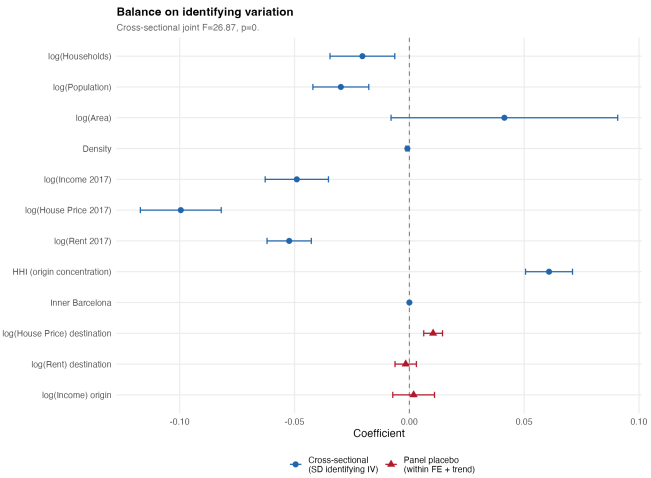
future identifying variation; a residual positive association with pre-period destination house prices remains but is small in magnitude ($\hat{\beta} = 0.010$).

Figure A.7. Balance Tests: Raw Exposure and Identifying Variation

(a) Raw IV exposure



(b) Identifying variation (residual)



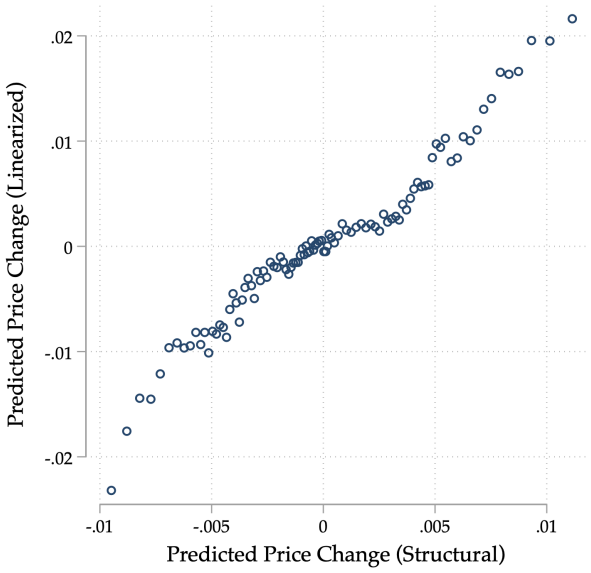
Notes: Each panel reports coefficients from cross-sectional regressions of 2017 baseline tract characteristics on measures of IV exposure, along with 95% confidence intervals. **Panel (a)** uses total raw exposure IV_{it}. **Panel (b)** uses the standard deviation of the identifying variation $\widetilde{IV}_{it}$, computed as the residual of $\log \text{IV}_{it}$ after removing tract fixed effects, month fixed effects, and the predicted-trend component $\log \widehat{IV}_{it}$, following the BHJ decomposition in Borusyak, Hull and Jaravel (2022). The cross-sectional regressions in Panel (b) include the standard deviations of both $\widetilde{IV}_{it}$ and $\widehat{IV}_{it}$ jointly, so each coefficient is purged of the other component. The bottom rows report panel placebo tests regressing pre-period house prices (destination), rent (destination), and income (origin) on the residual instrument with tract and month fixed effects. $N = 1{,}050$ census tracts for cross-sectional tests; $N \approx 25{,}000$ for panel placebo tests.

G.6 Predicted Linearized GE Versus Structural GE Effects

Figure A.8 shows predicted linearized GE versus structural GE effects by regressing the two on each other after running a first stage regression with location and time fixed effects.

Figure A.8. Comparison between Linearized Predicted Effects and Full Structural Model (Raw and Trimmed), residual with n,t FE

(a) Origin Shocks - Trimmed



(b) Destination Shocks - Trimmed

