Disaggregated Economic Accounts*

Asger Andersen Emil Toft Hansen Kilian Huber

Niels Johannesen Ludwig Straub

Please do not circulate.

October 2022

Abstract

We propose and implement a system of disaggregated economic accounts. Such a system comprehensively records bilateral flows between small groups of consumers, producers, and the government, including consumption, labor compensation, dividends and firm surplus, intermediate and foreign trade, and taxes and transfers. We develop a conceptual framework and methods allowing researchers to measure disaggregated accounts using different "bottom-up" and "top-down" approaches. In our implementation, we measure a disaggregated system for the Danish economy by assigning all consumers and producers to small region-by-industry cells. Unlike existing accounting systems, we trace the entire circular flow of money among consistently defined small groups, allowing us to establish facts on how different groups are connected. We develop a disaggregated model and show how disaggregated accounts improve our understanding of the distributional and aggregate consequences of trade shocks.

*Andersen (asger.lau.andersen@econ.ku.dk), Hansen (eth@econ.ku.dk), and Johannesen (niels.johannesen@econ.ku.dk) are at the Center of Economic Behavior and Inequality (CEBI) at the University of Copenhagen. Huber is at the University of Chicago (kilianhuber@uchicago.edu). Straub (ludwigstraub@fas.harvard.edu) is at Harvard University. We are grateful to Robert Barro, Vasco Carvalho, Erik Hurst, Greg Kaplan, and many seminar audiences for helpful comments. We thank Adam Sheridan for his contributions and Danske Bank for supporting the data work. Agustin Barboza, Jason Jia, Rodolfo Rigato, and Samuel Zhao provided outstanding research assistance. Data processing of Danske Bank customer records was conducted by authorized Danske Bank personnel, following the bank's strict data privacy guidelines. All individual data used in this analysis have been anonymized and no individual customer can be traced in the data. We gratefully acknowledge funding from the Becker Friedman Institute, Danish Innovation Fund, Danish National Research Foundation, Domenic and Molly Ferrante, Economic Policy Research Network, Fama-Miller Center, Washington Center for Equitable Growth, and William S. Fishman Fund. The activities of CEBI are financed by the Danish National Research Foundation Grant DNRF134.

I Introduction

Modern national accounting began in the 1930s, when Simon Kuznets in the US and Richard Stone in the UK developed systematic approaches to recording aggregate economic activity. Their national accounts focused on aggregate flows between all consumers and firms, most notably national consumption, income, and production. In the 1940s, Richard Stone made a first step toward measuring flows at a more disaggregated level. He assigned all producers to distinct industries and incorporated a table of cross-industry trade into a system of national accounts, inspired by Leontief's (1928) input-output analysis. These cross-industry tables have facilitated important research in the decades since, including work showing how cross-sectoral production linkages affect aggregate outcomes.[1]

However, since Stone, there has been little progress in disaggregating national accounts further. Today's standardized guidelines for national accounts still rely on Stone's seminal concept and, as a result, contain only few bilateral flows from subgroups of consumers to subgroups of producers and vice versa. Early luminaries like Richard Cantillon (1755) and François Quesnay (1758) had already outlined a vision of a completely disaggregated system, but were limited to analyzing a handful of subgroups (e.g., farmers, artisans, property owners). A lack of data and conceptual guidelines on how to link producers and consumers have made a more complete disaggregation challenging. In particular, economists have not constructed nationally representative datasets on consumption linkages, i.e., which producers receive consumer spending from which consumers, and have not developed methods to connect consumption categories and production industries. Recent advances in national accounting record how income and wealth is distributed across different consumers (Piketty, Saez and Zucman 2018), but do not measure bilaterally which producer groups pay which consumers.

The absence of a fully disaggregated system means that we have little understanding of how different consumers and producer types are connected. In turn, this has limited research on how shocks propagate through consumption, income, and production links, how aggregate outcomes depend on the initial incidence of shocks, and how shocks affect producer and consumer groups heterogeneously. Moreover, policymakers cannot fully estimate the aggregate and distributional consequences of policy without systematic data on how affected consumers and producers transact with other groups.

In this paper, we propose and implement a system of *disaggregated economic accounts*. Our approach comprehensively disaggregates national accounting flows between consumers and producers, international trade, and transactions with the government. It does so by assigning all

[1] E.g., Horvath (2000); Jones (2011); Acemoglu et al. (2012); Baqaee and Farhi (2019*a*, 2020*b*); Caliendo, Dvorkin and Parro (2019); Bigio and La'O (2020); Rubbo (2020).

consumers and producers to consistently defined subgroups and then measuring bilateral flows between these subgroups. The sum of each set of bilateral flows equals a corresponding aggregate from the national accounts. We provide general guidelines for how researchers can measure such a system with various types of underlying microdata and for any type of subgrouping of consumers and producers. In our implementation, we measure a disaggregated system for the Danish economy for region-by-industry groupings. We combine aggregate data with detailed microdata from several sources, most notably transaction-level information on which establishments receive consumer spending from which consumers. Unlike existing accounting systems, we measure the entire circular flow of money among consistently defined, small groups, so that we can show how different consumer and producer groups are connected through consumption, income, and production flows.

Our first contribution is to develop a conceptual framework for a system of disaggregated accounts. Such a system assigns all adults to distinct "consumer cells" and all firm establishments to "producer cells." For example, in our implementation, we assign consumers to cells depending on their industry of employment and region of residence (including at least one cell for foreign consumers), while assigning establishments to producer cells depending on their industry and location (including at least one cell for foreign producers). However, the proposed framework is general and allows for any type of subgrouping.

A disaggregated system comprehensively measures bilateral financial flows between different cells. Flows go from consumer to producer cells (e.g., consumer spending of every consumer cell going to every producer cell), producer to consumer cells (e.g., labor and dividend income of every consumer cell paid by every producer cell), between producer cells (e.g., intermediate trade across all producer cells), from and to the government (e.g., every cell's taxes and transfers), and from and to a capital accumulation cell (e.g., every cell's saving and borrowing). Each type of flow is captured in one disaggregated dataset (e.g., consumer spending matrix). We show how flows are linked through accounting identities within each cell (e.g., income equals expenditures plus saving), across cell types (e.g., total sales of a producer cell equal consumer spending on that cell by all other cells), and with respect to national accounts (e.g., total consumption of all cells equals national consumption).

Based on the conceptual framework, we provide guidelines on how to measure a disaggregated system of bilateral flows. There are two general approaches for a specific flow. A "bottom-up" method uses microdata on individual flows (e.g., consumer payment transactions linking consumers to retail establishments) to calculate cell-to-cell flows. This approach requires accurate, large, and representative microdata, ideally from administrative sources. Alternatively, a "top-down" method distributes an aggregate flow across cells using an assignment algorithm. We develop several such algorithms that vary depending on the flow. Typically, algorithms use cell

sizes and distance relationships as inputs and can be calibrated to the country of study. We develop additional general tools: a cross-walk between merchant categories (MCC) and industry codes (NACE), which is crucial in defining a consistent set of producer industry cells; programs that ensure that the system is consistent with national accounting identities; and calculations that transform national accounting tables from output-based to sales-based flows, which make up the disaggregated system.

We implement our guidelines by measuring a comprehensive system of disaggregated accounts in Denmark for 5,900 region-by-industry cells. Cells are small, with the median consumer cell containing 658 adults, and capture a rich degree of heterogeneity in flows. We use the "bottom-up" method to measure bilateral flows of consumer spending, labor compensation, and government transfers and taxes, among others. To measure consumer spending flows, we rely on unique data containining region and industry of both consumer and establishment receiving payment, provided by the largest Danish retail bank. We use the "top-down" method to measure producer exports and intermediate trade, using as input a cross-industry intermediate trade table from the national accounts, among others.

With the set of disaggregated accounts in hand, we turn to showing how such accounts improve our understanding of the macroeconomy. We produce a set of facts derived from combining the disaggregated datasets. First, distance has a strong effect on consumer spending, labor compensation, and intermediate trade. The effect of distance depends on the nature of a transaction, e.g., it is large for everyday purchases and smaller for travel-related goods. Second, consumers "spend upward in density," i.e., the population density of a consumer cell's home region is almost always lower than the average population density of the cells receiving its spending. Third, consumers living in urban regions spend a larger fraction of their total consumption expenditures abroad.

To interpret these facts formally, we develop a spatial equilibrium model matching the level of disaggregation of our measurement. The model contains many region-by-industry consumer and producer cells, foreign consumers and producers, and a government. Consumers have heterogeneous preferences across the output of every producer cell, which we calibrate using the measured consumer spending matrix. Producers use heterogeneous quantities of labor from every consumer cell as input and pay consumer cells different amounts of labor compensation, dividends, and surplus income, calibrated using the measured matrices. Producers also trade with each other, calibrated using the disaggregated intermediate matrix. Finally, the government transacts heterogeneously with every cell, measured by a set of disaggregated datasets on government transactions.

The combination of disaggregated data and model allows us to study the distributional and aggregate consequences of a range of economic shocks. In this paper, we use the calibrated

model to answer a specific question: How do lower trade costs affect the welfare of urban versus rural consumers? We find that the equilibrium gains from trade are larger for urban consumers, even though their employers (urban producers) export a slightly smaller share of their output. The negative relationship between relative export shares and relative gains is explained by the structure of disaggregated flows across regions. Specifically, the initial trade cost shock directly raises the incomes of rural residents. Since rural consumers spend a larger share of their income in urban regions, the incomes of urban residents also increase after the trade cost shock. This increase in urban incomes has little effect on rural consumers' income, however, because urban residents tend to spend in cities and abroad. Ultimately, after the initial gains to rural residents have circulated through the full disaggregated system, the income gains to urban residents are larger than those of rural consumers. Overall, this result highlights how a disaggregated system can improve our understanding of how shocks with heterogeneous direct exposure have subtle consequences.

II Disaggregated Economic Accounts in Relation to Existing Measurement

Cantillon (1755) and Quesnay (1758) first attempted to record transactions between subgroups of consumers and producers in 18th century France. Their vision presents the national economy as a complex circular flow of money, with different consumer and producer groups (e.g., farmers, artisans, property owners) connected through bilateral consumption, income, and output flows. They emphasize that a satisfactory system of measurement should meet three principles. First, it satisfies the accounting identities linking the consumption, income, and output of each group. Second, it captures transactions between consumer and producer groups comprehensively, so that measured flows add up to national aggregates. And third, it accounts for heterogeneity in bilateral flows across disaggregated consumer and producer groups.

Economists in the early 20th century developed these ideas further. Lahn (1903), Foster (1922), and Knight (1933) visualized aggregate accounting identities using diagrams of the circular flow of money. Kuznets, Epstein and Jenks (1941, at the US National Bureau of Economic Research) and Meade and Stone (1941, at the British Statistical Office) comprehensively measured the components of the circular flow (i.e., national consumption, income, and output). Leontief (1928; 1966) instead focused on disaggregation and measured intermediate trade between production industries.

Synthesizing and innovating on decades of work, Stone (1961) incorporated all these contributions into a general system of national accounts.[2] Today's accounts, globally standardized by the United Nations System of National Accounts (SNA), still follow his concept. That means, however, that existing systems do not fully implement the principles of Cantillon and Quesnay because they

[2]Stone also chaired the first United Nations Committee on National Income Statistics in 1947, which defined the first international benchmark for national accounts.

only disaggregate flows between producer groups. There exist neither a framework nor actual measurement of a system that comprehensively documents bilateral consumption and income flows between consumer and producer groups.

The system of disaggregated economic accounts developed in this paper fills this gap. We reconnect the classical principles of comprehensive aggregation, accounting identities, and disaggregated bilateral flows in one measurement system. We disaggregate the full circular flow, including consumption, income, and output. We devote particular attention to measuring bilateral consumption flows, since there exists little data showing how different consumer groups spend on different producer groups. This gap is particularly striking given that measuring consumption is a chief aim of national accounting, as emphasized by Kuznets, Epstein and Jenks (1941) and Barro (2021).

Complementing our work, three recent measurement innovations are noteworthy. First, distributional national accounts document the distribution of income and wealth across consumer groups (Piketty, Saez and Zucman 2018; Blanchet et al. 2021). Second, financial and online transaction data can be used to measure economic activity in high frequency (Chetty et al. 2020; Andersen et al. 2022; Vavra 2021) and inform consumption dynamics (Aladangady et al. 2021; Ehrlich et al. 2020; Buda et al. 2022).[3] They also allow researchers to investigate business entry and exit (Glaeser, Kim and Luca 2019), spending patterns across space (Davis et al. 2019; Agarwal, Jensen and Monte 2020; Dunn and Gholizadeh 2020; Baker and Kueng 2022), and living standards across regions (Diamond and Moretti 2021). Inspiring our consumer spending matrix, Baker (2018) demonstrates the power of transaction and accounts data for understanding consumer behavior. Third, government registers document income flows between producers and consumers (e.g., Card, Heining and Kline 2013; Adão et al. 2022) and intermediate trade between producers (e.g., Huneeus 2018; Dhyne et al. 2021; Bernard et al. 2022).

Three features distinguish disaggregated economic accounts from these existing approaches. First, we capture the entire circular flow of money. This means that we jointly and consistently measure consumption, income, and output flows, rather than focusing on a subset of the national accounts. Second, we focus on bilateral flows between producer and consumer groups, rather than just heterogeneity on the consumer or producer side. Finally, the disaggregated system is comprehensive, so that it covers national aggregates entirely, and it is consistent with national accounting identities within and across groups.

[3]Most transaction-based consumption datasets focus on credit cards and therefore do not capture all types of spending (i.e., cards, direct debits, bank transfers to firms, and mobile payments), apart from Buda et al. (2022) and our Danish data.

III Concept and Methods for Disaggregated Economic Accounts

In this section, we outline the concept of disaggregated economic accounts and the methods used to construct a disaggregated system. The ultimate aim is to measure bilateral flows between small subgroups of consumers and producers, so that measured flows capture the full circular flow of money, are consistent with accounting identities, and add up to national aggregates.

III.A Defining Disaggregated Cells

A disaggregated system assigns all consumers and firm establishments to non-overlapping groups, which we term consumer and producer cells, respectively. Cell classifications can be chosen flexibly depending on the research question and setting at hand. For example, consumer cells could depend on region, work industry, age, or income as well as interactions of these characteristics, whereas producer cells could depend on region or industry. At the extreme, the most disaggregated cell definition is at the level of individual consumers and establishments. At least one consumer cell represents foreign consumers and, likewise, one producer cell represents foreign producers, but there could be multiple cells, for instance, different country-by-industry cells.[4]

The system also includes a capital accumulation cell. It receives saving, interest, and financial transfers from all other cells, while issuing financial assets, loans, and interest payments. It also conducts investment transactions, by purchasing investment goods from producers and selling investment goods to producers.[5] Moreover, the system contains a government cell representing the entire public sector as well as the small group of non-profit institutions serving households (NPISH).

III.B Relation to the UN System of National Accounts

Our disaggregated system, and in particular the definitions of the disaggregated flows, builds on the global standard for aggregate national accounts defined in the United Nations System of National Accounts (SNA). However, an important difference between the SNA and disaggregated accounts lies in the definition of the units of analysis. The SNA groups all individuals and organizations

[4]The consumer and producer cells do not need to be defined symmetrically. While we choose a region-by-industry cut for both consumer and producer cells in our implementation (Section IV), one could, for instance, also choose a region cut for consumers and an industry cut for producers.

[5]Note that national accounts differentiate between the finance industry, which sells financial services (i.e., financial advice and interest rate spreads generated by financial firms), versus the capital accumulation accounts, which record savings, loans, and investment. Our implementation disaggregates the financial industry into many regional cells and contains just one capital accumulation cell. An alternative system could include multiple capital accumulation cells, one each for type of capital-accumulating institution (e.g., commercial versus investment banks etc.) and for different production industries trading investment goods with each other (vom Lehn and Winberry 2022). We leave such disaggregation to future work.

into five "institutional sectors:" non-financial corporations, financial corporations, government, NPISH, and households (including unincorporated businesses). The SNA government, NPISH, and household sectors are both producers and consumers of output. In contrast, our disaggregated system assigns all establishments engaged in production to producer cells, which means that the production sides of all five SNA sectors are part of the disaggregated producer cells. In contrast, consumer, government, and capital accumulation cells in the disaggregated system do not produce but mostly buy producers' output and receive payments (e.g., labor compensation and divdends) from producers.

Moreover, we assign NPISH to the government cell, since NPISH are very small and, just like the government, purchase output from producers and provide it to consumers free of charge (or at low nominal fees). Finally, the UN SNA contain capital and asset accounts within each of the five sectors, which record investment and asset accumulation in detail. We do not disaggregate these flows in this intitial concept for disaggregated accounts and, instead, the capital accumulation cell jointly carries out all these transactions.

III.C Components of Disaggregated Economic Accounts

A full system of disaggregated accounts contains several individual datasets. Each disaggregates one flow from the aggregate national accounts into a comprehensive set of bilateral flows between pairs of cells. In Table I, we list all disaggregated datasets needed to measure flows between consumer cells, producer cells, a government cell, and a capital accumulation cell. Of course, for many research and policy questions, researchers may find it sufficient to disaggregate only a subset of these flows.

Table I: Components of disaggregated economic accounts

	Disaggregated dataset	Financial outflow from	Financial inflow to
1	Domestic consumer spending matrix	Consumers	Producers
2	Foreign consumer spending	Consumers	Rest of world
3	Consumer taxes	Consumers	Government
4	Consumer interest, transfers, and saving	Consumers / Capital acc.	Capital acc. / Consumers
5	Consumer labor compensation matrix	Producers	Consumers
6	Consumer dividend, mixed income, and surplus matrix	Producers	Consumers
7	Government benefits to consumers	Government	Consumers
8	Intermediate trade matrix	Producers	Producers
9	Government dividend and surplus income	Producers	Government
11	Producer net taxes	Producers	Government
12	Producer net interest, transfers, and saving	Producers	Capital acc.
13	Producer imports	Producers	Rest of world
14	Domestic government spending	Government	Producers
15	Domestic capital accumulation spending	Capital acc.	Producers
16	Producer exports	Rest of world	Producers
17	Government imports	Government	Rest of world
18	Government net interest, transfers, and saving	Government	Capital acc.
19	Capital accumulation imports and trade balance	Rest of world / Capital acc.	Capital acc. / Rest of world
20	Consumption of government output	Provided free of charge to consumers	

Notes: The table lists the datasets (matrices and vectors) that make up a system of disaggregated economic accounts. A matrix is a dataset composed of multiple cells along both the outflow and inflow dimensions, while a vector contains only one inflow or one outflow dimension. The consumption values in "Consumption of government output" are provided free of charge, so that no money flows from consumers to the government cell in exchange for the goods. The government cell includes all NPISH (non-profit institutions serving households). The outflows and inflows are defined in terms of financial flows, so that goods flow in the opposite direction.

Each disaggregated dataset is comprehensive, in the sense that the sum of its individual observations equals the national aggregate (e.g., the total of the consumer spending matrix equals national consumer spending). The observations in different disaggregated datasets are linked by accounting identities that hold within each cell (i.e., all inflows equal all outflows), as we describe in detail Section IV.

III.D "Bottom-Up" and "Top-Down" Disaggregation Approaches

There are two approaches to disaggregating a national accounting flow. First, the "bottom-up" approach uses detailed microdata on individual transactions between consumers and producers,

such as individual spending transactions. Adding up all transactions between two cells and, if appropriate, weighting and scaling to match the underlying population produces a cell-to-cell bilateral flow. The second disaggregation approach is "top-down." It decomposes an aggregate quantity into bilateral flows using an algorithm based on cell characteristics like group size, geographic distance, and other microdata on consumer and producer behavior.

If microdata representative of the population are available, the bottom-up approach captures bilateral flows accurately. In the absence of such microdata, the top-down approach can provide a suitable approximation.

III.E Measurement Challenges

Disaggregated accounts raise several measurement challenges. A first challenge is the construction of producer cells. Different datasets do not report the activities of producer establishments in a consistent way. Consumption datasets typically classify consumer-facing establishments using the type of merchant (e.g., Merchant Category Code, MCC). In contrast, firm and employment datasets usually classify firm establishments into production industries (e.g., NACE or NAICS). We develop a new cross-walk allowing researchers to define producer cells consistently across consumption and firm/employment datasets (see Appendix E.C).

A futher challenge is that national accounts data need to be transformed before they can inform a disaggregated system. For instance, standard national accounts measure trade among producers in terms of output and consumption in terms of gross sales. This is important in the case of retail and wholesale producers, where output does not equal gross sales and where the raw trade and consumption data therefore do not correspond. We show how researchers can transform national accounts tables to ensure consistency across data sources.

Finally, researchers need to tailor top-down algorithms to the requirements of the national accounting flow that needs to be disaggregated. We develop algorithms for each flow and estimate flexible parameters that researchers can use to inform their top-down algorithms (e.g., the effect of distance on consumer spending for consumers living in different types of regions).

IV Measurement of Disaggregated Economic Accounts in Denmark

We implement a comprehensive measurement exercise for the Danish economy in 2018. We outline the main steps in this section, with methodological details described in Appendices C to P.

IV.A Region-Sector Cells

We assign all Danish adults to consumer cells (indexed by i) based on their region of residence and main industry of work. There are 2,772 consumer cells in total. The cells are formed from

the interaction of 99 regions, one for each of the Danish municipalities plus one for the rest of the world, and 28 industries paying labor compensation to consumers (listed in Table A.I). The classification includes industries selling directly to consumers (e.g., food away from home, entertainment, grocery stores, drug stores), non-retail industries selling mostly to firms (e.g., wholesale, agriculture, manufacturing, business services), and four industries for the non-working parts of the population (e.g., retired, students, unemployed, out of workforce).

We assign firm establishments to 2,673 producer cells (indexed by j), based on their region and production industry. There are 24 producer industries paying labor compensation to consumers as well as three producer industries providing housing services without any employees (private landlords, owner-occupied housing, government-owned housing).

The individual cells are small, with 658 adults in the median consumer cell and 47 establishments in the median producer cell. The region-by-industry breakdown is useful because, as we will show below, region and industry strongly predict heterogeneity in all types of flows and because many shocks of interest affect regions and industries heterogeneously. There is also a practical benefit to our cell definition, as we can observe region and industry across all underlying datasets.

IV.B Disaggregated Flows from and to Consumers

The accounting identity for consumer cell i says that total outflows equal total inflows,

$$\begin{aligned} &Domestic\,consumer\,spending_i + Foreign\,consumer\,spending_i \\ &+ Consumer\,taxes_i + Interest,\,transfers,\,and\,saving\,paid_i \\ &= Labor\,comp_i + Producer\,dividends,\,mixed\,inc,\,and\,surplus_i \\ &+ Government\,benefits_i + Interest\,and\,transfers\,rec_i. \end{aligned} \tag{1}$$

The disaggregated datasets break down every component of equation 1 into bilateral flows. We describe how we measure each disaggregated flow in turn.

IV.B.1 Disaggregated Consumer Spending Matrix

The consumer spending matrix measures how much consumer cell i spends on goods produced by every producer cell, both domestically and in the rest of the world. The closest corresponding position in the UN SNA is P.3, national final consumption expenditure. However, P.3 includes both spending flowing to producers and product taxes (e.g., value added taxes, import duties). In contrast, our spending matrix only records spending flowing to producers, whereas taxes appear in a separate dataset.

We take four steps to construct the spending matrix. First, we rely on transaction-level data from Danske Bank, the largest retail bank in Denmark, to measure cell-to-cell gross spending using a bottom-up approach. Second, we adjust a few specific cell-to-cell flows in cases where government data can improve the bank data. Third, we scale to match the population. Finally, we convert gross into net spending by deducting product taxes. The full details of the four-step procedure appear in Appendix E.

In the first step, our Danske Bank sample consists of all adults who held their main bank account at Danske Bank and conducted at least one spending transaction per month in 2018 and 2019. The sample covers 20% of the population and is representative in terms of age, income, and asset holdings (Table A.II). We observe the full range of consumer and merchant microdata necessary to construct the spending matrix only in 2018 and 2019. We include both years in the sample to maximize observations.

Crucially, text strings in the bank's internal system allow us to identify the consumers and merchant establishments involved in a wide range of individual transactions, including cards, direct debits, bank transfers to firms, and mobile payments. We infer consumers' home region from customer records and their work industry from incoming salary payments. We extract the merchant's address and category code (MCC, indicating the type of merchant) for transactions involving credit and debit cards, bank transfers, direct debits, and mobile applications. We then develop a novel cross-walk between MCC and industry codes (ISIC/NACE), so that we can identify the industry of the merchant in terms of the industry codes used in the disaggregated production and income datasets. Thanks to the standardized underlying classifications, the cross-walk can easily be adopted by other researchers who want to jointly study consumption, production, and income.

In a few cases, we do not observe merchants' region and/or industry. We assign these transactions to merchants in proportion to the observed spending of the same consumer cell using the same means of payment. Moreover, we assume that consumers spend cash withdrawals in proportion to in-store card payments (separately for withdrawals in Denmark and abroad). Cash withdrawals are 7% of aggregate transaction value, so the matrix is not sensitive to this assumption. To complete the first step, we aggregate all spending transactions going from a consumer cell to a producer cell and then scale by the ratio of consumers in the Danske Banke sample to the population.

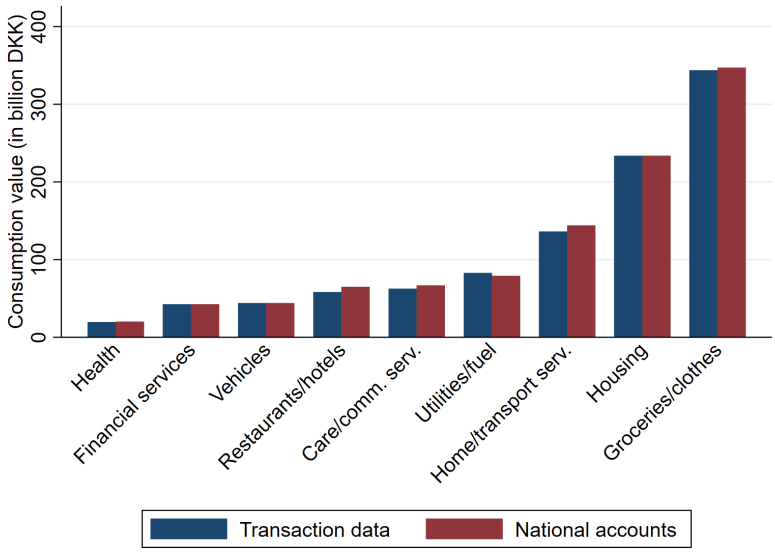
In the second step, we augment the matrix with government data on housing, financial services, vehicles, and water and waste services. Spending on these four goods is often not comprehensively measured in bank transactions. However, we can measure such spending by combining government and bank data. Notably, the government income register directly records the rental value of owner-occupied housing, interest expenses, and vehicle registrations for every

individual, and allows us to infer whether individuals are renters or owners.[6]

In the third step, we scale every cell-to-cell spending observation by a common scaling factor, so that aggregate gross spending in the matrix matches national consumption in 2018 (SNA P.3). Aggregate spending in the unscaled matrix is 3% larger than national consumer spending, largely because we include both 2018 and 2019 in the sample.

We compare the matrix produced after the third step with national accounts consumption data. Both include product taxes paid. Spending patterns by industry are very similar (Figure I). It is not clear to what extent remaining differences reflect errors in the matrix or national accounts, since both contain some statistical error due to sampling and assumptions.[7] We also have access to data on aggregate card spending in the Danske Bank for a longer period. The evolution of card spending is similar in both series (Figure II). These results suggest that the Danske Bank data match broad patterns in the national accounts.

Figure I: Spending on receiving industry in the spending matrix and national accounts

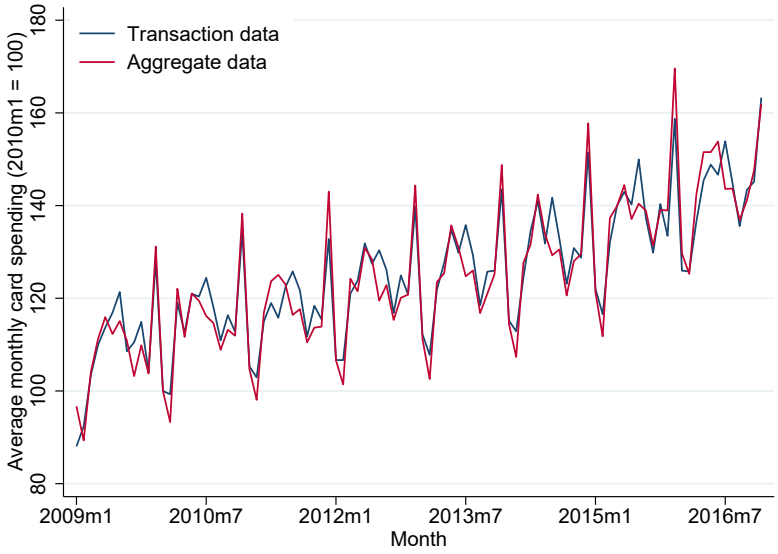


Notes: The figure compares spending aggregated by receiving industry from the spending matrix with national accounts consumption data for 2018. Housing spending is constructed using a bottom-up approach for owner-occupied housing and a top-down approach for rented housing. Vehicles, financial services, and water and waster (part of utilities) are constructed using a top-down approach, so aggregates in the transaction data and national accounts match by construction. The remaining categories are constructed using a bottom-up approach, so there is no mechanical reason that these aggregates should match. For details, see Appendix E.

[6] According to national accounting conventions, homeowners living in their own property rent to themselves, so that imputed rents are counted as consumer spending and rental income. Differences in home ownership rates therefore do not affect national consumption.

[7] For instance, national accounts do not observe spending directly, but back out spending from retail indices, which requires making assumptions on spending patterns by foreigners. In contrast, we directly observe spending by Danes and can exclude spending by foreigners.

Figure II: Card spending over time in the Danske Bank data and national accounts



Notes: The figure compares aggregated card spending over time in the Danske Bank data and national accounts.

In the fourth and final step, we subtract product taxes from each gross flow to produce the final spending matrix. We use national accounts data to calculate the product tax rate paid on each industry's products and the import tax rate by foreign industry (as detailed in Appendix F).

IV.B.2 Disaggregated Consumer Taxes

Danish consumers pay three types of taxes to the government: product taxes (SNA D.21), non-product taxes on income and pension wealth returns (SNA D.5), and social contributions (SNA D.61). We already measured product taxes paid by each cell to construct the spending matrix (details in Appendix F).

We observe individual income tax returns in the government registers and use a bottom-up summation to calculate income taxes paid per cell, ultimately scaling each cell-level observation by a common scaling factor to match the national aggregate exactly. We do not directly observe pension return taxes, since they are paid as an aggregate by pensions companies. We therefore use a top-down approach to distribute pension return taxes across cells, under the assumption that taxes paid are proportional to accumulated pension contributions. Finally, we use a bottom-up approach to measure various pension contributions, which make up social contributions, also scaling to match the national aggregate. Details are in Appendix G.

IV.B.3 Disaggregated Consumer Interest, Transfers, and Saving Paid

Consumer interest, transfers, and saving flow to the capital accumulation cell (details in Appendix H). We measure consumer interest payments (SNA D.41) bottom-up by aggregating individual-level interest payments on financial liabilities as reported in the government tax registers. We apply an adjustment factor to all interest payments, following the aggregate the national accounts, to ensure that we measure interest net of the consumption value of financial services, which is a component of the consumption matrix. Apart from interest, two very small flows go to the capital accumulation cell—rental payments for natural resources (SNA D.45) and other current transfer (SNA D.7)—which we disaggregate using a top-down approach. Finally, we measure saving of a consumer cell as the difference between all other inflows and outflows.

IV.B.4 Disaggregated Labor Compensation Matrix

The labor compensation matrix records income paid to consumer cell i by each producer cell (SNA D.1). We construct the matrix bottom-up, drawing on tax registers containing individual labor income (including employer pension contributions) and address and industry of the workplace (details in Appendix I). The aggregate of the matrix is slightly lower than in the national accounts, mostly because our sample contains only Danish adults. We scale the resulting matrix entries to match the national aggregate.

IV.B.5 Disaggregated Consumer Dividend, Mixed Income, and Surplus Matrix

We record dividend, mixed income, and surplus payments received by consumer cell i from corporate and non-corporate firms in each producer cell (details in Appendix J). To assign dividends of Danish corporations (SNA D.42), we use government data on the stock ownership of every Danish consumer. We sum stock wealth in each consumer cell and assign dividends of Danish corporations in proportion to the cells' share in national stock wealth. We thereby assume that all consumer cells hold a fully diversified portfolio of Danish corporations.

Non-corporate firms are privately-owned and mixed income accrues to their owners (SNA B.3G). We observe mixed income received and workplace of every business owner in the government registers. This allows us to construct a matrix that directly measures how much each consumer cell receives in mixed income from each producer cell. Operating surplus (SNA B.2G) corresponds to the (imputed) surplus of homeowners letting housing to themselves. We disaggregate it using each consumer cell's owner-occupied housing rental income, which is directly observed in government income registers.

IV.B.6 Disaggregated Government Benefits to Consumers

Consumers' income depends on the government through various benefit programs, including unemployment, social security, pensions, child support, and student benefits (SNA D.62). We aggregate these benefits bottom-up by summing individual-level benefits from the government registers and ultimately scaling to exactly match the national aggregate (details in Appendix K). We disaggregate other transfers (SNA D.7), which include disaster or accident relief, in proportion to each cell's population share. Last, national accounts include an adjustment to avoid double counting changes in pension fund entitlements (SNA D.8). We disaggregate this position by constructing individual pension contributions and payouts from the tax registers.[8]

IV.B.7 Disaggregated Consumer Interest and Transfers Received

Interest and transfers flow from the capital accumulation cell to consumers (details in Appendix L). We disaggregate interest (SNA D.41) and natural resource rents (SNA D.45) using a procedure analogous to the one for interest and rent paid (see Section IV.B.3). We distribute national pension investment income (SNA D.44) top-down in proportion to accumulated pension contributions.

IV.C Disaggregated Flows from and to Producers

The accounting identity for producer cell j is,

$$\begin{aligned}
&Domestic\,intermediates_j + Labor\,comp_j + Producer\,dividends,\ mixed\,inc,\ and\,surplus_j \\
&+ Dividends\,and\,surplus\,of\,government\,producers_j \\
&+ Producer\,net\,taxes_j + Producer\,net\,interest,\ transfers,\ and\,saving\,paid_j + Imports_j \\
&= Domestic\,intermediates_j + Domestic\,consumer\,spending_j \\
&+ Domestic\,government\,spending_j + Domestic\,capital\,acc\,spending_j + Exports_j.
\end{aligned}$$

Producer cells include all establishments producing market and non-market output, including financial and non-financial corporations, unincorporated businesses, government-owned firms (which are usually partly government-owned and sell output at market prices), and government-operated firms (which are run by government employees and provide mostly non-market output to consumers, e.g., in public administration).

[8] In our disaggregated system, all pension-related flows originate and go to the government cell (SNA D.61, D.62, and D.8), which has the advantage that all double-counting of pension flows naturally nets out in the government cell. Aggregate national accounts in different countries deal with pensions in different ways. In the Danish national accounts, pensions flow between consumers and government as well as between consumers and the financial corporation sector, which creates the double-counting issue.

IV.C.1 Disaggregated Intermediate Trade Matrix

The intermediate trade matrix measures flows of intermediate goods between producer cell j and every other domestic producer cell, net of product taxes. Its national counterpart is intermediate consumption (SNA P.2). Unlike for consumer spending, we do not have detailed bottom-up microdata but instead use a top-down approach composed of four steps (details in Appendix M).

We begin with a national input-output table at the most disaggregated level possible, of 117 industries. First, we face the challenge that national input-output tables do not report the sales value of goods sold by consumer-facing merchants, but only the "trade margins" earned by these merchants (final value of goods sold minus their purchase value). We generate an alternative input-output table that measures actual sales values for merchants and is therefore consistent with our consumer spending matrix. Second, input-output tables report trade across industries but not across regions. We therefore use data on cargo trade across Danish provinces to estimate a gravity model. We find an elasticity of trade with respect to distance of -0.6, which is roughly constant across different types of goods.

In the third step, we convert the input-output table to a trade matrix at the region-by-industry level, still at the level of the 177 industries. We use an iterative algorithm that ensures the trade matrix is consistent with the estimated distance elasticity and with the employment shares of different regions within the same industry. For example, when disaggregating intermediate trade from the agricultural industry to the restaurant industry, we ensure that cells with a greater share of agricultural employment sell more intermediates to cells with a greater share of restaurant employment, all while the distance elasticity remains constant. To disaggregate exports and imports across cells, we only use employment shares and not the distance elasticity.

The fourth and final step aggregates the matrix to the level of our industry-by-region producer cells. Since we started with fine industries, the resulting intermediate trade matrix contains substantial heterogeneity in both domestic and foreign trade flows within our industries.

IV.C.2 Disaggregated Labor Compensation, Dividend, Mixed Income, and Surplus Payments

Labor compensation (SNA D.1), dividends (SNA D.42), mixed income (SNA B.3g), and surplus (SNA B.2G) payments by producer cell j to each consumer cell appear in the matrices described above. We identify the share of dividends paid to the government by each producer cell (SNA D.42) using hand-collected data on the location and industries of Danish government-owned corporations (details in Appendix N). We measure the share of each producer cell's surplus accruing to the government using the share of government employees in each cell.

IV.C.3 Disaggregated Producer Sales

Intermediate sales appear in the intermediate trade matrix, whereas sales to Danish consumers appear in the domestic consumer spending matrix. We additionally record sales of producer cell j to the government ("domestic government spending") by assuming that a producer cell's share in the total sales of its industry to the government equals its employment share in the industry (details in Appendix M).[9] We follow an analogous procedure to measure sales of investment goods, inventories, and valuables by producer cell j to the capital accumulation cell ("domestic capital accumulation spending").

IV.C.4 Disaggregated Producer Net Taxes and Net Interest, Transfers, and Saving

We measure product taxes paid by each producer cell j by disaggregating product taxes paid by industry in the national input-output table (details in Appendix F). We calculate the net of non-product taxes paid and subsidies received by producers by summing aggregate positions the national accounts of the financial, incorporated, and unincorporated production sectors.[10] The distribution rules for net non-product taxes and subsidies depend closely on accounting profits in Denmark. Hence, we distribute non-product taxes and subsidies across cells in proportion to the accounting profits of each cell (total sales minus intermediates minus labor compensation). Finally, we calculate the net of interest, transfers, and saving by each producer cell as the difference between total inflows and outflows.[11]

[9] In national accounts, government-operated establishments sell their output to the government (a process called "government consumption"). Since these establishments do not sell at market prices, we follow national accounting conventions and equate the value of their sales to the cost of production (i.e., total value of intermediates, labor, and capital usage costs).

[10] Producer taxes paid include the following SNA positions: D.29, other taxes on production; D.5, current taxes on income, wealth; D.62, social benefits other than social transfers in kind; D.8, adjustment for the change in pension entitlements, payable. Subsidies received include: D.39, other subsidies on production; D.61, net social contributions.

[11] The closest analog to producer net interest, transfers, and saving in the aggregate national accounts is the sum of the following SNA positions in the consumer, non-financial corporate, and financial corporate accounts: interest paid - received (net of D.41) + reinvested earnings on direct foreign investments other current transfers paid - received (net of D.43) + other investment income paid - received (net of D.44) + other current transfers paid - received (net of D.7) + natural resource rents paid (D.45) + gross saving (B.8g) + distributed income of corporations paid to rest of world (part of D.42) - distributed income of corporations received (D.42). This sum based on the aggregate national accounts differs slightly from "producer net interest, transfers, and saving" in our disaggregated accounts because we do not measure labor compensation from and to foreign countries, because we record inflows of consumer-facing merchants in terms of sales and not trade margins (see step 1 in Section IV.C.1), and because we treat pension-related flows differently (see Section IV.B.6).

IV.D Flows from and to Government

We include one cell for the Danish government, so the government flows are aggregate values for the entire economy and have no subscript,

$$\begin{aligned}
&Domestic\,government\,spending + Government\,imports \\
&+ Government\,benefits \\
&+ Government\,net\,interest,\ transfers,\ and\,saving\,paid \\
&= Consumer\,taxes + Producer\,net\,taxes \\
&+ Dividends\,and\,surplus\,of\ government\,producers.
\end{aligned}$$

The main function of the government cell is to purchase output from domestic government-operated establishments (so-called "domestic government spending") and from abroad ("government imports") and to provide this output to consumers free of charge or at low nominal fees (so-called "consumption of government output"). Government-operated establishments are primarily in public administration, defense, education, and healthcare. We already disaggregated the value of government spending flowing to each producer cell in Section IV.C.3 and we observe government imports directly in the national accounts of the government sector.

We also measure the consumption of government output by each consumer cell. Since there are no financial flows associated with this consumption, it does not appear in the consumer accounts (equation 1) or the spending matrix. We instead record consumption of government output in a separate dataset by combining individual-level government records with Danske Bank data (details in Appendix O).

We measure total government benefits paid to all consumer cells by summing the consumer cell-level measures from Section IV.B.6. Similarly, we record taxes collected from consumers (measured in Sections IV.B.2 and IV.B.6) and producers (Section IV.C.4) by summing cell-level figures. We recorded dividends and surplus from government-owned corporations in each producer cell in Section IV.C.2. Finally, we measure government net interest, transfers, and saving outflows as the difference between all other outflows and inflows.[12]

[12]The closest analog to government net interest, transfers, and saving in the aggregate national accounts is the sum of the following government SNA positions: interest paid - interest received (net of D.41) + other current transfers paid - other current transfers received (net of D.7) + gross saving (B.8g) - other investment income (D.44) - natural resource rents received (D.45). However, this sum based on the aggregate national accounts does not equal "net interest, transfers, and saving" in our disaggregated accounts because we do not disaggregate taxes, benefits, and subsidies received by institutional sectors other than producer and consumers and because we treat pension-related flows slightly differently (see Section IV.B.6).

IV.E Flows from and to the Rest of the World and the Capital Accumulation Cell

We measure inflows and outflows for two remaining cells, the rest of the world and capital accumulation. This step ensures that all consumer, producer, and government flows disaggregated so far have a well-defined counterparty cell. The accounting identity for the rest of the world is,

$$\begin{aligned}&Foreign\,consumer\,spending + Producer\,imports\\&+ Government\,imports + Capital\,acc\,imports + Trade\,balance\\&= Producer\,exports.\end{aligned}$$

We have so far measured all these flows except capital accumulation imports, which we observe as imports of investment goods, valuables, and inventory in the national input-output table. The trade balance is total exports minus total imports (including foreign consumer spending).[13]

The accounting identity for the capital accumulation cell is,

$$\begin{aligned}&Consumer\,interest,\,transfers,\,and\,saving\,rec + Domestic\,capital\,acc\,spending\\&+ Capital\,acc\,imports + Trade\,balance\\&= Consumer\,interest,\,transfers,\,and\,saving\,paid\\&+ Producer\,net\,interest,\,transfers,\,and\,saving\,paid\\&+ Government\,net\,interest,\,transfers,\,and\,saving\,paid.\end{aligned}$$

We have already accounted for all these flows in the previous sections.

IV.F Overview of the Disaggregated Circular Flow in Denmark

We next provide a summary and visualization of disaggregated flows in Denmark. In Table II, we list all disaggregated flows along with the closest related SNA position (see table notes). We have disaggregated every individual flow from the SNA household sector account, while we combine flows related to producer and government transactions with the capital accumulation cell in joint flows (24, 30). These flows are not central to our analysis of facts and application below, but could be relevant to other questions and we therefore leave their disaggregation to future work.

[13]The Danish trade balance in our data is slightly below the aggregate trade balance measured in the national accounts (SNA P.6 - P.7). The difference is that the spending matrix implies a slightly higher value for foreign consumer spending than the national accounts, and thus slightly higher total imports. An advantage of our approach is that we can directly observe foreign spending by Danish consumers in the Danske Bank data and that our spending measures do not include spending by foreign consumers in Denmark. In contrast, national accounts rely on balance of payments statistics, retail turnover, and consumer surveys.

Table II: Overview of all flows in the Danish disaggregated economic accounts

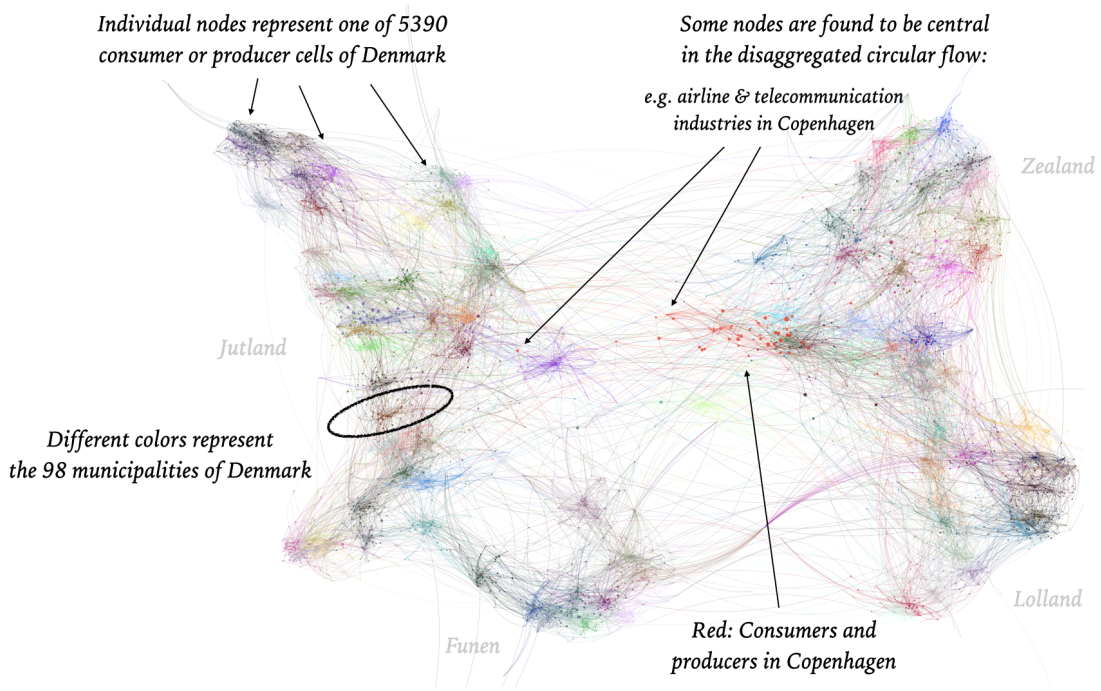
	Disaggregated flow name	Related SNA code	Outflow from	Inflow to	Disaggregated dataset	Total value (bn DKK)
1	Domestic consumer spending	P.3[a]	Consumers	Producers	Domestic consumer spending matrix	771.9
2	Foreign consumer spending	P.3[a]	Consumers	Rest of world	Foreign consumer spending	81.9
3	Consumer product taxes paid	D.21	Consumers	Government	Consumer taxes	173.2
4	Consumer non-product taxes paid	D.5	Consumers	Government	Consumer taxes	566.4
5	Consumer social contributions paid	D.61	Consumers	Government	Consumer taxes	181.1
6	Consumer interest paid	D.41	Consumers	Capital acc.	Consumer interest, transfers, and saving	29.7
7	Consumer natural resource rents paid	D.45	Consumers	Capital acc.	Consumer interest, transfers, and saving	3.4
8	Consumer other transfers paid	D.7	Consumers	Capital acc.	Consumer interest, transfers, and saving	44.8
9	Consumer gross saving	B.8g	Consumers	Capital acc.	Consumer interest, transfers, and saving	130.0
10	Labor compensation	D.1	Producers	Consumers	Consumer labor compensation matrix	1,141.8
11	Mixed income from non-corporate producers	B.3g	Producers	Consumers	Consumer dividend, mixed income, and surplus matrix	80.7
12	Surplus of corporate producers to consumers (dividends)	D.42	Producers	Consumers	Consumer dividend, mixed income, and surplus matrix	38.4
13	Surplus of owner-occupied housing to consumers	B.2g	Producers	Consumers	Consumer dividend, mixed income, and surplus matrix	83.3
14	Consumer social benefits received	D.62	Government	Consumers	Government benefits to consumers	422.2
15	Consumer adjustment for pension entitlements received	D.8	Government	Consumers	Government benefits to consumers	92.5
16	Consumer interest received	D.41	Capital acc.	Consumers	Consumer interest, transfers, and saving	5.3
17	Consumer pension investment income	D.44	Capital acc.	Consumers	Consumer interest, transfers, and saving	75.5
18	Consumer natural resource rents received	D.45	Capital acc.	Consumers	Consumer interest, transfers, and saving	3.4
19	Consumer other transfers received	D.7	Capital acc.	Consumers	Consumer interest, transfers, and saving	39.2
20	Domestic intermediates	P.2[b]	Producers	Producers	Intermediate trade matrix	1,414.5
21	Dividends and surplus of government-owned/operated producers to government	D.42	Producers	Government	Government dividend and surplus income	67.9
22	Producer product taxes paid	D.21	Producers	Government	Producer net taxes	71.9
23	Producer net non-product taxes paid and subsidies received	D.29 + D.5 + D.62 + D.8 - D.39 - D.61	Producers	Government	Producer net taxes	82.8
24	Producer net interest, transfers, and saving	(D.41 + D.43 + D.44+ D.7) [net outflow] + D.45 + B.8g - D.42 + D.42 to rest of world[c]	Producers	Capital acc.	Producer net interest, transfers, and saving	422.4
25	Producer imports	P.7[b]	Producers	Rest of world	Producer imports	792.3
26	Domestic government spending	P.3[b]	Government	Producers	Domestic government spending	572.3
27	Domestic capital accumulation spending	P.1[b]	Capital acc.	Producers	Domestic capital acc. spending	359.5
28	Producer exports	P.6[b]	Rest of world	Producers	Producer exports	1,077.9
29	Government imports	P.7[b]	Government	Rest of world	Government imports	4.3
30	Government net interest, transfers, and saving	(D.41 + D.7) [net outflow] + B.8g - D.44 - D.45[d]	Government	Capital acc.	Government net interest, transfers, and saving	52.0
31	Capital accumulation cell imports	P.7[b]	Capital acc.	Rest of world	Capital acc. imports and trade balance	98.9
32	Aggregate trade balance	P.6 - P.7[e]	Capital acc.	Rest of world	Capital acc. imports and trade balance	100.6
33	Consumption of government output	P.3	Provided free of charge to consumers		Consumption of government output	578.6

Notes: The table lists all the individual flows that make up the measured disaggregated economic accounts in Denmark. "Related SNA code" indicates the closest analog to the disaggregated flow in the standard national accounts. a: disaggregated flow excludes product taxes, unlike SNA (Section IV.B.1). b: disaggregated flow measures output of consumer-facing producers in terms of sales, whereas SNA measures trade margins (Section IV.C.1). c: disaggregated flow and SNA position defined slightly differently (Footnote 11). d: disaggregated flow and SNA position defined slightly differently (Footnote 12). e: foreign spending higher in disaggregated flow than in SNA (Footnote 13).

To visualize the new disaggregated system, we construct a disaggregated analog to the circular flow of money. The classical circular flow contains a handful of nodes for groups at the national level. In contrast, our disaggregated circular flow in Figure III contains 5,390 nodes, one for each region-by-industry consumer and producer cell in Denmark.[14] Nodes lying in the same region share the same color. Node size on the plot is proportional to the cell's economic size, measured as the square root of all inflows into the cell.

The disaggregated system allows us to calculate the total amount of money flowing from each cell to every other cell through any type of transaction (consumer spending, income, intermediate trade). We visualize cell-to-cell flows by drawing a link from a source cell to a receiving cell if the flow between the two cells is among the five largest outflows for the source cell or the single largest inflow for the receiving cell.[15] We let an algorithm ("ForceAtlas 2", Jacomy et al. 2014) arrange the nodes, so that cells with larger flows are located next to each other.

Figure III: The disaggregated circular flow of money



Notes: Nodes are all consumer and producer cells in Denmark. We draw a link between two cells if the cell-to-cell flow is among the top five outflows per cell or the top inflow per cell. Nodes are arranged according to the "ForceAtlas 2" algorithm of Jacomy et al. (2014).

A few patterns are noteworthy. First, nodes in the same region (and color) are close together,

[14] We deliberately do not plot nodes for the government, capital accumulation, and foreign cells, since these cells are orders of magnitude larger than individual consumer and producer cells, making it difficult to see flows between consumer and producer cells

[15] The plot looks similar but more crowded if these restrictions are relaxed.

highlighting the importance of geographic location for economic interactions. Second, the shape of the graph is strikingly similar to the geography of Denmark: the large cluster of nodes on the left is the continental Western part of Denmark, the small cluster in the bottom is the central island Funen (with major city Odense), and the large Eastern island Zealand with the capital Copenhagen (red) lies on the right. Third, there are also notable deviations from the geography of Denmark. For example, the producer cells for airline and shipping industries in Copenhagen (red central nodes) sit much more centrally in the plot than the location of Copenhagen would suggest, mirroring their central position in the disaggregated circular flow.

V Facts Based on Disaggregated Economic Accounts in Denmark

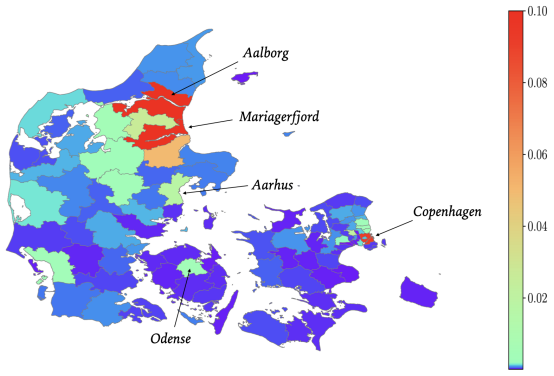
We analyze the newly measured system of disaggregated economic accounts and present several facts about how money circulates through the Danish economy.

V.A Most Flows are Regionally Concentrated and Sparse

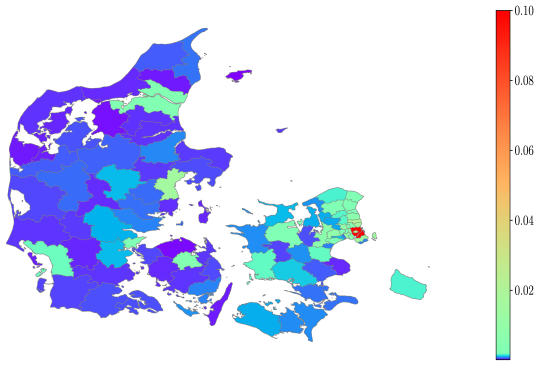
We begin by exploring the role of geographic distance for domestic consumer spending flows in the disaggregated system. Figure IV plots the geographical distribution of domestic consumer spending for restaurant workers living in rural Mariagerfjord (Panel IVa) and in the capital city Copenhagen (Panel IVb). In both cases, we see that spending is regionally concentrated and falls with distance. However, deviating from the pattern of regional concentration, a large share of spending from Mariagerfjord goes to the large cities Aalborg, Aarhus, Odense, and Copenhagen. In contrast, little spending from Copenhagen flows to rural regions and essentially no spending to far-away rural regions.

Figure IV: Spending shares across regions: Examples

(a) Restaurant workers in Mariagerfjord



(b) Restaurant workers in Copenhagen



Note: We plot the fraction of spending received by producers in each region, for spending by restaurant workers in Mariagerfjord (Panel a) and Copenhagen (Panel b). The scale is truncated at 0.10. We include only spending to Danish producers in the calculations.

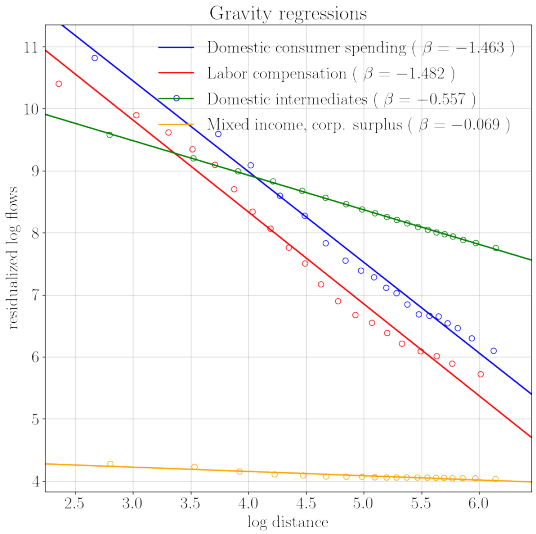
We formalize the role of distance by estimating gravity specifications for consumer spending. We regress spending from a consumer cell i on a producer cell j on distance and a full set of consumer and producer cell fixed effects,

$$\log \text{ flow}_{i \to j} = \alpha_i + \delta_j + \beta \times \log \text{ distance}_{ij} + \epsilon_{ij}. \tag{2}$$

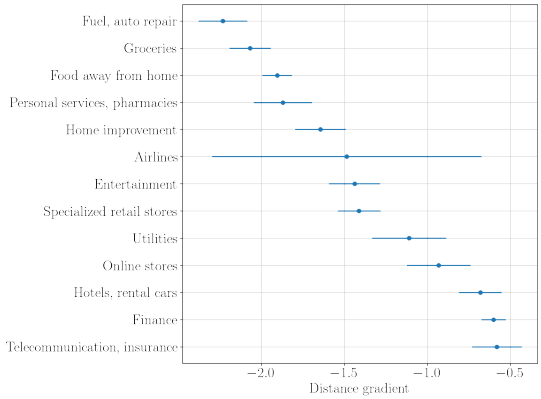
The fixed effects control for all cell-specific factors that determine total spending out of or into a cell, such as population, average income, or industry. Panel a of Figure V shows a binned scatter plot and Table A.V the associated regression table. The relation between spending and distance is negative and close to linear, with a gradient of -1.5. Panel b shows that there is significant heterogeneity depending on the types of goods purchased (see also Agarwal, Jensen and Monte 2020). An advantage of the new consumption matrix is that it allows studying the full set of consumption goods, not just goods purchased with cards. Spending on fuel, auto repair, and groceries, which requires regular physical transfer of goods, falls steeply with distance. Spending on telecommunications and insurance, which requires no physical transfers, and spending on hotels and rental cars, which is often part of travel, still decline with distance but to a lower extent.

Figure V: Gravity

(a) Across types of flows



(b) Across receiving industries



Notes: The panels show binned scatter plots estimated separately for different flows (panel a) and receiving industries (panel b) using equation (2). We plot averages of $\overline{\alpha} + \overline{\delta} + \beta \times \log \text{ distance}_{ij} + \epsilon_{ij}$ for 20 evenly-sized bins of log distance ij. Flows are in log DKK and distance is in log km. Flows are residualized using fixed effects for source cells and receiving cells. We add the mean of log flows to the residualized variable, which facilitates interpretation of level differences. We exclude observations with zero distance or zero flow. We include only spending to Danish producers in the calculations. Distance is measured using haversine distance between region centers. The solid line is the line of best fit, estimated using the cell-level data.

We also display gravity results for labor compensation, intermediate inputs, and mixed plus surplus income (sum of mixed income, surplus of corporate producers, and surplus of owner-occupied housing payments to consumers). Labor compensation displays an average distance gradient close to that of consumer spending, although the labor compensation gradient is flatter at low and high distances and steeper for middle range distances. Trade in intermediate inputs declines less with distance, as the gravity coefficient in the cargo data is -0.6. Finally, mixed plus surplus income require only financial transactions, which are easily sent across space.

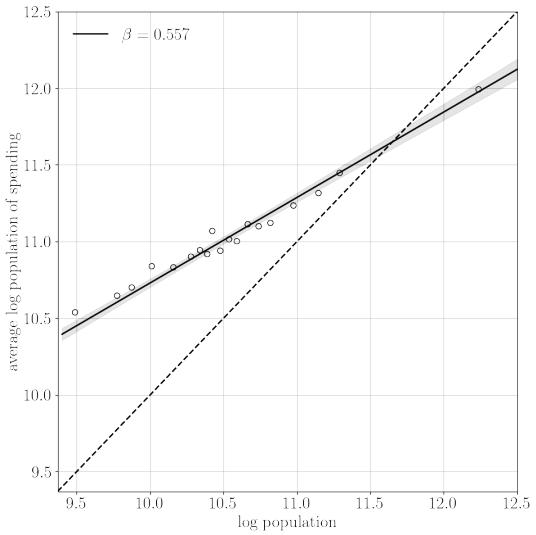
In general, the disaggregated flows are sparse. For most cells, outgoing flows tend to go to a few receiving cells while incoming flows tend to come from a few source cells. For instance, the largest 10% of the 7.3 million pairwise consumer spending flows account for over 99.5% of aggregate consumer spending (as spending mostly goes to nearby regions as well as near and far cities).

V.B Consumers "Spend Upward in Density"

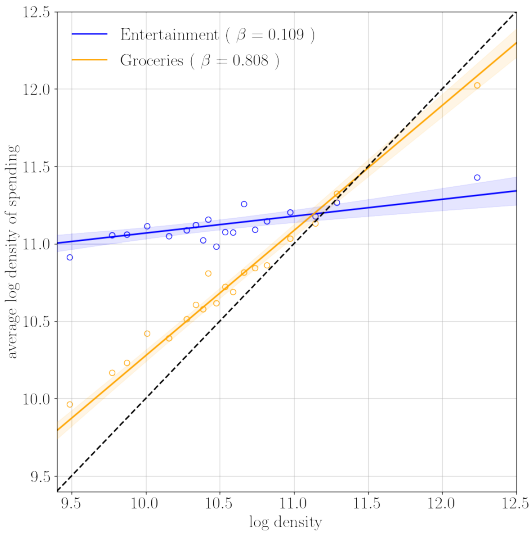
Figure IV already illustrated that a larger share of spending goes to cities than to rural regions. We next show that spending within Denmark systematically travels "upward in density," from sparsely populated to dense regions. Panel a of Figure VI plots the average population density of receiving producer cells by the density of consumer cells. Almost all points lie above the 45 degree line, indicating that consumers tend to spend in denser regions than their home region.

Figure VI: Population density of home and receiving region

(a) All categories



(b) Only entertainment and groceries



Notes: The panels show the average log population density of regions receiving consumer spending (vertical axis) for 20 evenly-sized bins of consumers' home region log population density (horizontal axis). We include only spending to Danish producers in the calculations. Distance is measured using haversine distance between region centers. The solid line is the line of best fit, estimated using the cell-level data.

There is heterogneity in the degree of "spending upward in density" depending on the receiving industry, shown in Panel VIb. Consumers spend on groceries in their home region to a relatively large extent, implying that the estimated line for groceries is close to the 45 degree line, although still above it. In contrast, spending on entertainment happens in similar locations for almost all consumers, implying that the estimated line for entertainment is relatively flat. Since these locations are more urban than the average region, most bins lie above the 45 degree line.

"Spending upward in density" represents a statistically and economically significant deviation from the standard gravity equation. We show that this pattern cannot be explained by differences in size, industrial composition, or other attributes of cities by estimating a modified gravity equation,

$$\log \text{flow}_{i\to j} = \alpha_i + \delta_j + \beta \times \log \text{distance}_{ij} + \gamma_{\text{destination}} \times \log \text{distance}_{ij} \times \log \text{density}_i + \gamma_{\text{origin}} \times \log \text{distance}_{ij} \times \log \text{density}_j + \epsilon_{ij}. \quad (3)$$

Table III shows that spending into dense regions is less sensitive to distance, indicating the consumers from near and far tend to spend in cities whereas spending in rural areas is more likely

to come from closer consumers (column 2). In addition, we find that spending by rural consumers is less sensitive to distance, indicating that rural consumers are more likely to spend at distant retailers, relative to urban consumers (column 3). Moving from the lowest to the highest density consumer region, the distance effect falls from -1.99 to -1.18.

Table III: Gravity regressions with interactions

	log spending			
	(1)	*(2)*	*(3)*	*(4)*
log distance	-1.463^{***}	-1.781^{***}	-1.765^{***}	-2.098^{***}
	(0.028)	(0.119)	(0.046)	(0.126)
log distance × log dest. pop.		0.573^{***}		0.585^{***}
		(0.209)		(0.208)
log distance × log origin pop.			0.551^{***}	0.564^{***}
			(0.068)	(0.068)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Observations	2561029	2561029	2561029	2561029
R^2	0.332	0.333	0.333	0.334
F stat.	1268417.4	637166.1	637012.6	426738.8

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Distances are measured as haversine distances between region centers. For the interaction terms, we normalize log density to range from 0 to 1 across regions. The interaction coefficient therefore shows the change in the distance gradient when moving from lowest to highest density. We include only spending to Danish producers in the calculations.

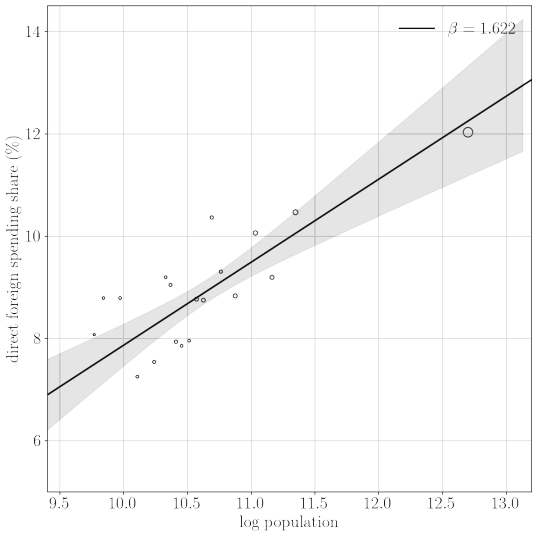
V.C Rich, Urban Consumers Spend More Abroad

We next document which Danish consumers spend abroad. Figure VII plots direct foreign spending shares by consumer cells' population density and average spending. Across the two panels, we see that density and average spending predict direct foreign spending positively: rich, urban consumers spend a larger share of their total spending abroad.[16]

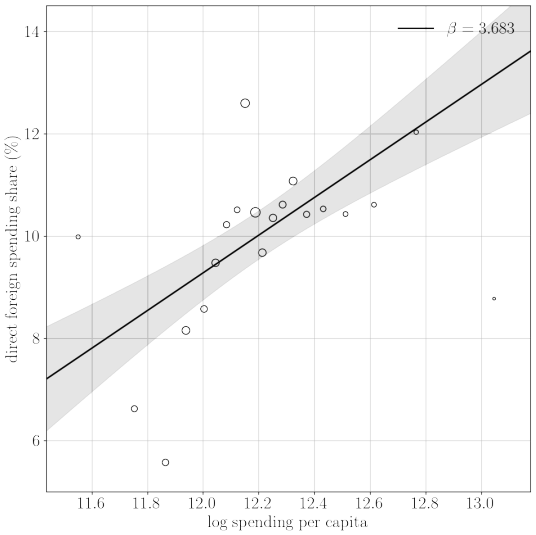
[16]In Table A.VII, we show that population and average spending predict foreign spending also when controlling for driving or geographic distance to the Danish border.

Figure VII: Foreign spending by population and total spending

(a) By population



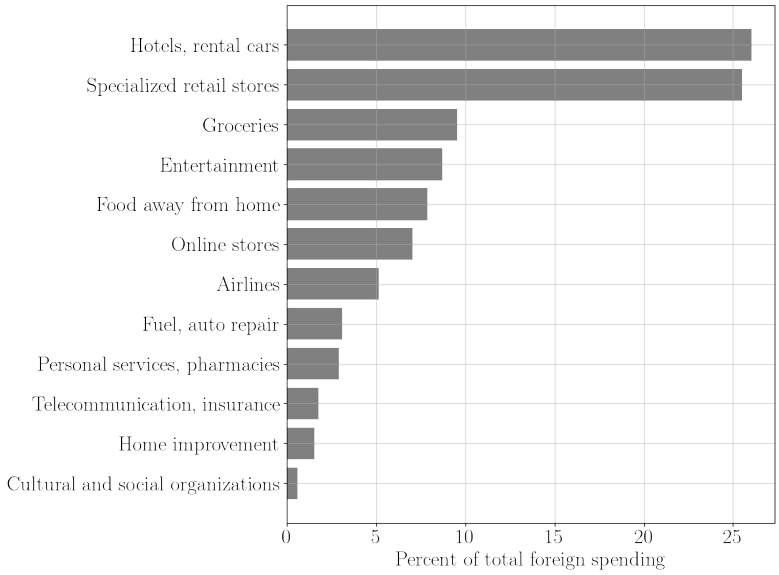
(b) By total spending per capita



Notes: The figures show the direct foreign spending share of consumer cells by their home region's log population density. Both plots are binned scatter plots with 98 bins mirroring the number of regions in our analysis. Weights are proportional to the number of bank customers in a municipality.

The spending matrix also contains information on which types of goods are purchased abroad. Figure VIII shows that spending on hotels, rental cars, and airlines accounts for around 26 percent of total foreign spending, entertainment and food away from home for around 15 percent, and spending on groceries and specialized retail stores for about 35 percent.

Figure VIII: Foreign spending shares by spending category



Taken together, the facts established so far imply that rural consumers allocate a larger share of their spending to their home region and cities. In contrast, urban consumers allocate a larger share to their home region, other cities, and abroad. These patterns suggest that shocks to the spending of rural consumers affect rural and urban producers more strongly, while shocks to urban consumers impact urban and foreign producers. We will explore the asymmetric consequences of such shocks below using a formal model. Before doing so, however, we study how other flows, apart from consumer spending, differ between rural and urban regions.

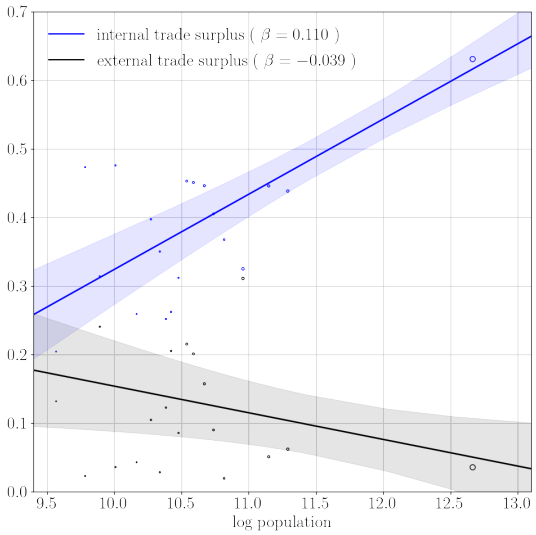
V.D Regional BOP and GDP: Rural regions export abroad, urban regions domestically

Having studied consumer spending patterns by density, we now investigate all other flows that enter and leave different regions. We use our disaggregated economic accounts to construct regional balance of payments (BOP) and GDP statistics. We compute five cross-regional flows, which jointly add up to zero: (i) *external trade surplus*, defined as a region's trade surplus with the rest of the world (producer exports - producer imports - foreign consumer spending); (ii) *internal trade surplus*, the trade surplus with the rest of the Danish economy (domestic intermediates + domestic consumer spending + domestic government spending + domestic capital accumulation spending (all inflows/sales) - domestic intermediate outflows/purchases); (iii) *net factor payments* (net of: labor compensation received + producer dividends, mixed income, and surplus received + natural resources rents received); (iv) *net transfers* (net of: social benefits received - taxes - social contributions); and (v) *net borrowing* (negative of: consumer interest, transfers, and saving + net producer interest, transfers, and saving). We compute (nominal) regional GDP as total production

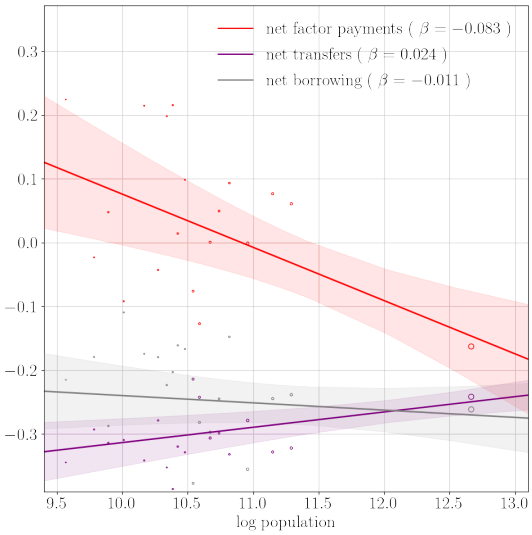
value net of domestic and foreign intermediate purchases.

Figure IX: Regional balance of payments (normalized by GDP)

(a) External and internal trade surplus



(b) Other items of regional balance of payments



Notes: The panels show the five components of the regional balance of payments (BOP) statistics, normalized by regional GDP. The five components add up to zero for each region. The "internal trade surplus" includes investment and government spending, see footnote 17.

Figure IX shows how each of the five components, normalized by regional GDP, varies with a region's population density. The external trade surplus is largest in rural regions, indicating that rural regions export a significant fraction of their output to foreign countries (Panel IXa). The internal trade surplus is largest in urban regions, indicating that those regions export a large share of their output to the rest of Denmark.[17] Net transfers and net borrowing do not vary significantly, whereas net factor payments flow into rural regions and out of urban regions, on average (Panel IXb). Taken together, these facts reveal that, on net, money flows into rural cells from abroad, while it flows out of rural cells to more urban domestic cells (e.g., through "spending upward in density").

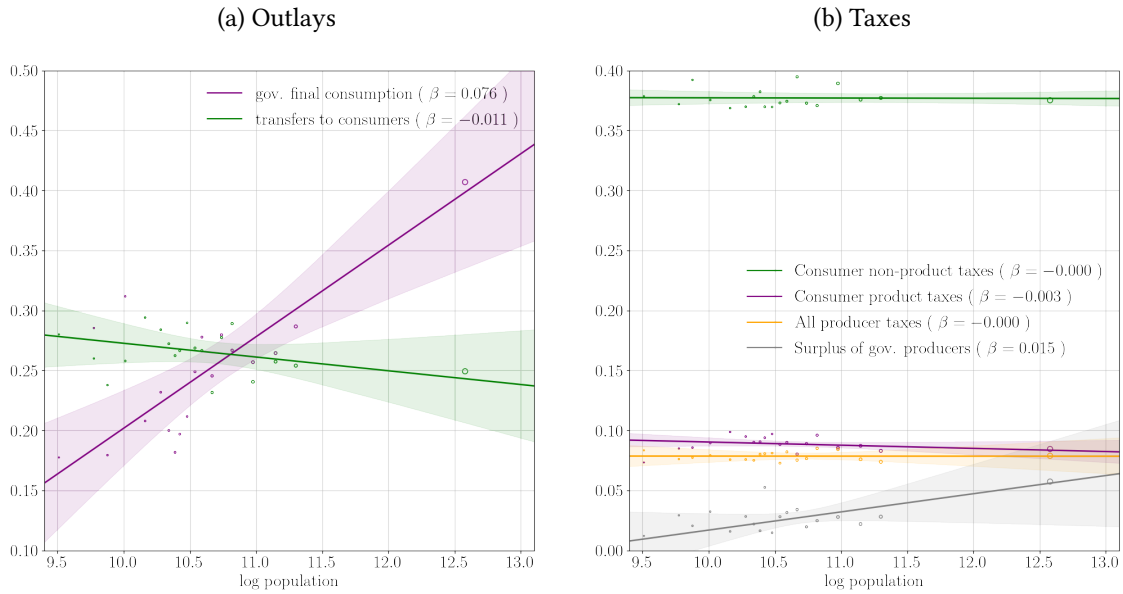
V.E Government redistributes towards rural regions

The regional BOP statistics also allow us document which regions rely more on transactions with the government transfers. The government purchases a larger share of urban regions' output and pays a larger share of benefits to rural regions. However, the overall size of purchases is larger

[17]The internal trade surplus does not average out to zero in Panel IXa because it includes capital accumulation spending and government spending, which are paid for by the capital accumulation and government cells, respectively. Since those two cells do not have a region assigned to them, they do not show up in Panel IXa.

than that of transfers, implying that, in total, the government plays a more important role in urban regions.

Figure X: Government consumption and benefits across regions (by region GDP)

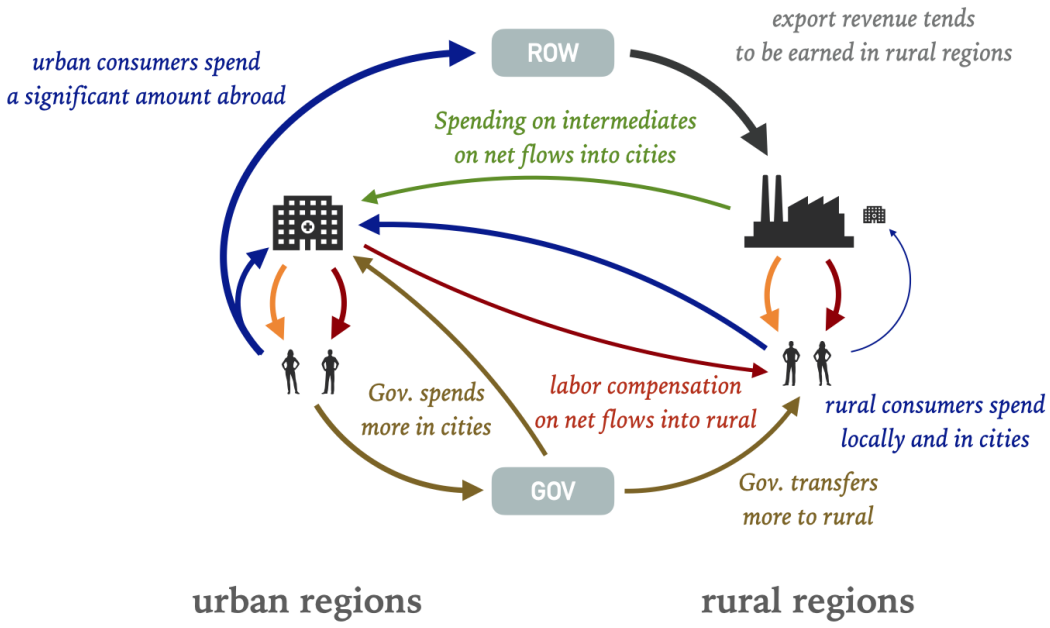


Notes: The panels show all flows between individual regions and the government, normalized by region GDP, as function of a region's density.

V.F Stylized overview of the circular flow by density

Figure XI provides a stylized overview of the facts we have documented. Money enters the economy via exports, predominantly flowing into the rural regions of Denmark and contributing to the income of rural workers. Rural consumers then spend their income "upward in density," implying that, on net, money leaves rural regions toward more urban regions. There, it is earned by consumers working predominantly in service industries. Part of urban income is redistributed by the government toward rural regions. The remainder is either spent locally or abroad, closing the circular flow.

Figure XI: Stylized overview of the circular flow by density



VI A Model of Disaggregated Economics Accounts

We have so far shown how to measure a system of disaggregated economic accounts and that it reveals new facts on the heterogeneous flows connecting different consumer and producer groups. We now turn toward understanding how disaggregated accounts inform macroeconomic analyses. On their own, the data illustrate the complex interdependencies making up the macroeconomy. However, the data alone do not reveal how shocks affect different cells in general equilibrium where producers and consumers optimize and prices endogenously adjust.

In this section, we therefore construct a model that matches the structure of the disaggregated accounts. Our aim is to develop a flexible model that allows researchers to investigate how shocks with heterogeneous direct incidence across consumer and producer cells propagate across the economy in the medium run. In general, we highlight that a disaggregated approach—the combination of disaggregated data and model—strongly enriches our understanding of how shocks affect the economy, both in terms of distributional consequences across cells but also in terms of aggregate outcomes. Moreover, the disaggregated approach holds promise because it permits governments to predict the distributional and aggregate consequences of policies. In two specific applications at the end of the paper, we will illustrate the types of insights generated by the disaggregated approach by focusing on trade shocks in Denmark.

VI.A Overview

We model a small open economy composed of many region-industry cells, which are linked via cross-border trade in intermediate goods, factor payments, consumer spending, profit flows, taxes and transfers, and government spending. The model is static and builds on previous work, among others by Caliendo and Parro (2015) and Baqaee and Farhi (2019*b*). It is supposed to capture the medium-run behavior of an economy and abstracts from nominal rigidities. Because we focus on medium-run effects, we abstract from the capital accumulation cell, i.e., there is no saving and investment in the economy. We leave a dynamic extension of the model to future work.

There exists a set $\mathcal{I}$ of consumer cells (corresponding to our region-by-industry of work cells in the data, indexed by $i \in \mathcal{I}$) and a set $\mathcal{J}$ of producer cells (corresponding to our region-by-industry cells in the data, indexed by $j \in \mathcal{J}$). The (large) representative consumer that populates the rest of the world is labeled $i = \mathcal{R} \notin \mathcal{I}$ and the foreign representative producer cell is $j = \mathcal{R} \notin \mathcal{J}$. We write $\mathcal{I} \cup \{\mathcal{R}\}$ and $\mathcal{J} \cup \{\mathcal{R}\}$ whenever the rest of the world is included in the set of indices at consideration. Each cell lies in an industry s, which is an element of the set $\mathcal{S}_{\mathcal{I}}$ for consumers and the set $\mathcal{S}_{\mathcal{J}}$ for producers. We denote by $s(i)$ and $s(j)$ the industries associated with cells i and j, respectively.

VI.B Setup

We describe consumer and producer cells in turn.

Consumer cells: utility. A mass L_i of consumers lives within each consumer cell i, all sharing the same preferences and maximizing utility over consolidated consumption C_i,

$$U_i = \log C_i = \alpha_{i\mathcal{R}} \log c_{i\mathcal{R}} + (1 - \alpha_{i\mathcal{R}}) \sum_{s \in \mathcal{S}_{\mathcal{J}}} \alpha_{is} \log \left(\sum_{j:s(j)=s} \alpha_{isj}^{\frac{1}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{4}$$

where $\sum_{s \in \mathcal{S}_{\mathcal{J}}} \alpha_{is} = \sum_{j:s(j)=s} \alpha_{isj} = 1$. $c_{i\mathcal{R}}$ is consumer cell i's consumption of foreign goods and c_{ij} is consumer cell i's consumption of goods produced in producer cell j. This is a nested CES utility, with a Cobb-Douglas specification across industries and a CES specification with elasticity $\sigma > 0$ within industries.

Consumer cells: labor supply. Consumer industries $\mathcal{S}_{\mathcal{I}}$ are of two types: *work industries* or *non-work industries.* If consumer cell i is in a work industry, it has a total labor endowment of N_i, for which it earns a wage W_i. Total labor income is then $N_i W_i$. Consumer cells in non-work industries do not supply labor.

Consumer cells: budget constraint. In addition to labor income (**??**), consumer cells also earn profit income $\sum_{j\in\mathcal{J}} \kappa_{ij}\Pi_j$. The total profit of producer cell j is Π_j. The share of producer cell j's profits earned by consumer cell i is $\kappa_{ij} \in [0,1]$. Further, consumer cell i pays a proportional income tax rate τ_i, a domestic product tax τ_i^{vat}, and receives a government transfer T_i. Total nominal pre-tax income of consumer cell i is therefore

$$Y_i = N_i W_i + \sum_{j\in\mathcal{J}\cup\{\mathcal{R}\}} \kappa_{ij}\Pi_j + T_i. \tag{5}$$

The consolidated budget constraint of cell i is then

$$\sum_{j\in\mathcal{J}\cup\{\mathcal{R}\}} \left(1+\tau_i^{vat}\right) P_j c_{ij} \leq (1-\tau_i)\, Y_i. \tag{6}$$

Producer cells. There is a representative firm in each producer cell $j \in \mathcal{J}$. It produces quantity Q_j of its good j with technology

$$Q_j = Z_j \left(\sum_{i\in\mathcal{I}} \lambda_{ij}^{\frac{1}{\zeta}} N_{ij}^{\frac{\zeta-1}{\zeta}}\right)^{\frac{\zeta}{\zeta-1}\gamma_j^N} \prod_{s\in\mathcal{S}_\mathcal{J}} \left(\sum_{j'\in\mathcal{J}\cup\{\mathcal{R}\}:s(j')=s} \omega_{jsj'}^{\frac{1}{\eta}} X_{j'j}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}\omega_{js}\gamma_j^X}, \tag{7}$$

where $\sum_s \omega_{js} = \sum_{j':s(j')=s} \omega_{jsj'} = \sum_i \lambda_{ij} = 1$. Producer cell j uses N_i units of labor supplied by consumer cell i in the same industry, $s(i) = s(j)$, with total j-specific labor income share γ_j^N.[18] Cell j uses $X_{j'j}$ units of intermediate goods, purchased from firm cell j', for each $j' \in \mathcal{J}\cup\{\mathcal{R}\}$. The elasticity of substitution between producer cells within the same industry is $\eta > 0$. The elasticity of substitution between labor supplied by different consumer cells is $\zeta > 0$. Total factor productivity is given by Z_j. We assume weakly decreasing returns to scale, $\gamma_j^N + \gamma_j^X \leq 1$, allowing for potentially fixed factors. Pretax profits are given by

$$\Pi_j^{pre} = P_j Q_j - \sum_{i\in\mathcal{I}} W_i N_{ij} - \sum_{j'\in\mathcal{J}\cup\{\mathcal{R}\}} P_{j'} X_{j'j}. \tag{8}$$

Profits are taxed at the corporate tax rate $\tau_j^{corp} \geq 0$. After-tax profits are then

$$\Pi_j = \left(1-\tau_j^{corp}\right)\Pi_j^{pre}. \tag{9}$$

[18] We set $\lambda_{ij} = 0$ if $s(i) \neq s(j)$.

Rest of the world. Domestic consumers and producers buy foreign goods at price $P_{\mathcal{R}}$. Export demand for domestically produced goods is

$$x_j = \tilde{x}_j \cdot P_j^{-\gamma}, \tag{10}$$

where the elasticity of exports to the terms of trade is equal to $\gamma > 0$. Trade will be balanced in equilibrium.

Government. The government pays (nominal) transfers T_i and spends on domestically produced goods G_j, financed by tax revenue. The government budget constraint is given by

$$\sum_{j\in\mathcal{J}} P_j G_j + \sum_{i\in\mathcal{I}} (1-\tau_i)\, T_i = \sum_{i\in\mathcal{I}} \tau_i W_i \sum_{j\in\mathcal{J}} N_{ij} + \sum_{i\in\mathcal{I}} \tau_i \sum_{j\in\mathcal{J}} \kappa_{ij} \Pi_j + \sum_{j\in\mathcal{J}} \tau_j^{corp} \Pi_j^{pre} + \tau^{vat} \sum_{j\in\mathcal{J}\cup\{\mathcal{R}\}} P_j c_{ij}. \tag{11}$$

Equilibrium. We define equilibrium as in any competitive model with flexible prices.

Definition 1. A *competitive equilibrium* in the economy consists of prices and wages $\{P_j, W_i\}$ and an allocation $\{Q_j, N_{ij}, X_{j'j}, \Pi_j, T_i, G_j, Y_i, c_{ij}, x_j\}$ such that (a) income is given by (5); (b) all consumer cells maximize utility (4) subject to (6); (c) all producer cells maximize profits (9); (d) the government's budget (11) is balanced; (e) labor markets clear for each producer cell,

$$N_i = \sum_j N_{ij}; \tag{12}$$

and (f) the goods market clears for each producer cell,

$$Q_j = \sum_{\tilde{j}\in\mathcal{J}} X_{j\tilde{j}} + x_j + \sum_i c_{ij} + G_j.$$

VI.C Calibration

We calibrate the model to match the disaggregated flows measured in Section IV. For the baseline calibrated equilibrium, we normalize all prices and wages to 1, that is, $P_j = W_i = P_{\mathcal{R}} = 1$. This is without loss since (a) the units of production can always be chosen to set $P_j = 1$ (by choosing Z_j); and (b) the units of labor supply can always be chosen to obtain $W_i = 1$ (by choosing N_i).[19] With this normalization, the average spending share of consumer cell i on producer cell j can be

[19]Note that prices and wages will change in our comparative statics results below. These are mere normalizations and do *not* represent rigidities.

expressed as

$$\frac{P_j c_{ij}}{\sum_j P_j c_{ij}} = \begin{cases} \alpha_{i\mathcal{R}} & \text{if } j = \mathcal{R} \\ (1 - \alpha_{i\mathcal{R}})\, \alpha_{is}\alpha_{isj} & \text{if } j \neq \mathcal{R} \end{cases}.$$

We calibrate the average spending shares $\alpha_{i\mathcal{R}}, \alpha_{is}, \alpha_{isj}$ of consumer cell i to match the corresponding spending shares from the disaggregated consumer spending flows (both domestic and foreign spending).

Similarly, we measure the share of producer cell j's wage bill going to consumer cell i, $\frac{N_{ij}W_i}{\sum_i N_{ij}W_i} = \lambda_{ij}$, in the disaggregated labor compensation flows. The distribution of profit income κ_{ij} is calibrated by matching, for each producer cell j, the distribution of dividend, mixed income, and surplus of owner-occupied housing flows across consumer cells. We calibrate τ_i to match the ratio of consumer non-product taxes and social contributions to total income for each consumer cell i; and τ_i^{vat} to match the ratio of consumer product taxes to total consumer cell income. We calibrate τ_j^{corp} to match, as fraction of j's total output $P_j Q_j$, the sum of: dividends and surplus paid to the government, producer product and non-product taxes paid, less subsidies.

We choose cell-level transfers T_i to match social benefits received by i plus adjustments for pension entitlements, as share of total government outlays. We choose government spending G_j to match government final consumption expenditure on producer cell j, as share of total government outlays.

We set the labor share γ_j^N in line with the ratio of labor compensation paid by producer cell j to output by producer cell j, and γ_j^X in line with the ratio of intermediate input purchases to output. We set intermediate input shares $\omega_{js}\omega_{js'} = \frac{P_{j'}X_{j'j}}{\sum_{j'} P_{j'}X_{j'j}}$ in line with the disaggregated input-output flows. We choose the relative magnitudes of $\tilde{x}_j$ to match the distribution of exports across producer cells j. We choose the level of $\tilde{x}_j$ to match aggregate GDP.

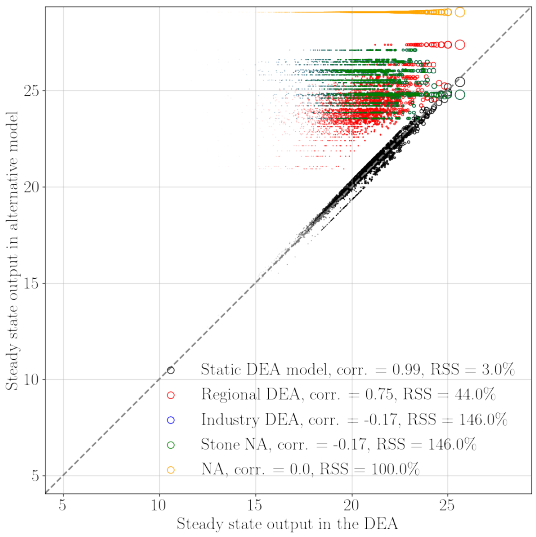
This leaves us with a calibrated model economy that matches the disaggregated accounts we measured in the previous sections. The only difference arises from the lack of dynamics, and hence of a capital accumulation sector, in our model economy. In Figure XII, we illustrate the proximity by comparing nominal output across producer cells as well as domestic spending across consumer cells. As the black markers show, the correlation between model and data is very close to 1 for those variables, with a residual sum of squares of only 3% of the total variation in the data.

As a comparison, we also include the predictions of four alternative models in Figure XII. All four are calibrated like the model here, just at different levels of aggregation. "NA" refers to "national accounts", a model with only a single consumer cell and a single producer cell. "Stone NA" refers to a model with only a single consumer cell and producer cells differentiated by industry, but not by region. "Industry DEA" refers to model in which both consumer cells and producer cells are differentiated by industries of work, and industries of production, respectively. "Regional

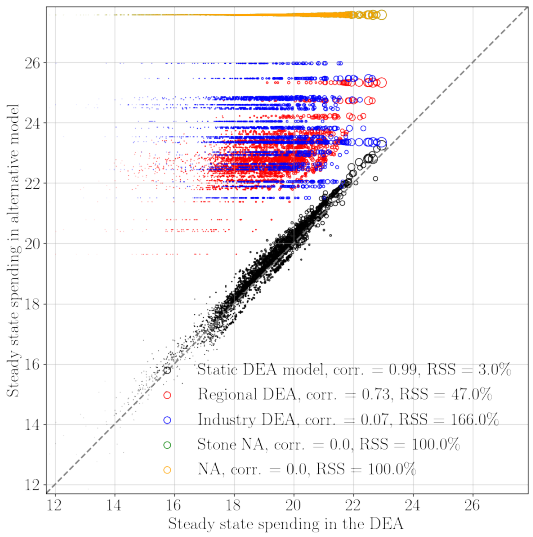
DEA" refers to a model in which both consumer cells and producer cells are differentiated by region of residence, and region of production, respectively. As can clearly be seen, the fit for those alternative models is significantly worse compared to the static DEA model we developed in this section.

Figure XII: Match between model economy and disaggregated accounts

(a) Nominal output by producer cell



(b) Domestic consumer spending



Notes: Panel a shows a scatter plot of log producer cell output in the model relative to log producer cell in the disaggregated accounts (bubble size is proportional to producer cell output in the disaggregated accounts). Panel b shows a binned scatter plot of log consumer spending in the model versus the disaggregated accounts (bubble size is proportional to domestic spending in the disaggregated accounts). "corr" stands for correlation. "RSS" stands for the residual sum of squares. It captures the variance in the data not explained by the model, $\text{var}(y^{data} - y^{model})$, for any variable y, in percent of variance in the data $\text{var}(y^{data})$. If larger than 100%, this implies that the variation in the model is not useful to explain the data.

The only parameters that cannot directly be calibrated from the disaggregated accounts are the four elasticities σ, ζ, η, and γ. Our baseline choice is $\sigma = \zeta = \gamma = 2$, $\eta = 1$. The values for σ, ζ, and η are from Adão et al. (2022) and for γ from Boehm, Levchenko and Pandalai-Nayar (2022). We provide robustness to those choices in Section **??**.

VII Conclusion

The idea of a comprehensively disaggregated circular flow of money goes back to at least 18th century France. Richard Cantillon and François Quesnay presented systems of measurement that assigned all citizens to a handful of distinct groups and then showed how money traveled across these groups through consumption, income, and production links. Conceptual and measurement

challenges meant that this idea lay dormant for almost 200 years, until Wassily Leontief developed ways to measure cross-industry production links and Richard Stone integrated these flows into a system of national accounts. Despite these seminal contributions, the core vision of Cantillon and Quesnay remained unfulfilled. Until now, no framework exists that allows researchers to comprehensively measure flows between small subgroups of consumers and producers, so that (i) bilateral transactions add up to national aggregates, (ii) accounting identities across different levels of aggregation are satisfied, (iii) rich heterogeneity in bilateral flows across different consumer and producer groups is accurately recorded, and (iv) the full circular flow of consumption, income, and production links is measured.

We propose and implement a new system of measurement that fulfills these criteria. The advent of administrative microdata on consumption and income transactions as well as modern computing tools allow us to take a step toward realizing the vision of a fully disaggregated circular flow. We describe general guidelines and practical tools that allow researchers to measure a disaggregated system using either bottom-up or top-down approaches. We implement the guidelines in Denmark and measure a disaggregated system using a variety of aggregate and microdata sources.

The new data allow us to present facts on how money circulates across the circular flow. We show, for instance, how the impact of distance on flows depends on the nature of transactions. We also highlight that the consumer spending behavior of urban consumers differs systematically from rural consumers, as spending tends to flow from rural to urban regions and from urban regions abroad.

We show that these facts have practical bite by combining a disaggregated model with the new data. We show that the welfare gains from trade are larger for urban residents because they receive an income increase from the spending of rural consumers, while rural consumers do not receive an income boost from urban consumers in return. Ultimately, after the initial gains from trade have circulated through the full disaggregated system, the income gains to urban residents are larger than those of rural consumers. Overall, we underscore that a disaggregated system enriches our understanding of the distributional effects of economic shocks.

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica*, 80(5): 1977–2016.

Adão, Rodrigo, Paul Carrillo, Arnaud Costinot, Dave Donaldson, and Dina Pomeranz. 2022. "Imports, Exports, and Earnings Inequality: Measures of Exposure and Estimates of Incidence." *Quarterly Journal of Economics*.

Agarwal, Sumit, J. Bradford Jensen, and Ferdinando Monte. 2020. "Consumer Mobility and the Local Structure of Consumption Industries."

Aladangady, Aditya, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson, Paul Lengermann, and Claudia Sahm. 2021. "From Transaction Data to Economic Statistics: Constructing Real-Time, High-Frequency, Geographic Measures of Consumer Spending." *Big Data for Twenty-First-Century Economic Statistics*, 115–145. University of Chicago Press.

Andersen, Asger Lau, Emil Toft Hansen, Niels Johannesen, and Adam Sheridan. 2022. "Consumer Responses to the COVID-19 Crisis: Evidence From Bank Account Transaction Data." *Scandinavian Journal of Economics*.

Baker, Scott R. 2018. "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data." *Journal of Political Economy*, 126(4): 1504–1557.

Baker, Scott R., and Lorenz Kueng. 2022. "Household Financial Transaction Data." *Annual Review of Economics*.

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

Baqaee, David Rezza, and Emmanuel Farhi. 2019*b*. "Macroeconomics with Heterogeneous Agents and Input-Output Networks."

Baqaee, David Rezza, and Emmanuel Farhi. 2020*b*. "Productivity and Misallocation in General Equilibrium." *Quarterly Journal of Economics*, 135(1): 105–163.

Barro, Robert J. 2021. "Double Counting of Investment." *Economic Journal*, 131(638): 2333–2356.

Bernard, Andrew B, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes. 2022. "The Origins of Firm Heterogeneity: A Production Network Approach." *Journal of Political Economy*, 130(7): 1765–1804.

Bigio, Saki, and Jennifer La'O. 2020. "Distortions in production networks." *Quarterly Journal of Economics*, 135(4): 2187–2253.

Blanchet, Thomas, Lucas Chancel, Ignacio Flores, Marc Morgan, Facundo Alvaredo, Anthony B. Atkinson, Luis Bauluz, Matthew Fisher-Post, Bertrand Garbinti, Jonathan Goupille-Lebret, Clara Martínez-Toledano, Theresa Neef, Thomas Piketty, Anne-Sophie Robilliard, Emmanuel Saez, Li Yang, and Gabriel Zucman. 2021. *Distributional National Accounts Guidelines.* World Inequality Lab.

Boehm, Christoph E, Andrei A Levchenko, and Nitya Pandalai-Nayar. 2022. "The Long and Short (Run) of Trade Elasticities."

Buda, Gergely, Vasco M. Carvalho, Stephen Hansen, Álvaro Ortiz, Rodrigo Tomasa, and José Vicente Rodríguez Mora. 2022. "National Accounts in a World of Naturally Occurring Data: A Proof of Concept for Consumption."

Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *The Review of Economic Studies*, 82(1): 1–44.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro. 2019. "Trade and labor market dynamics: General equilibrium analysis of the china trade shock." *Econometrica*, 87(3): 741–835.

Cantillon, Richard. 1755. *Essai sur la Nature du Commerce en Général.* Fletcher Gyles.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics*, 128(3): 967–1015.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and Opportunity Insights Team. 2020. "The Economic Impacts of COVID-19: Evidence From a New Public Database Built Using Private Sector Data." NBER Working Paper 27431.

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.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, Magne Mogstad, and Felix Tintelnot. 2021. "Trade and Domestic Production Networks." *Review of Economic Studies*, 88(2): 643–668.

Diamond, Rebecca, and Enrico Moretti. 2021. "Where Is Standard of Living the Highest? Local Prices and the Geography of Consumption." NBER Working Paper 29533.

Dunn, Abe, and Mahsa Gholizadeh. 2020. "The Geography of Consumption and Local Economic Shocks: The Case of the Great Recession."

Ehrlich, Gabriel, John C. Haltiwanger, Ron S. Jarmin, David Johnson, and Matthew D. Shapiro. 2020. "Reengineering Key National Economic Indicators." *Big Data for Twenty-First-Century Economic Statistics*, 25–68. University of Chicago Press.

Foster, William T. 1922. "The Circuit Flow of Money." *American Economic Review*, 12(3): 460–473.

Glaeser, Edward L., Hyunjin Kim, and Michael Luca. 2019. "Nowcasting the Local Economy: Using Yelp Data to Measure Economic Activity." *Big Data for Twenty-First-Century Economic Statistics*, 249–273. University of Chicago Press.

Horvath, Michael. 2000. "Sectoral shocks and aggregate fluctuations." *Journal of Monetary Economics*, 45(1): 69–106.

Huneeus, Federico. 2018. "Production Network Dynamics and the Propagation of Shocks."

Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. "ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software." *PloS one*, 9(6): e98679.

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics*, 3(2): 1–28.

Knight, Frank H. 1933. "The Economic Organization."

Kuznets, Simon, Lillian Epstein, and Elizabeth Jenks. 1941. *National Income and Its Composition, 1919-1938, Volume I.* National Bureau of Economic Research.

Lahn, J. J. O. 1903. *Der Kreislauf des Geldes und Mechanismus des Sozialebens.* Puttkammer & Mühlbrecht.

Leontief, Wassily. 1928. *Die Wirtschaft als Kreislauf.* Laupp.

Leontief, Wassily. 1966. *Input-Output Economics.* Oxford University Press.

Meade, James E., and Richard Stone. 1941. "The Construction of Tables of National Income, Expenditure, Savings and Investment." *Economic Journal*, 51(202/203): 216–233.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics*, 133(2): 553–609.

Quesnay, François. 1758. *Tableau Economique.*

Rubbo, Elisa. 2020. "Networks, Phillips Curves and Monetary Policy."

Stone, Richard. 1961. *Input-Output and National Accounts.* Organisation for European Economic Co-operation.

Vavra, Joseph. 2021. "Tracking the Pandemic in Real Time: Administrative Micro Data in Business Cycles Enters the Spotlight." *Journal of Economic Perspectives*, 35(3): 47–66.

vom Lehn, Christian, and Thomas Winberry. 2022. "The Investment Network, Sectoral Comovement, and the Changing US Business Cycle." *Quarterly Journal of Economics*, 137(1): 387–433.

Appendix for "Disaggregated Economic Accounts"

Appendix A Tables and Figures

Table A.I: Classification of industries

	Industry	Pays income to consumers	Produces output	Sells directly to consumers
1	Food away from home	Yes	Yes	Yes
2	Entertainment	Yes	Yes	Yes
3	Groceries	Yes	Yes	Yes
4	Personal services, pharmacies	Yes	Yes	Yes
5	Cars, fuel, car repair, public transport	Yes	Yes	Yes
6	Hotels, rental cars	Yes	Yes	Yes
7	Airlines	Yes	Yes	Yes
8	Telecommunication, insurance	Yes	Yes	Yes
9	Online stores	Yes	Yes	Yes
10	Utilities	Yes	Yes	Yes
11	Specialized retail stores	Yes	Yes	Yes
12	Home improvement	Yes	Yes	Yes
13	Consulting, information technology, media	Yes	Yes	No
14	Business administration and janitorial services	Yes	Yes	No
15	Manufacturing	Yes	Yes	No
16	Wholesale	Yes	Yes	No
17	Finance, real estate	Yes	Yes	Yes
18	Cultural and social organizations	Yes	Yes	Yes
19	Agriculture, mining	Yes	Yes	No
20	Construction	Yes	Yes	No
21	Shipping, transport	Yes	Yes	No
22	Out of workforce and others	Yes	No	No
23	Retired	Yes	No	No
24	Health	Yes	Yes	No
25	Students	Yes	No	No
26	Education	Yes	Yes	No
27	Public administration	Yes	Yes	No
28	Unemployed	Yes	No	No
29	Private landlords	No	Yes	Yes
30	Owner-occupied housing	No	Yes	Yes
31	Government-owned housing	No	Yes	Yes

Notes: The table lists the industry classification used throughout the paper.

Table A.II: Summary statistics on the population and Danske Bank sample

	Full population	Danske Bank sample
Number of adults	4,367,226	858,409
Mean age	48.56	49.97
Mean income (DKK)	298,834	281,039
Age distribution		
18-39	0.35	0.34
40-59	0.35	0.32
60+	0.30	0.34
Income distribution		
Quintile 1	0.20	0.22
Quintile 2	0.20	0.23
Quintile 3	0.20	0.21
Quintile 4	0.20	0.18
Quintile 5	0.20	0.17
Ratio of liquid assets to income distribution		
Quintile 1	0.20	0.20
Quintile 2	0.20	0.21
Quintile 3	0.20	0.19
Quintile 4	0.20	0.20
Quintile 5	0.20	0.20

Notes: The table compares summary statistics for the Danish population from administrative registers with our sample of Danske Bank customers.

Table A.III: Gravity regressions

	log spending *(1)*	log wages *(2)*	log profits *(3)*	log interm. inputs *(4)*
log distance	-1.478^{***}	-1.482^{***}	-0.016^{***}	-0.600^{***}
	(0.028)	(0.013)	(0.001)	(0.006)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Observations	2394700	1654401	5331473	5233780
R^2	0.336	0.279	0.003	0.073
F stat.	1208001.1	637191.3	14772.4	413933.5

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Distances are measured as haversine distances between region centers.

Table A.V: Gravity regressions

	log spending *(1)*	log labor comp. *(2)*	log intermediates *(3)*	log mixed inc., surplus *(4)*
log distance	-1.463^{***}	-1.482^{***}	-0.557^{***}	-0.069^{***}
	(0.028)	(0.013)	(0.005)	(0.011)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Observations	2561029	1654401	5544292	5468081
R^2	0.332	0.279	0.059	0.002
F stat.	1268417.4	637191.3	348752.8	12746.0

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Distances are measured as haversine distances between region centers.

Table A.VII: Direct foreign spending shares

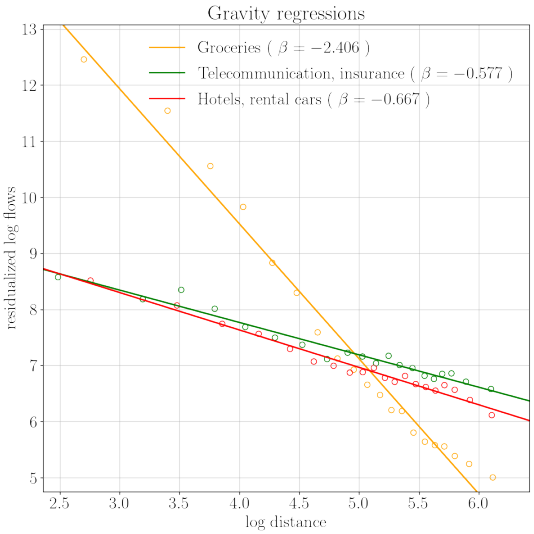
	direct foreign spending share (%)							
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
log population	1.622^{***}		1.766^{***}			1.664^{***}	1.616^{***}	1.648^{***}
	(0.264)		(0.244)			(0.229)	(0.225)	(0.218)
log spending per capita		3.683^{***}	4.724^{***}			3.737^{***}	3.887^{***}	5.009^{***}
		(0.595)	(0.576)			(0.641)	(0.618)	(0.960)
distance to border				-1.871^{***}		-1.329^{***}		
				(0.319)		(0.256)		
duration to border					-2.746^{***}		-1.897^{***}	-1.861^{***}
					(0.429)		(0.320)	(0.313)
log pop. × log spending p.c.								2.755^{**}
								(1.139)
const	-8.362^{***}	-34.926^{***}	-67.494^{***}	17.997^{***}	21.324^{***}	-48.573^{***}	-47.713^{***}	-61.803^{***}
	(2.820)	(7.271)	(7.333)	(1.527)	(1.931)	(7.888)	(7.731)	(11.796)
Observations	2741	2741	2741	2741	2741	2741	2741	2741
R^2	0.226	0.060	0.323	0.125	0.142	0.382	0.386	0.412
F stat.	37.7	38.3	64.0	34.3	41.0	47.1	48.5	37.3

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Foreign spending shares are measured in percentage points. For the log density × log income per capita interaction, the two individual variables were demeaned beforehand. Weighted least squares with weights that are proportional to the number of bank customers per cell used in the analysis. Robust standard errors.

Figure A.I: Estimated gravity coefficients by industry

(a) Domestic consumer spending



(b) Labor compensation

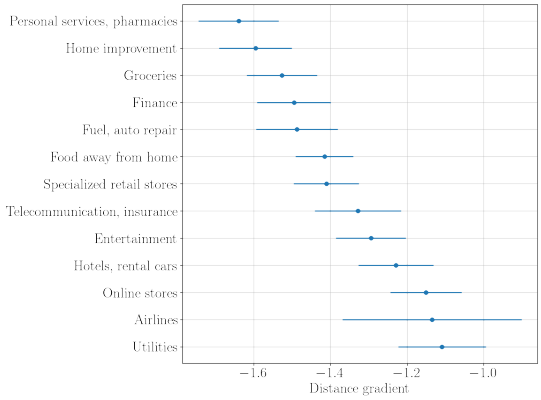
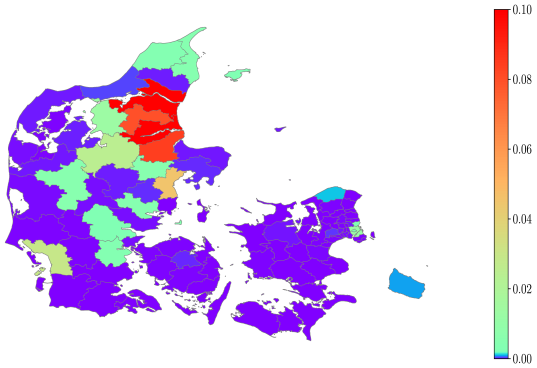
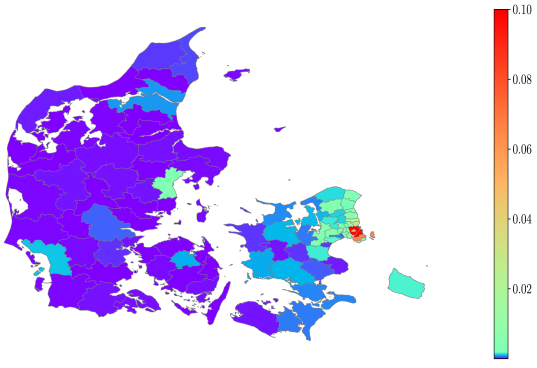


Figure A.II: Labor compensation shares across regions: Examples

(a) Restaurant workers in Mariagerfjord



(b) Restaurant workers in Copenhagen



Note: Scale truncated at 0.10. These maps show the percentage of labor compensation earned in the various regions of Denmark.

Figure A.III: Correlation between labor compensation shares and domestic consumer spending shares

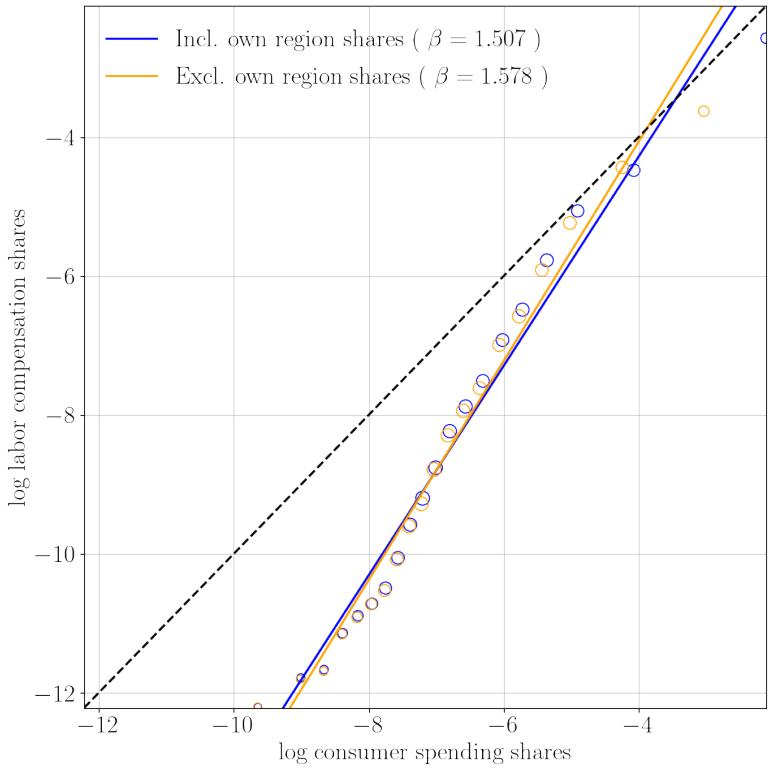


Figure A.IV: Population density of home and receiving region: all industries

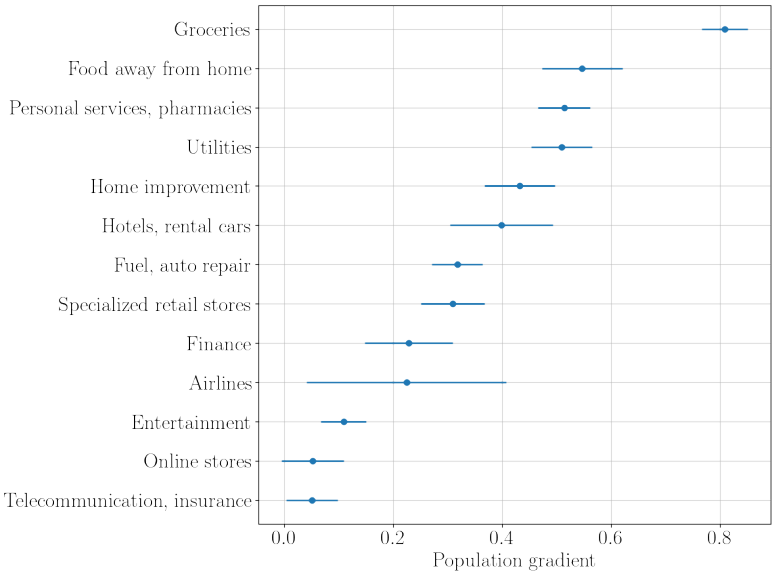


Figure A.V: Share of total sales going into cities, across spending categories

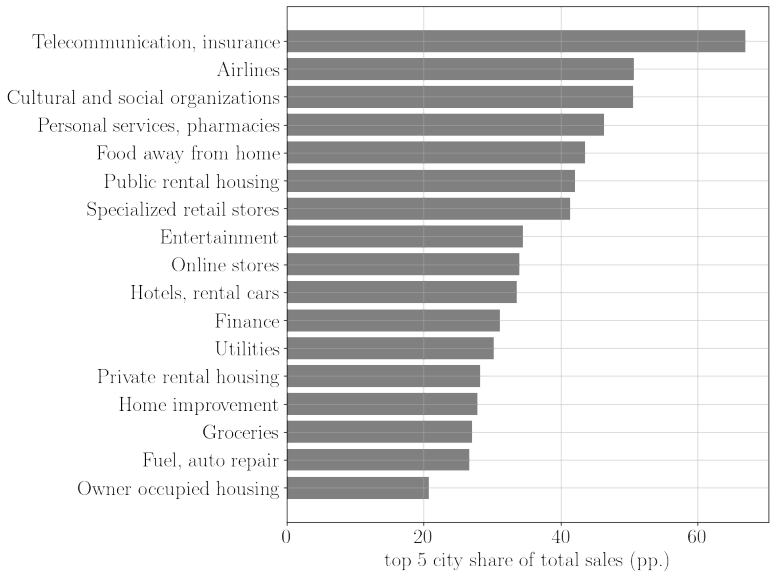


Figure A.VI: Relative consumer sales shares, top 5 cities vs rest

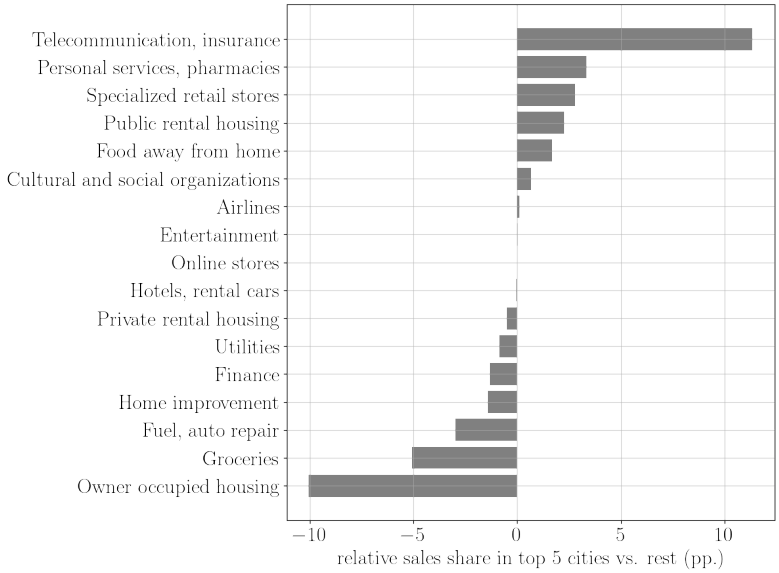
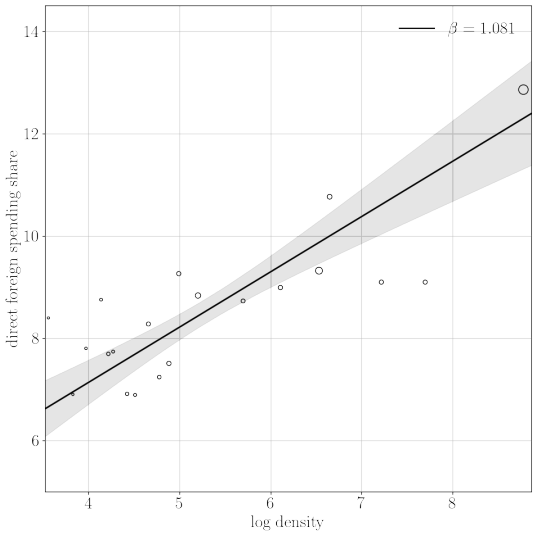
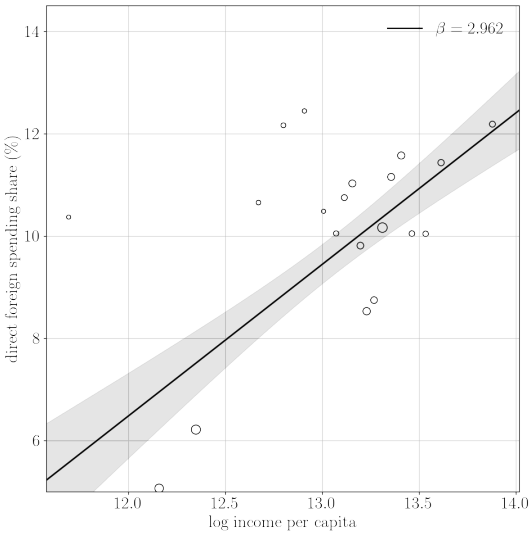


Figure A.VII: Foreign spending by population density and total income

(a) By population density



(b) By total income per capita



Notes: The figures show the direct foreign spending share of consumer cells by their home region's log population density. Both plots are binned scatter plots with 98 bins mirroring the number of regions in our analysis. Weights are proportional to the number of bank customers in a municipality.

Figure A.VIII: Log population slopes by foreign spending category

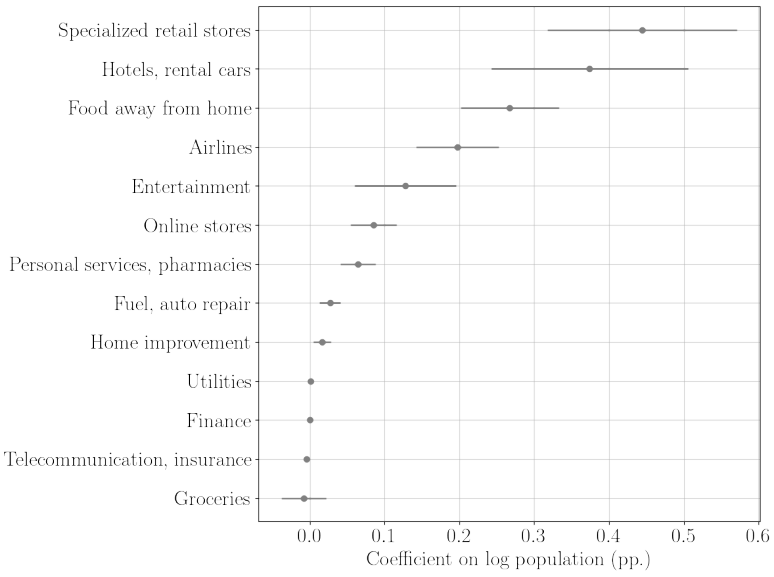
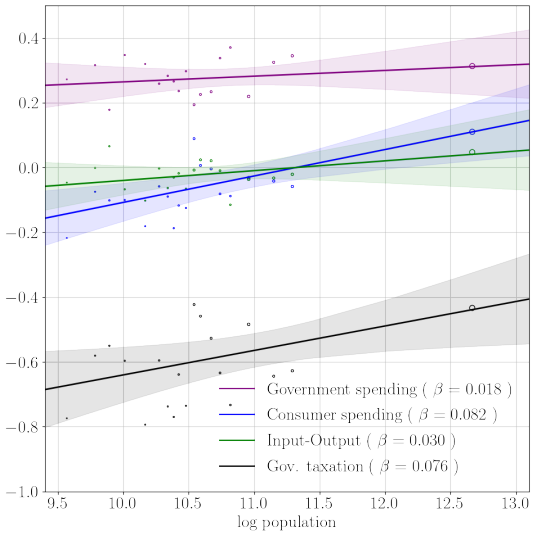


Figure A.IX: Flows towards rural and urban regions

(a) Flows towards urban regions



(b) Flows towards rural regions

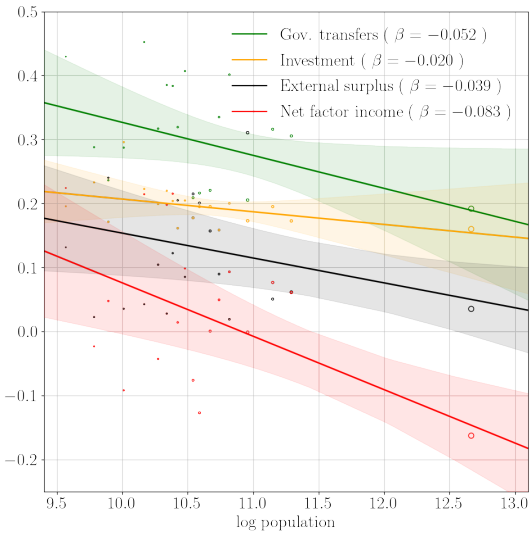


Figure A.X: Share of total IO sales going into cities, across spending categories

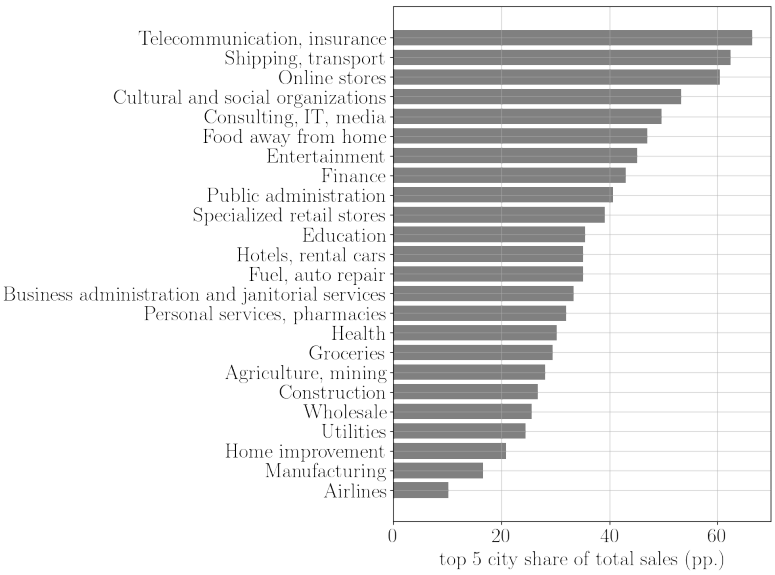


Figure A.XI: Relative IO sales shares, top 5 cities vs rest

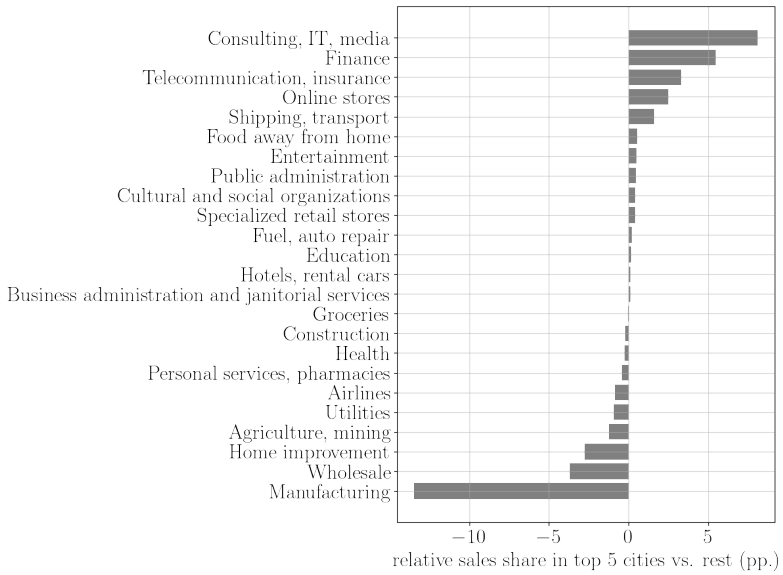


Figure A.XII: Share of industry sales accounted for by government spending

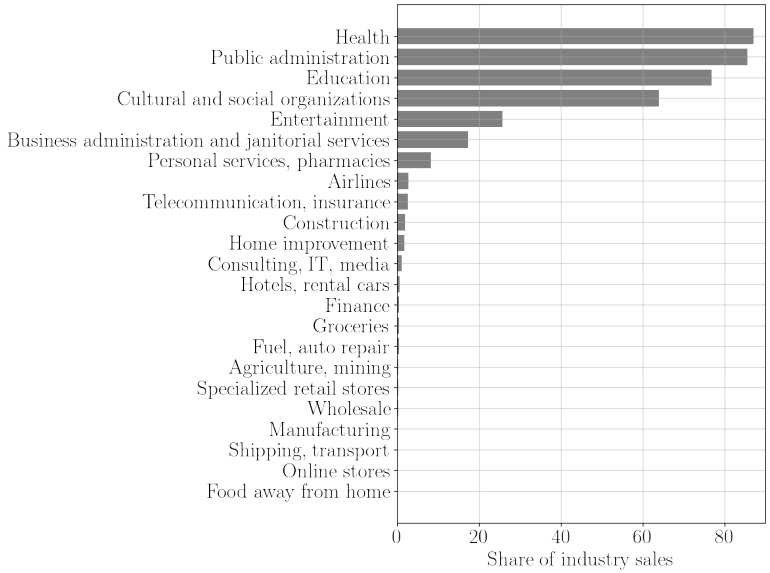


Figure A.XIII: Distribution of government spending across industries

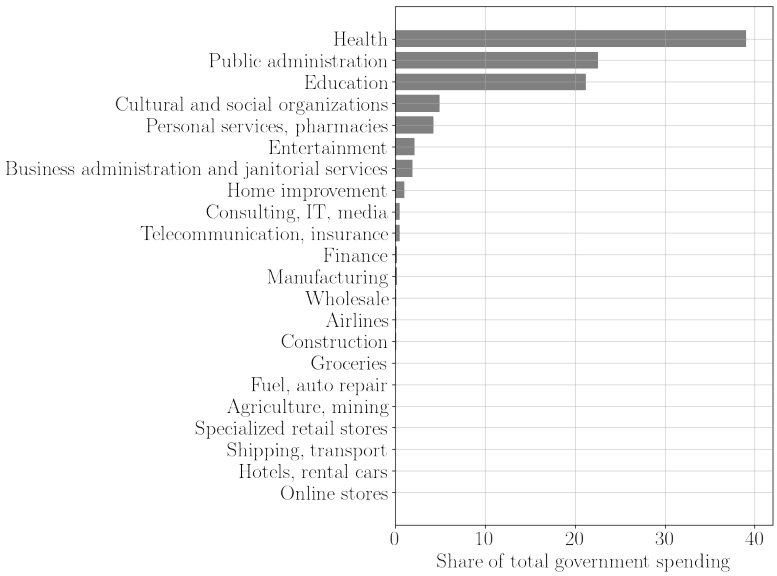


Figure A.XIV: Shares of government surplus and dividends by industry

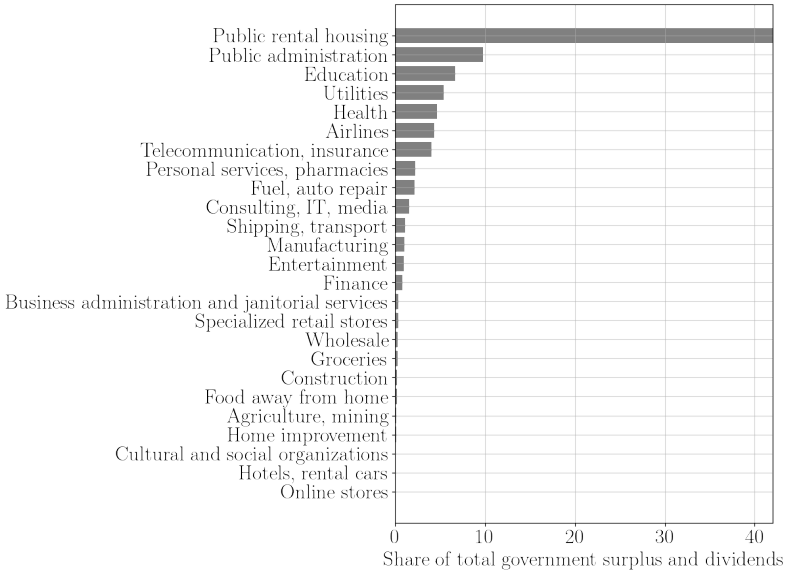


Figure A.XV: Shares of total government benefits received by industry

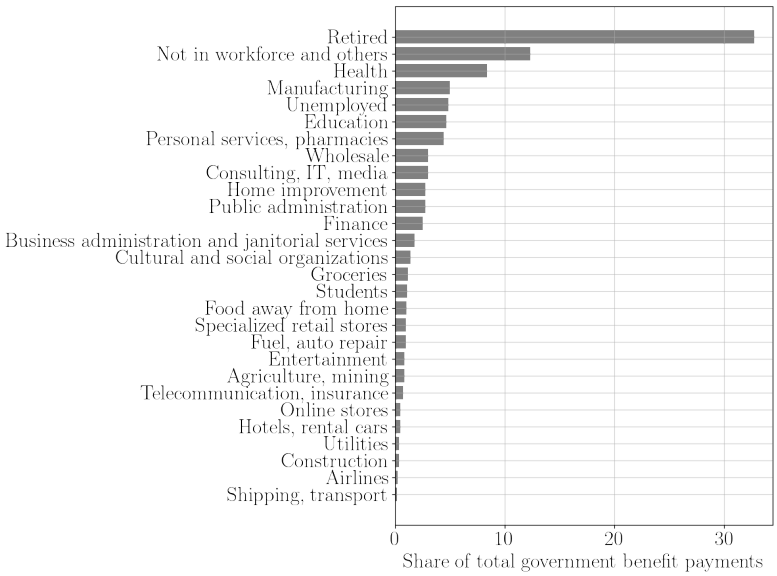
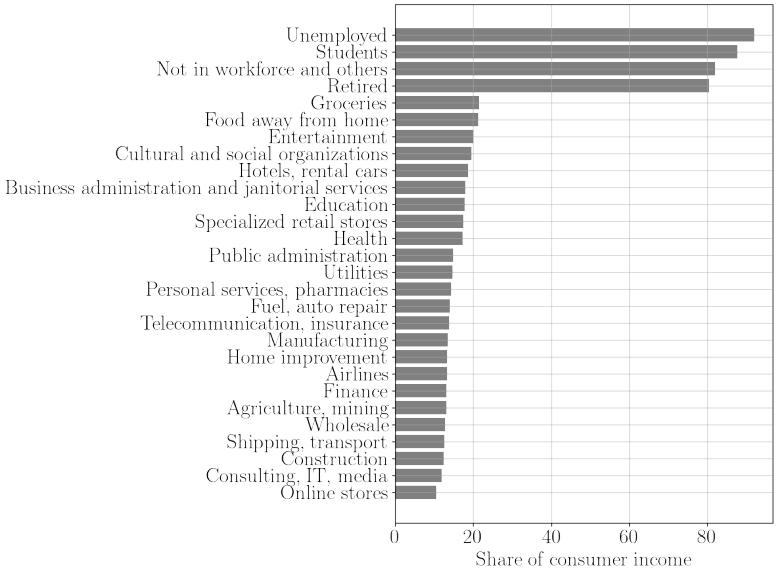


Figure A.XVI: Government benefits received as shares of total consumer income



Appendix B Aggregate Accounts by Sector (T-Tables) Based on Danish Disaggregated Economic Accounts

Table A.IX: Aggregate consumer account

Outflows			Inflows		
Outflow	Outflow to	Value (bn DKK)	Inflow	Inflow from	Value (bn DKK)
Domestic consumer spending	Producers	771.9	Labor compensation	Producers	1141.81
Foreign consumer spending	Rest of the World	81.94	Mixed income from non-corporate producers	Producers	80.72
Consumer product taxes paid	Government	173.18	Surplus of corporate producers to consumers (dividends)	Producers	38.48
Consumer non-product taxes paid	Government	566.35	Surplus of owner-occupied housing to consumers	Producers	83.27
Consumer social contributions paid	Government	181.06	Consumer social benefits received	Government	422.2
Consumer interest paid	Capital accumulation	29.67	Consumer adjustment for pension entitlements received	Government	92.48
Consumer natural resource rents paid	Capital accumulation	3.44	Consumer interest received	Capital accumulation	5.32
Consumer other transfers paid	Capital accumulation	44.82	Consumer pension investment income	Capital accumulation	75.48
Consumer gross saving	Capital accumulation	130.05	Consumer natural resource rents received	Capital accumulation	3.44
			Consumer other transfers received	Capital accumulation	39.21
Total value outflows		1982.42	**Total value inflows**		1982.42

Table A.X: Aggregate producer account

Outflows			**Inflows**		
Outflow	Outflow to	Value (bn DKK)	Inflow	Inflow from	Value (bn DKK)
Labor compensation	Consumers	1141.81	Domestic consumer spending	Consumers	771.9
Mixed income from non-corporate producers	Consumers	80.72	Domestic intermediates	Producers	1414.53
Surplus of corporate producers to consumers (dividends)	Consumers	38.48	Domestic government spending	Government	572.26
Surplus of owner-occupied housing to consumers	Consumers	83.27	Domestic capital accumulation spending	Capital accumulation	359.51
Domestic intermediates	Producers	1414.53	Producer exports	Rest of the World	1077.94
Dividends and surplus of government-owned/operated producers to government	Government	67.86			
Producer product taxes paid	Government	71.94			
Producer net non-product taxes paid and subsidies received	Government	82.79			
Producer net interest, transfers, and saving	Capital accumulation	422.41			
Producer imports	Rest of the World	792.33			
Total value outflows		4196.14	**Total value inflows**		4196.14

Table A.XI: Aggregate government account

Outflows			Inflows		
Outflow	Outflow to	Value (bn DKK)	Inflow	Inflow from	Value (bn DKK)
Consumer social benefits received	Consumers	422.2	Consumer product taxes paid	Consumers	173.18
Consumer adjustment for pension entitlements received	Consumers	92.48	Consumer non-product taxes paid	Consumers	566.35
Domestic government spending	Producers	572.26	Consumer social contributions paid	Consumers	181.06
Government imports	Rest of the World	4.26	Dividends and surplus of government-owned/operated producers to government	Producers	67.86
Government net interest, transfers, and saving	Capital accumulation	51.99	Producer product taxes paid	Producers	71.94
			Producer net non-product taxes paid and subsidies received	Producers	82.79
Total value outflows		1143.19	**Total value inflows**		1143.19

Table A.XII: Aggregate rest of the world account

Outflows			Inflows		
Outflow	Outflow to	Value (bn DKK)	Inflow	Inflow from	Value (bn DKK)
Producer exports	Producers	1077.94	Trade surplus	Capital accumulation	100.55
			Foreign consumer spending	Consumers	81.94
			Producer imports	Producers	792.33
			Government imports	Government	4.26
			Capital accumulation imports	Capital accumulation	98.86
Total value outflows		1077.94	**Total value inflows**		1077.94

Table A.XIII: Aggregate capital accumulation account

Outflows			**Inflows**		
Outflow	Outflow to	Value (bn DKK)	Inflow	Inflow from	Value (bn DKK)
Consumer interest received	Consumers	5.32	Consumer interest paid	Consumers	29.67
Consumer pension investment income	Consumers	75.48	Consumer natural resource rents paid	Consumers	3.44
Consumer natural resource rents received	Consumers	3.44	Consumer other transfers paid	Consumers	44.82
Consumer other transfers received	Consumers	39.21	Consumer gross saving	Consumers	130.05
Domestic capital accumulation spending	Producers	359.51	Producer net interest, transfers, and saving	Producers	422.41
Capital accumulation imports	Rest of the World	98.86	Government net interest, transfers, and saving	Government	51.99
Trade surplus	Rest of the World	100.55			
Total value outflows		682.38	**Total value inflows**		682.38

Appendix C Overview of Data

We rely on several types of data to construct the disaggregated economic accounts. First, we use customer data from Danske Bank with information about individual customers and about their transactions such as payments and transfers. Second, we use administrative data from a range of government registers such as the population register and the income register with information about individuals and the employment register with information about employment relations. Third, we use publicly available tabulations of aggregate statistics, such as housing statistics, health statistics and securities statistics.

We use these different types of data to disaggregate the flows between consumers, producers and the government described in the national accounts to the level of region-by-industry cells. In the following appendices, we provide more details about the data sources and the various steps of the disaggregated measurement. Table A.XIV lists the disaggregated datatsets, data sources, and the Appendix where laying out the details of the disaggregated measurement.

We strive to obtain consistency across all the disaggregated measurement by relying, whenever possible, on the same sample of individuals, the entire adult population in 2018, and a uniform assignment of individuals and firm establishments to cells based on government registers. The only instance where we need to deviate from these general rules is when we construct of the spending matrix. Here, we work with the sample of Danske Bank customers, roughly 20% of the national population, as we do not have data on spending for customers of other banks. Moreover, we need to assign individuals and firm establishments to cells using the bank's internal data, as confidentiality concerns prevent us from merging the bank data and government registers. Finally, we use data for both 2018 and 2019 to maximize sample size and then scale to match 2018 national accounts.

Table A.XIV: Overview of Data Sources

Disaggregated dataset	Microdata from Danske Bank	Microdata from government registers	Aggregate statistics	Details
Region-by-industry cells	Customer records, transaction data on incoming transfers	Population register (BEF), labor market register (AKM), income register (IND), employment registers (BFL, IDAN)		Appendix D
Consumer spending matrix	Customer records, transaction data on card payments, cash withdrawals, bill payments and mobile applications	Income register (IND), credit register (URTE), auto register (DMR)	National Housing Statistics (BOL101)	Appendix E
Consumer taxes (product)			Input-Output table (NIO3)	Appendix F
Consumer taxes (non-product)		Income register (IND), pension contribution register (INPI), population register (BEF)		Appendix G
Consumer interest, transfers, and saving (paid)		Population register (BEF), income register (IND)		Appendix H
Labor compensation matrix		Income register (IND), pension contribution register (INPI), employment registers (BFL)		Appendix I
Consumer dividend, mixed income, and surplus matrix		Income register (IND), employment register (IDAN)	General Firm Statistics (GF5), Registered Securities Statistics (DNVPDKS)	Appendix J
Government benefits to consumers		Income register (IND), pension contribution register (INPI), population register (BEF)		Appendix K
Consumer interest, transfers, and saving (received)		Income register (IND)		Appendix L
Intermediate trade matrix			Input-Output tables (NIO1, NIO2, NIO3), Cargo transport statistics (Tables NVG23, SKIB481)	Appendix M
Government dividend and surplus income		CVR register (hand-collected)		Appendix N
Consumption of government and NPISH output	Transaction data on payments	Population register (BEF), education register (UDD), income register (IND)	Public Expenditure Statistics (SYGU1), Health Statistics (INDAMP01), Education Statistics (UDDAKT20), Child Care Statistics (BOERN4), Culture and Leisure Statistics (BIB1)	Appendix O

Appendix D Defining Region-by-Industry Cells

The unit of our measurement is region-by-industry cells, the interaction of geographical regions and economic industries. Specifically, we define 99 regions (98 Danish municipalities and one foreign region) and 31 industries (Table A.I). We choose this definition of region and industries because we can observe it consistently across all underlying datasets and because it reveals a large degree of heterogeneity in disaggregated flows.

There are 24 industries that employ consumers. They include retail industries selling directly to consumers (e.g., food away from home, entertainment, grocery stores, drug stores), non-retail industries transacting mostly with firms (e.g., wholesale, agriculture, manufacturing, business services), and government-operated industries (e.g., public administration, health, education). We map six-digit NACE industry codes to the 24 industries with employment (the European NACE is based on the standard United Nations ISIC). Additionally, there are 4 separate industries for the non-working parts of the population (e.g., retired, students, unemployed, out of workforce) and 3 industries providing housing without employees.

In the government registers, we assign all adults to a region based on their home address at the start of 2018 observed in the administrative population register (BEF) and to an industry based on the NACE code of the firm establishment responsible for the largest part of their 2018 labor income in the labor market register (AKM). Individuals with no labor income are assigned to an industry based on their age, observed in the administrative population register (BEF), and other income sources, observed in the income register (IND). Specifically, individuals with no labor income are assigned to the industry "retired" if they are older than 65 years; to the industry "students" if they receive a government stipend (for which higher education students are almost universally eligible); to the industry "unemployed" if they receive unemployment or cash benefits; and otherwise to the industry "out of workforce." We assign firm establishments to producer region-by-industry cells using information from the employment registers (BFL and IDAN).

In the Danske Bank data, we identify consumers' industry by extracting the name of the employer from incoming salary payments and by identifying incoming government transfers (retired, students, government stipend, unemployed). We use the banks' customer records to identify their region of residence. The address register in the bank is linked to the government's address registers and updated on a monthly basis. To ensure that moves across regions do not distort the spending patterns, we update an individual's region every month when constructing the spending matrix. However, consistent with the assignment in the government registers, we define the industry of main employment on an annual basis, as the industry paying the largest share of consumer income.

Appendix E Measuring the Disaggregated Consumer Spending Matrix

Appendix E.A Data and Sample

To construct the spending matrix, we rely on data from Danske Bank for 2018 and 2019. We observe transaction-level information on consumers and merchants only for this period. Our sample consists of adult customers who conducted at least one transaction per month and registered their main bank account at Danske Bank throughout 2018 and 2019.[A1] For each customer, we observe all incoming and outgoing transactions. Card payments accounted for 47% of the total value of payment transactions. The most common card was Dakotan (debit cards issued by Nets A/S), followed by MasterCard (debit and credit cards) and Visa (debit and credit cards). Cash withdrawals accounted for 7% of transaction value, while bill payments (direct debits and bank transfers to merchants) accounted for 45%. Mobile applications, such as Apple Pay or the Denmark-specific MobilePay, made up 1%.

Appendix E.B Identifying the Merchant Region

We extract the address of the merchant establishment (i.e., the store) from the string that accompanies payment transactions in the bank's internal computer system. The information for card and mobile payments differs slightly by payment type.

- Dakotan statements include a unique ID number for each merchant establishment for transactions in Denmark and the country name for transactions abroad. We match the Danish merchant IDs to the exact merchant address using a table issued by Nets A/S.
- MasterCard includes a detailed merchant address directly in the transaction string in the following format: merchant ID, shop name, street name, house number, postal code, country.
- Visa reports the merchant ID, shop name, and town for each transaction. The merchant IDs used by Visa and MasterCard generally coincide. Based on the MasterCard data, we can therefore construct a table matching merchant ID and detailed address. In very few cases, a merchant ID gets used twice for a Danish merchant and a foreign merchant. In these cases, we assume that the transaction was with the Danish merchant.
- Some transactions in MobilePay and Apple Pay contain merchant addresses, but many do not. There is no clear pattern in which transactions contain addresses, which is likely due to the recent development of these transactions.

[A1] All adults register one "main" account with the Danish government, through which they conduct all financial interactions with the government.

For all card and mobile payments, we extract the merchant address and convert it to a consistent format using an API service provided by the government agency Styrelsen for Dataforsyning og Effektivisering (`http://dawa.aws.dk/dok/api/adgangsadresse#datavask`). This conversion identifies precise geocodes for each merchant while accounting for misspelled addresses and addresses that appear twice due to minor differences in formatting or spelling. The API compares the merchant address from a transaction with its database of official addresses. It iterates in a Levenshtein manner (i.e., it calculates the number of letters/digits that must be exchanged before one string is equal to another). We force the Levenshtein process to consider only addresses with exactly identical postcodes. Municipalities are combinations of several postcodes. By restricting to the same postcode, we ensure that the Levenshtein process cannot change the municipality information, the key information that we use to construct the spending matrix.

If the API cannot match the address unambiguously (so called C-match), we remove the first line of the address, which often combines abbreviations of merchant and street name, making it difficult to automatically recognize. We also check whether the address contains the name of a shopping mall, rather than an official address. If so, we replace the name of the mall with the mall's address and rerun the API process. Finally, we manually research the official address of the 100 most common unmatched addresses.

Using this procedure, we identify the official shop address for 95% of card and mobile spending. We assume that the remaining 5% go proportionally to the same regions as other card and mobile spending. These remaining 5% also include cases where mobile applications (e.g., MobilePay) and online services (e.g., PayPal) do not directly send the purchase amount to a sales terminal, but transfer to a central account first before paying the merchant.

For bill payments, we use a slightly different approach. We directly observe the merchant's postal code for recurrent bill payments, which make up 67% of all payments. These observed postal codes allow us to infer the merchant region for the majority of remaining bill payments. Specifically, we split merchants into 48 industries. We calculate the number of transactions from each consumer region going to each of these 48 industries. To minimize noise, we keep industries where at least 50% of incoming bill payments contain postal code information and where we observe at least 200 incoming transactions from every consumer region. (Industries receiving 80% of total bill transaction value satisfy these two requirements.) For these industries, we then assume that bill payments flow to the same postal codes as bill payments with observed postal codes from the same consumer region to the same industry. For the remaining industries (covering 20% of bill transaction value), we assume that bill payments flow to the same postal codes as card payments from the same consumer region to the same industry.

We generally do not observe how cash withdrawals are spent, but we assign them to merchants in proportion to observed card spending in Denmark if the withdrawal was in Denmark and to

observed card spending abroad otherwise. In a few cases, country information is missing, so we assume that withdrawals without decimal points (e.g., 100.00 DKK) are withdrawn and spent in Denmark and that other withdrawals (e.g., 100.76 DKK) are spent abroad.

Appendix E.C Identifying the Merchant Industry

We extract the merchant category code (MCC), a classification for the type of goods sold, for each merchant in the bank transaction data. To create a disaggregated system with consistently defined producer cells, we need to map MCCs to the industry codes used in the employment and production data. However, no such mapping exists so far. We therefore create a new cross-walk between MCCs and NACE industries.

First, we observe each merchant's Danish business identification number (CVR) and MCC in the bank's system. We link the CVR to the Danish Central Business Register where we can retrieve the merchant's industry (at the level of 741 NACE codes). Second, we manually assign MCCs to our 31 industries (only 14 of which are consumer-facing and are assigned MCCs). We then identify which of our industries appear most frequently among merchants in each of the 741 NACE codes. In very few cases, two industries appear equally often and we manually research the largest firms to identify the right match.

We create an alternative mapping between a merchant classification system called PCAT and our industries. The PCAT usually appears as part of the electronic transaction information for bill payments and can easily be mapped to our industries. MCC and PCAT are missing for some bill payments, amounting to 8% of total transaction value through bills. We assume that these payments go proportionally to the same industries as other bill payments by the same consumer cell.

Appendix E.D Online Spending

We identify whether card transactions took place in a physical store or online. Dakotan transactions include a straightforward indicator for online transactions. Mastercard and Visacard transactions contain ISO 8583 information, an internationally standardized message sent by a sales terminal in a transaction. If the POS7 code (input method) equals 1, 6, K, or L or if the POS5 (cardholder) code equals 5, the transaction is online. We treat payments using mobile applications (e.g., MobilePay) and online services (e.g., PayPal) that do not report a physical merchant ID and address as online transactions.

In constructing the spending matrix, we generally treat online spending identically to spending in a physical store. That means that we identify the region-industry cell of the merchant and assign the incoming spending to this cell. However, we adjust the regional (but not the sectoral)

distribution of online spending for industries where we know that consumption of the final good takes place entirely in a physical location. In these cases, we assign all online spending using the regional distribution of spending on physical merchants. The online spending on these industries often goes to the central payment terminal of a parent company before being assigned to the physical merchant. For instance, online purchases of cinema tickets are often booked through a central company terminal in Copenhagen, even though consumption happens in local movie theaters. The full list of industries where we adjust the regional distribution is: food away from home; entertainment; medical and other specialized merchants; commuting; vehicle repair; hotels; rental cars; home improvement services. (These industries are generally sub-categories of our 31 final industries.)

Appendix E.E Augment the Spending Matrix with Government Data

Consumer spending on four types of goods are not captured well in bank transaction data: housing, financial services, vehicles, and water and waste services. We replace the transaction-based values for these four goods in the spending matrix with adjusted values derived from combining Danske Bank data with government registers.

Appendix E.E.1 Consumption of housing

We use different methods to assign consumption of owner-occupied housing and rented housing. First, owner-occupied rents are notoriously difficult to measure because they involve no financial transaction. However, the administrative income register (IND) contains the imputed rental value of owner-occupied housing for every individual. We thus allocate the national consumption of owner-occupied housing to consumer cells in proportion to the imputed rental value of their owner-occupied housing. Expenditure on owner-occupied housing flows to the producer cell for owner-occupied housing (our industry number 30) located in the same region as the home owners. The operating surplus of an owner-occupied housing producer cell in a region (SNA B.2G) then goes back to the consumer cells owning homes in that same region.

Second, we do not observe all rental payments in the bank transaction data. We instead distribute the national consumption of rented housing across consumer cells in proportion to their estimated rental costs. We observe some rental payments in the bank data, which we use to estimate:

$$rental\ payment_p = \alpha + \vartheta\, region_p + \phi\, industry_p + \psi_1\, age_p + \psi_2\, age_p^2 + \varepsilon_p, \tag{A.1}$$

where p is an individual renter. Using the estimated fixed effects ϑ and ϕ, we predict the average rental cost in each consumer cell. We observe ownership of real estate in the administrative

income register (IND) and assume that all consumers who do not own real estate are renters, which allows use to calculate the number of renters by consumer cell. Combining estimated rental payments with the number of renters allows us to estimate total rental costs by consumer cell, which we use to allocate national consumption of rented housing across cells.

Finally, we use aggregate statistics to assign rental payments to producer cells. In the National Housing Statistics (Table BOL101), we observe the number of rental housing units in each region owned by different owner types: individuals, non-profit building societies, limited liability companies, housing societies, and public authorities. The surplus of individual owners accrues to the "private landlord" industry (our industry number 29) and those of corporate owners accrue to the "finance, real estate" industry (our industry number 17). As there is no information about the geographical location of the owners, we assume that the geographical distribution of individual owners of rental units in a given region follows the geographical distribution of mixed income in the region and that the geographical distribution of the individuals behind corporate owners of rental units in a given region follows the geographical distribution of dividend payments. The remaining owner types are public or non-profit organization. The surplus of these owners goes to the "government-owned housing" industry (our industry number 31) in the same region where the housing is located.

Appendix E.E.2 Consumption of financial services

The consumption of financial services in the national accounts is composed of the value of financial advice provided by financial firms and the interest rate spreads applied by financial firms. While we observe payments for financial advice in the bank transaction data, it is difficult to disentangle the interest rate spreads from the raw value of interest payments in transaction data. Instead, we allocate the national consumption of financial services across consumer cells in proportion to their interest expenses. The tax register (IND) contains interest expenses for every individual. We aggregate interest expenses at the level of each consumer cell.

Producer cells in our "finance, real estate" industry (number 17) receive consumers' expenditures on financial services. To identify which regional producer cells receive expenditures from which consumer cells, we use loan-level data from the administrative credit register (URTE). Specifically, for each bank loan and each interest payment, the credit register contains an identifier for the bank branch that has recorded the loan and the interest payment. There are around 3,000 bank branches in Denmark. We do not observe the location of branches directly, so we define the region of each branch as the most common region of consumers holding loans recorded at the branch. We then compute how the interest payments of each consumer cell are distributed across bank branches in different regions. Finally, we assume that a consumer cells' expenditure of financial services is distributed across producer cells in proportion to its distribution of interest

payments.

Appendix E.E.3 Vehicle purchases

We do not observe all vehicle purchases in the bank transaction data because many purchases do not flow directly to the vehicle producer but through financial firms. We therefore use a top-down approach to assign national vehicle purchases to consumer cells in proportion to each cell's share of total spending on new cars. We estimate each consumer cell's total spending on new cars by combining bank transaction data on annual spending at vehicle dealers with information on vehicle registrations from the administrative auto register (DMR). We use data over the period 2014-2016, as this is the most recent period where we can combine transaction data from Danske Bank and administrative data from the auto register. We therefore assume that relative levels of vehicle spending are unchanged between 2014-2016 and 2018.

We first estimate the average vehicle price in each consumer cell. We use a sample of individuals where we observe just one new car registration in a given year and spending at vehicle dealers of at least 50,000 DKK in the bank transaction data in the same year. We then regress individual-level spending at vehicle merchants on industry-by-year and region-by-year fixed effects:

$$spending_{p,y}^{vehicle_merchants} = \theta_{i(p),y} + \eta_{r(p),y} + \epsilon_p.$$

where p is an individual in industry $i(p)$ and region $r(p)$ and year is y. We predict the average price of new cars in each consumer cell using the estimated fixed effects. We can directly calculate the number of newly registered vehicles in the government vehicle registers. Combining the estimated price with the number gives an estimate of total spending on new cars by consumer cell for each year in 2014-2016. We compute each cell's share of total spending on new cars in each year and then average across years. We use these shares to allocate national vehicle purchases in 2018 across consumer cells.

Producer cells in our "cars, fuel, car repair, public transport" industry (number 5) receive consumers' expenditures on vehicles. To identify which regional producer cells receive spending from which consumer cells, we use Danske Bank data on card transactions. Since vehicle purchases are relatively infrequent, we collapse to the regional level and assume that all consumers within the same region distribute their vehicle spending across regions in the same proportion.[A2]

[A2] If there are less than 50 transactions for a region, we group that region with a neighboring region. This leads us to group Læsø with Frederikshavn; Langeland and Ærø with Svendborg; Fanø with Esbjerg; Ringkøbing-Skjern with Herning; Lemvig with Holstebro; Morsø with Thisted; and Samsø with Odder.

Appendix E.E.4 Water and waste services

Rental payments often include consumption of water and waste services in Denmark, which implies that we cannot separately identify spending on water and waste in the transaction data. We therefore allocate the national consumption of water and waste services to consumer cells in proportion to their spending on other utilities. We assume that water and waste is produced locally, setting the region of the producer cell receiving the payments equal to region of the consumer cell.

Appendix F Measuring Consumer and Producer Product Taxes Paid

Product taxes are paid by buyers upon the purchase of a good. The most important product taxes are import taxes, product-specific taxes (e.g., on fuel, energy, cigarette, and alcohol), and valued added taxes (VAT).

We first focus on measuring product taxes paid by consumers. We observe import and product-specific taxes paid by consumers in the Danish national accounts table NIO3. It reports the total amount of import and product-specific taxes paid by consumers on purchases of goods produced by 117 distinct industries. These 117 industries do not map directly into our 31 industries. We therefore break down the 117 industries into the most granular grouping in the Danish national accounts (741 industries), assume that taxes paid are proportional to employment shares, and aggregate to our 31 industries. We then calculate the implied tax rate by combining information on total consumer spending on each industry from the consumer spending matrix with the tax data. The implied import and product-specific tax rates range from 0 percent (e.g., for "personal services, pharmacies" and "cultural and social organizations") to 50 percent in utilities (due to Denmark's very high energy tax rates).

The standard VAT rate in Denmark is 25 percent. We set this tax rate for consumer spending on all industries except a few industries whose products are VAT-exempt: airlines, finance, health, education, public administration, and all housing. In addition, spending on two of our industries is partially exempt: insurance (part of our industry 8, "telecommunication and insurance") and culture (part of our industry 18 "cultural and social organizations"). We use data from NIO3 on total VAT paid by consumers by industry and calculate that the average VAT rate is 7 percent (for industry 8) and 5 percent (for industry 18).

We observe product taxes paid by Danish producers, the government and capital accumulation cells, and foreign countries in the Danish Input-Output table, described in detail in Appendix M. Notable facts are that VAT does not get paid on exports and that firms get reimbursed for any VAT they paid on intermediates.

Appendix G Measuring Disaggregated Consumer Non-Product Taxes

Consumer taxes comprise national accounting flows from the consumer sector. First, current taxes on income, wealth, etc., payable (SNA D.5) includes income taxes paid directly by consumers as well as a tax on pension wealth returns paid by pension funds on their behalf. We disaggregate each part separately. In the income register (IND), we observe total annual income tax liabilities, excluding pension returns taxes, for each individual in the population. We use this as the disaggregated measure of directly paid income taxes, scaling it by a factor **1.09** to match the national accounts value when aggregating across consumer cells. For the pension returns taxes paid by pension funds, we have no direct micro level measure. We therefore apply a top-down approach assuming that DKK pension wealth returns, and hence also returns taxes, are proportional to total accumulated pension contributions since 1995, which is the first year for which we have microdata on pension contributions in the pension contribution register (INPI).

Second, we use a bottom-up approach to disaggregate social contributions, payable (SNA D.61). In the income register (IND) and the pension contributions register (INPI), we extract information on total annual pension contributions, including contributions to a mandatory retirement savings program (ATP), as well as on membership fee payments to unemployment insurance funds. We aggregate these variables to the cell level and scale the cell totals by a factor **1.2** to make the national total match the value from national accounts.

Appendix H Measuring Disaggregated Consumer Interest and Transfers Paid

Consumer financial and rental expenditures include consumers' interest payments (SNA D.41) and payments related to renting of land and subsoil resources (SNA D.45). We disaggregate the former using a bottom-up approach. From the administrative income register (IND), we have information on each individual's interest payments on all financial liabilities. The sum of these interest payments exceeds the value of position D.41 in the national accounts because the micro-level measure includes the full nominal amounts paid by consumers to lenders, whereas the national accounts value is net of Financial Services Indirectly Measured (FISIM). We handle this by **scaling** the micro-level variable so that it matches the national accounts value in the aggregate, the implicit assumption being that the ratio between FISIM and total nominal interest payment is the same across consumer cells. Since we have no microdata information about payments related to renting of land and subsoil resources, we use a top-down approach to disaggregate this national accounts flow. Each consumer cell is assigned a share of the aggregate value corresponding to its population share derived from the administrative population register (BEF). Other current transfers, payable (SNA D.7) includes non-life insurance premium payments and miscellaneous current transfers. We disaggregate it top-down, assigning to each cell a share of the national total

equal to its population share derived from the administrative population register (BEF).

Appendix I Measuring the Disaggregated Labor Compensation Matrix

The labor compensation matrix records flows of labor income from producer cells to consumer cells. We use the income register (IND) and the pension contributions register (INPI) to measure total annual labor compensation, including employer-paid contributions to pension schemes, received by each individual. We define the producer cells based on the industry of the individual's main employer observed in the labor market register (AKM) and the region of the firm establishment where most of the payments come from observed in the employment register (BFL). We then aggregate these payments to a matrix at the level of consumer and producer cells. We scale all values by a factor of **1.02** to match the national accounts flow Compensation of employees, receivable (SNA D.1).

Appendix J Measuring the Disaggregated Dividend, Mixed Income, and Surplus Matrix

The matrix reveals dividends, mixed income, and surplus flowing from each producer cell to each consumer cell. We first calculate how each producer cell is split into a "corporate" and a "non-corporate" part and then, for each part separately, estimate the distribution of dividends, mixed income, and surplus across consumer cells.

In the first step, we use aggregate statistics to split firms in each industry into a corporate and a non-corporate part (Table GF5 in "General Firm Statistics"). We categorize sole proprietorships as "non-corporate," corporations with limited liability and partnerships as "corporate," and associations, funds and other rare legal forms as "other." For each industry, we are thus able to compute the share of total sales attributable to "non-corporate" and "corporate" firms, respectively. The sole exception is the financial industry where we compute the split based on employment, as sales are not reported. We convert the industries reported in Table GF5 to our industries using employment shares as weights.

In the next step, we estimate how corporate ownership is distributed across consumer cells. We measure foreign ownership using aggregate statistics (Table "DNVPDKS" in the "Registered Securities Statistics"). The share of Danish stocks held by foreigners was 50% in 2018. We assume that this fraction is constant across all producer cells. We then distribute the ownership of the remaining 50% across domestic consumer cells in proportion to their stock market wealth. To implement this distribution, we aggregate the market value of stocks from the administrative income register (IND) for each consumer cell.

Next, we determine how non-corporate ownership is distributed across consumer cells. Here, we construct a self-employment matrix similar to the labor supply matrix discussed above. To construct the matrix, we link information about establishments operated by sole proprietorships from the employment register (IDAN), including the municipality where the establishments are located, to the business income of the self-employed individual operating the establishment.

We distribute the income of stock corporations and privately-owned firms using distinct methods. For stock corporations, we rely on data from the income register (IND) about individual stock dividend income, which we use to disaggregate distributed income of corporations, receivable (SNA D.42). We **scale** so that the total across consumer cells matches the national aggregate.

Finally, the national accounting flow operating surplus, gross (SNA B.2G) corresponds to operating surplus from owner-occupied housing. We aggregate individual-level imputed rental values of owner-occupied housing as reported in the income register (IND), which produces the industry's total output. We scale this output by a **factor 0.67** to match the national value for B.2G. The implicit assumption is that the ratio between gross operating surplus and output in the owner-occupied dwellings industry is constant across consumer cells.

Appendix K Measuring Disaggregated Government Benefits to Consumers

National accounts describe three types of transfers to consumers. First, we aggregate all government income transfers and private pension savings payouts from the income register (IND) to calculate the cell-level measure of social benefits other than social transfers in kind, receivable (SNA D.62). We scale by a factor **1.03** to match the national accounts value. Second, other current transfers, receivable (SNA D.7) consist of miscellaneous current transfers, for example disaster and accident relief. We disaggregate this position top-down using cell population shares obtained from the administrative population register (BEF). Third, adjustment for the change in pension entitlements, receivable (SNA D.8) represents an accounting adjustment to make net changes in pension entitlements on which consumers have a definite claim appear in the calculation of consumer saving. We disaggregate it by combining data from the income register and the pension contributions register (INPI) to construct a micro-level measure of pension contributions net of payouts. We then scale this measure to match the national accounts aggregate value.

Appendix L Measuring Disaggregated Consumer Interest and Transfers Received

First, we disaggregate interest, receivable (SNA D.41) bottom-up by using individual-level information on interest income from the income register (IND) and scaling so that the total across consumer cells matches the national accounts aggregate. Second, other investment income, receivable (SNA D.44) includes investment income from insurance policies and pension entitlements.

We disaggregate this using a top-down approach where each consumer cell is assigned a share of the national accounts value proportional to its pension contributions accumulated since 1995. Third, rent, receivable (SNA D.45) consists of income from renting land and subsoil resources. We disaggregate it top-down using population shares.

Appendix M Measuring the Disaggregated Intermediate Trade Matrix

The disaggregated intermediates matrix describes how producers in one cell are connected to producers in other region-sector cells through their demand for production inputs. The Disaggregated Government Consumption Production Vector describes how producer cells contribute to the production of goods and services used for government consumption. To construct them, we start from a standard input-output table for the Danish economy at the most disaggregated level with 117 industries (Tables NIO1, NIO2, and NIO3 at statbank.dk). The input-output table illustrates how the output produced in one industry is used as intermediate inputs in other industries or in final use categories such as private consumption, government consumption, capital formation, or exports. It also shows how output from a given industry is produced from intermediate inputs from other domestic industries and from abroad (net of product taxes and import duties, which are added as additional primary inputs), and gross value added.

To convert the standard input-output table to a format suitable for our purposes, we need to address four challenges. First, the 117 industries do not map directly into our industry classification. Second, the input-output table has no geographical dimension: It describes inputs from firms in industry s to firms in industry t at the national level, but not from firms in industry s and region i to firms in region j and industry t. Third, the national accounting convention of defining production in the retail industry as the net trade margins makes the table unsuitable for combining with the disaggregated consumer spending matrix (which reports actual money flows) in order to follow the flow of money from consumers to retailers and further on to upstream producers. Fourth, the input-output table does not distinguish between residents' and non-residents' spending in the moneys flows from private consumers to domestic producers, bringing it at odds with the disaggregated consumer spending matrix. The following subsections describe how we handle each of these challenges.

Appendix M.A Adapting the IO table to DEA Industry Classification

To address the first challenge, we first disaggregate the input-output table based on the national accounts (NA) classification to a more granular subindustry grouping. We later re-aggregate to the industry classification used in our system of disaggregated economic accounts (DEA). For the initial disaggregation, we use that the industry codes in the microdata (which have 744

categories including the three zero-employment housing services industries) maps directly into both classifications.

Formally, let lower-case letters $\{a, b, c...\}$denote the NA industries in the standard input-output table and let upper-case letters $\{A, B, C...\}$denote the 27 output-producing DEA industries shown in Table A.I. Let j denote a granular industry at the level used in the microdata. Consider a particular NA industry $x \in \{a, b, c...\}$ and a particular DEA industry $Y \in \{A, B, C...\}$: We define subindustry x_Y as the set of granular-level industries that are subindustries of both x and Y, $x_Y = \{j | j \subset x, j \subset Y\}$. This produces 173 non-empty subindustries. These represent the highest level of industry aggregation compatible with both the NA classification and the DEA classification.

In the microdata, we aggregate salary payouts within each non-empty subindustryx_Y and compute the subindustry's salary share within NA industry x. We then disaggregate the flows reported in the original input-output table using these shares. For example, the original input-output table reports the value of flows from NA industry a to NA industry b. We assume that the flow stemming from subindustry a_Y is proportional to its share of total salary payouts to workers in NA industry a. Similarly, we assume that the flow to subindustry b_Z is proportional to its share of total salary payouts to workers in NA industry b. Concretely, we compute the flow from a_Y to b_Z as the total flow from a to b multiplied by the product of the salary shares of a_Y and b_Z within their respective NA industries.[A3]

The input-output table also reports money flows from final use categories (private consumption, government consumption exports, capital accumulaton, etc) to domestic producers, and from domestic producers to tax collectors and foreign producers. In both cases, we disagregate to the subindustry level on the domestic producer side. As a general rule, we again use salary shares for the disaggregation procedure, but with one important exception: The shipping industry (NA industry 500000) is large in Denmark, accounting for 16% of both imports and exports in 2018. It has two subindustries, one related to transport of passengers and one related to transport of goods, and the latter pays about twice as much in total salary payments as the former. Detailed foreign trade statistics show that trade values are about 100 times larger for goods than for passengers. On this background, we assign all imports and exports to/from the shipping industry to the goods transport subindustry.

It should be noted that the disaggregation procedure may introduce considerable measurement error if flows are not proportional to salary shares. But to the extent that granular industries that belong to the same industry in the NA classification also belong to the same industry in our DEA classification, which is very often the case, such errors will vanish when we later re-aggregate to

[A3]We make an exception to this general rule for the three zero-employment housing services industries where we use production shares rather than salary shares for the disaggregation.

the DEA industry level.

Appendix M.B Adding Geography

The next challenge is to disaggregate the national flows between subindustries and final use categories by pairs of regions. To do this, we use aggregate cargo transport statistics describing general trade patterns across localities, as well as micro-level data on the economic size of each region within a subindustry. Specifically, we assume that regions send and receive flows in proportion to their shares of total salary payouts in each subindustry, except in the three zero-employment housing services industries where we use regional spending shares from the disaggregated consumer spending matrix (see Appendix E). Further, we take the regions' geographical locations into account by imposing a gravity structure that assigns larger flows to pairs of regions that are close to each other than to those that are far apart.

The first step in this procedure is to obtain an estimate of the elasticity of intra-country trade with respect to distance. For that purpose, we use data from publicly available cargo transport statistics (Tables NVG23 and SKIB481) to estimate the following gravity model:

$$log(Goods_{ij}) = \alpha + \gamma log(Empl)_i + \eta log(Empl)_j + \beta log(Distance_{ij}) + \epsilon_{it} \tag{A.2}$$

where $Goods_{ij}$ denotes the weight of goods transported between provinces i and j by either road or sea; $Empl_i$ and $Empl_j$denote employment in provinces i and j respectively; and $Distance_{ij}$denotes the average driving distance between regions within province i and regions within province j (where all intra-province distances are set to 1km). We conduct the analysis at the province-level as the cargo data is not available at the more disaggregated region-level. To accommodate zeroes, we estimate the equation on its multiplicative form using a Poisson pseudo-maximum-likelihood estimator (Silva and Tenreyro, 2006). The model yields an estimated elastisticy of cargo transport with respect to distance of -0.61 (s.e. 0.03).

Equipped with this elasticity estimate, we disaggregate the national flows between subindustries to the regional level in the origin dimension as well as the destination dimension. Formally, let a_Y and b_Z denote subindustries and let i and j denote regions. We then assume that the flow from firms in subindustry a_Y, region i to firms in subindustry b_Z, region j is

$$Flow_{a_Y,b_Z,i,j} = Flow_{a_Y,b_Z} * \theta_{a_Y,b_Z,i} * \eta_{a_Y,b_Z,j} * Distance_{ij}^{-0.6} \tag{A.3}$$

where $Flow_{a_Y,b_Z}$is the flow from subindustry a_Y to subindustry b_Z at the national level and $Distance_{ij}$ is the distance between region i and region j.[A4] The parameters $\theta_{a_Y,b_Z,i}$ and $\eta_{a_Y,b_Z,j}$

[A4]Distances are measured as the distance between the two regions' centroids. We use Google Maps API service to calculate actual travel distances. We define a region's distance to itself as 1.

are origin- and destination region fixed effects within the a_Y-b_Z subindustry pair. These are set so that *i*) region i's total share of national a_Y-b_Z flows (i.e.$\sum_j Flow_{a_Y,b_Z,i,j}/Flow_{a_Y,b_Z}$) matches its share of total salary payouts in subindustry a_Z, and *ii*) region j's total share of industry a_Y-b_Z flows (i.e. $\sum_i Flow_{a_Y,b_Z,i,j}/Flow_{a_Y,b_Z}$) matches its share of total salary payouts in subindustry b_Z. In practice, we accomplish this through an iterative numerical procedure: Starting from initial guesses for $\theta_{a_Y,b_Z,i}$ and $\eta_{a_Y,b_Z,j}$, we compute the implied value of each $Flow_{a_Y,b_Z,i,j}$. We then adjust the guesses by a multiplicative constant that ensures that these implied values add up to their national level counterpart, $Flow_{a_Y,b_Z}$. Next, we update the guesses by multiplying them with the ratios of the regions' salary shares to the implied shares of national a_Y-b_Z flows. We repeat this procedure iteratively until the implied share of national a_Y-b_Z flows converges towards the relevant salary share for each origin and destination region.

For money flows from final use categories (private consumption, non-profits, government consumption, capital formation, exports) to domestic subindustries, we add a geographic dimension on the destination side only. Here, we simply assign each region a share of the total national flow equal to its share of total subindustry salary payouts. Conversely, for flows going from domestic subindustries to tax collectors and foreign producers, we add a geographic dimension on the origin side only: Once again, each region is assigned a share of the national flow equal to its salary share within the subindustry.

After disaggregating to the subindustry-region level, we aggregate to the industry-region cell level described in Appendix D by summing over subindustries belonging to the same DEA industry within each region.

Appendix M.C Redirecting flows from consumers through retailers

National accounting conventions imply that production in retail industries is given by trade margins, i.e. sales net of acquisition costs. Thus, if a retailer buys a product from a non-retail producer at price p_1 and sells it to a consumer at price p_2, the input-output table will display two money flows: 1) a flow of $p_2 - p_1$from consumers to retail, and 2) a flow of p_1 from consumers to the non-retail producer's industry. This makes the standard input-output table unsuitable for tracing the circular flow of spending starting from the Disaggregated Consumer Spending Matrix. The reason is that the spending matrix shows actual cash-flows from consumers to producers. Thus, in the example above, the spending matrix would show a flow of p_2 from the consumer to the retailer, but no flow from the consumer to the non-retail producer. Tracing the circular flow of spending from consumers and through the production network therefore requires an adjustment to make the two matrices compatible.

We adjust the input-output table by re-coding all flows from consumers to non-consumer facing industries (defined as those not selling directly to consumers) and foreign producers (except

flows related to Danish residents' tourist spending abroad, see subsection Appendix M.D). We replace each of these flows with two equal-sized sets of flows: One from a retail industry to the non-retail industry in question (or foreign producers in the case of imports), and one from private consumers to the same retail industry. Each set splits the value of the original flow across retailers' regions, using the geographic distribution of existing retail-to-non-retail flows of the same type. The choice of retail industry depends on the type of consumption good traded. Exploiting that the input-output table breaks flows from private consumers down on 72 consumption categories, we manually assign each category to the retail industry most likely to sell products within that category.[A5]

Note that this adjustment increases the total sum of flows in the table; for retail industries the total sum of ingoing money flows thus no longer corresponds to the total value of production as defined in national accounts. One way to think of this is that we apply a different definition of production for retailer's sales to consumers: The output value of such sales is now defined as the full amount of the sale (net of taxes such as VAT), while the retailer's acquisition cost is treated as an intermediate input from the non-retail producer of the traded product.

Appendix M.D Foreign residents' spending in Denmark

In the original input-output matrix, consumers' cross-border spending is captured by two separate flows from consumers to foreign producers: a positive representing residents' spending abroad, and a negative representing non-residents' spending in Denmark. The latter adjusts for the fact that the flows from consumers to domestic producers include spending from foreign residents and thus ensures that the total sum of flows from private consumers matches the national accounts value of resident consumers' private consumption expenditure (SNA P.31). To ensure consistency with national acoounts aggregates for imports (SNA P.7) and exports (SNA P.6), the matrix also contains a positive, same-sized flow from foreign consumers (i.e. exports) to foreign producers (i.e. imports).

For our purposes, it is important that the recorded flows from domestic consumers to domestic producers represent the spending of resident consumers only. If not, consistency with the disaggregated consumer spending matrix is compromised. Sales to foreign tourists, on the other hand, should be recorded as exports for the producer cells making the sales. To achieve this, we adjust the flows from private consumers to domestic producers downwards so that the total adjustment across producer cells matches the aforementioned negative flow from consumers to

[A5]For example, if we observe a flow of 1000 from private consumers to manufacturers in region X as payment for cheese,we replace it with 1) flows from grocery retailers in each of the 98 regions to manufacturers in region X, where the size of each flow is 1000 times the origin region's share of total existing flows from grocery retailers to manufacturers in region X; 2) flows from private consumers to grocery retailers in each of the 98 regions, where the size of each flow matches the corresponding retailer-to-manufacturer flow.

foreign producers. We then adjust the total flow from consumers to foreign producers upwards by the same amount. To distribute the total adjustment across producer cells, we draw on two data sources: First, we use the industry distribution of Danish residents' spending abroad from the disaggregated consumer spending matrix to compute a proxy for each industry's share of foreign tourist spending, thus assuming that Danish tourists' spending behavior in other countries is indicative of foreign tourists' spending in Denmark. Second, for the distribution across regions, we use regional data from Visitdenmark.dk on foreigners' overnight stays at hotels and hostel to compute a proxy for each region's share of foreign tourist spending. In an initial round of adjustments, we then allocate the total amount to be adjusted on cells in proportion to the product of these two shares, but subject to a constraint that the adjusted flow must be non-negative. The non-negativity constraint implies that less than the required amount is allocated in the first round, so we distribute the residual amount on the cells with positive tourist spending shares in proportion to their remaining private consumption flows. For each producer cell, we also adjust the flow from foreign consumers (i.e. exports) upwards by the exact same amount so that total flows from consumers (domestic and foreign) to the cell are unchanged. To complete the adjustment process, we adjust the flow from foreign consumers to foreign producers downwards by the same total amount.

Combined, these steps ensure that total consumer spending and total exports are both unaffected, but the split between flows to domestic producers and flows to foreign producers has changed for both. Similarly, total imports are unchanged, and so is total output for each producer cell, but the split between flows from domestic consumers and flows from foreign consumers is different in both cases.

Appendix N Measuring Disaggregated Government Dividend and Surplus Income

The matrix describes the income that the government receives from each producer cell. We start with a list of firms that sell to consumers and producers at market prices are (full or partly) owned by the government.[A6] We manually collect annual turnover for every firm from annual reports. We also collect information on establishment-level employment from the Danish business register (CVR). We combine these two datasets and split annual turnover regionally using each firm's distribution of employment across regions. We finally aggregate across industries (using the industry-code of the parent firm) and regions to get to our level of producer cells. We assume that the share of surplus received by the government is equal to the share of turnover by government-owned establishments in each producer cell.

[A6] `https://fm.dk/arbejdsomraader/statens-selskaber/organisering`

Appendix O Measuring Disaggregated Consumption of Government and NPISH Output

We measure which consumer cells use different types of government services. The government purchases these services from producer cells (specifically, government-operated establishments in each producer cell) and provides them to consumers free of charge. We assume a uniform per capita consumption for collective public goods (including police, national defense, and public administration). We use individual-level data on actual uses of public services to allocate individual public consumption (including education, healthcare, and social protection), as detailed below.

Appendix O.A Education

We assign the aggregate consumption of education services observed in the national accounts to consumer cells according to the number of students in primary, secondary, and tertiary education in a cell. The education register ("UDD") contains information about the education program in which each individual is currently enrolled (if any) as well as each individual's highest completed education. We categorize individuals as primary school students if they are currently enrolled in a program and have no completed education; as secondary school students if they are currently enrolled in a program and their highest completed education is primary school (10 years); as tertiary education students if they are currently enrolled in a program and their highest completed education is secondary school (13 years); and as non-students if they are currently not enrolled in a program.

We aggregate the number of students in primary, secondary and tertiary education to the level of consumer cells. As the cells only include the adult population, we assign the education consumption of minors to adults in the same household before aggregating, drawing on the intra-household links in the population register (BEF). For instance, two parents with three children, two of which are in primary school and one is in secondary school, each consume the equivalent of one year of primary education and half a year of secondary education.

Finally, we allocate the aggregate government consumption of education to cells in proportion to their share of students in primary, secondary and tertiary education and total government expenditure on education at each of these levels. Specifically, the estimated consumption of education services in region-by-industry cell i is given by

$$C_i^{edu} = \sum_{q=p,s,t} \frac{\#students_{i,q}}{\#students_q} \times expenditure_q$$

where q is the level of education (with p, s and t indicating primary, secondary, and tertiary, respectively) and expenditure is government consumption of education of level q.

Appendix O.B Healthcare

Government consumption of healthcare falls into six categories: outpatient services; hospital services; medical products, appliances and equipment; public health services; research and development; and other. We allocate government healthcare consumption summed over all six categories, as reported in national accounts, to consumer cells using publicly available statistics for the first two categories, which make up around 85% of the total.

"Outpatient services" capture government consumption flowing to primary healthcare providers, like general practitioners, specialist doctors, psychiatrists, and dentists. We obtain average primary healthcare expenditures by age, gender, and municipality (Table SYGU1 in the Public Expenditure Statistics). Based on a regression of average primary healthcare expenditure on a set of indicators for age, gender and municipality, we predict primary healthcare expenses for each individual in the population. We then aggregate the predicted expenditures to the level of consumer cells. Since children account for a non-negligible part of healthcare spending, we include the full population by assigning minors to the same consumer cells as the adults cohabiting with the child. If parents live together but work in different industries, we split the child's predicted healthcare expenditure equally between the two cells.

"Hospital services" capture government expenditure related to hospital treatments, including emergency room and outpatient hospital treatments. Again, we obtain information on the average number of days spent in hospital by age, gender, and municipality (Table INDAMP01 in the Health Statistics). Regressing average hospital days on a set of indicators for age, gender, and municipality, we predict the number of hospital days for each individual in the population. We then aggregate the predicted hospital days to the consumer cell level, again allocating the hospital days of minors to their parents' cells.

Finally, we combine the two indicators of healthcare usage to disaggregate total consumption of healthcare services of consumer cell i as:

$$C_i^{health} = \sum_{q=o,h} \left(\frac{usage_{i,q}}{\sum_i usage_{i,q}} \times \frac{exp_q}{exp_o + exp_h} \right) \times exp, \tag{A.4}$$

where q indexes the type of healthcare (with o and h indicating outpatient and hospital services, respectively), $usage_{i,q}$ denotes cell *i's* usage of type q (expenditure on primary care and the number of hospital days), exp_q is national government consumption of healthcare of type q, and exp is national government consumption of healthcare summed over all six categories.

Appendix O.C Social Protection

Government consumption of social protection falls into five categories: sickness and disability; old age; family and children; unemployment; and other. We allocate social protection services to consumer cells based on government microdata. Specifically, we allocate the category "sickness and disability" to cells in proportion to the number of individuals on sick leave or disability pension as observed in the income register (IND); the category "old age" in proportion to the number of individuals aged 80 or older as observed in the population register (BEF); the category "family and children" in proportion to the number of pre-school children as observed in the population register (BEF); the category "unemployment" in proportion to the number of unemployed; and the category "other" by population shares. The estimated consumption of social protection services in consumer cell i is thus:

$$C_i^{social} = \sum_{q=s,o,d,u,z} \frac{usage_{i,q}}{\sum_i usage_{i,q}} \times exp_q, \tag{A.5}$$

where q indexes the type of social protection (with s, o, d, u and z denoting sickness/disability, old age, family/children, unemployment and other, respectively), $usage_{i,q}$ denotes the relevant indicator for cell i's usage of type q (see above), and exp_q is national government consumption of social protection services of type q.

Appendix O.D Measuring Disaggregated Consumption of NPISH Output

We measure the consumption of output provided by non-profit organizations (NPISH) for different consumer cells. NPISH output falls into five categories: education; social work activities; libraries; museums and other cultural activities; sports activities (non-market); and activities of membership organizations.

We first disaggregate usage of the first four categories by consumer region. For education, we use regional data on the share of children attending private schools relatively to public schools (Table "UDDAKT20" in the Education Statistics). For social work, we use regional data on the share of privately-owned (as opposed to government-operated) daycare institutions (Table "BOERN4" in the Child Care Statistics). For libraries, we use regional data on library usage per capita (Table "BIB1" in the Culture and Leisure Statistics). For museums and other cultural activities, we use regional data on members of sports associations per capita.

To infer usage by industry of work, we rely on the Danske Bank data. For education, we proxy use of NPISH education with payments to private schools. For social work, we calculate payments to private child-care institutions. For libraries, we use payments to libraries. For sports activities, we use membership payments to sports associations. For all categories, we count the number of

transactions relative to the number of bank customers in each industry. We thereby estimate how likely consumers in each industry are to consume a given type of NPISH output.

We combine the information on consumption shares of NPISH by consumer region and industry to calculate consumption of NPISH output by consumer cell:

$$npish^{q}_{r,i} = \sum_{q} \frac{npish_{q,r} \times npish_{q,i} \times pop_{r,i}}{\sum_{r,s} npish_{q,r} \times npish_{q,i} \times pop_{r,i}} \times expenditure_{q},$$

where q is the NPISH category, r is region, i is industry, $pop_{r,i}$ is the cell's population, and expenditure is national NPISH output of type q.

For the final type of NPISH consumption, activities of membership organizations, we rely on Danske Bank data. This category consists mostly of trade unions and a small component of political or religious organizations. We disaggregate national consumption using as weights the share of individuals making payments to trade unions in each cell multiplied by the cell's population.

Appendix References

Silva, JMC Santos, and Silvana Tenreyro. 2006. "The log of gravity." *The Review of Economics and statistics*, 88(4): 641–658.