Search Frictions and Efficiency in Decentralized Transport Markets

Giulia Brancaccio[*], Myrto Kalouptsidi[†], Theodore Papageorgiou[‡], Nicola Rosaia[§¶‖]

December 2020

Abstract

We explore efficiency and optimal policy in decentralized transport markets, such as taxis, trucks and bulk shipping. We show that in these markets, search frictions distort the transportation network and the dynamic allocation of carriers over space. We identify the sources of externalities, derive explicit and intuitive conditions for efficiency and show how they translate into efficient pricing rules or optimal taxes and subsidies for the planner who cannot set prices directly. Using data from dry bulk shipping, we find sizeable social loss and spatial misallocation of carriers. Optimal policy restores efficiency by favoring locations that are central in the trade network and might be preferable to centralization.

[*]Cornell University, Department of Economics, 464 Uris Hall, Ithaca NY 14850, USA, giulia.brancaccio@cornell.edu

[†]Harvard University, Department of Economics, Littauer Center 124, Cambridge MA 02138, USA, myrto@fas.harvard.edu

[‡]Boston College, Department of Economics, Maloney Hall 343, Chestnut Hill, MA, 02467, USA, theodore.papageorgiou@bc.edu

[§]Harvard University, Department of Economics, Littauer Center, Cambridge MA 02138, USA, nicolarosaia@g.harvard.edu

[¶]We are thankful to Costas Arkolakis, Nick Buchholz, Pablo Fajgelbaum, Manolis Galenianos, Allen Head, Réka Juhász, Marc Melitz, Robin Lee, Ariel Pakes, and Tobias Salz. We gratefully acknowledge the support of the National Science Foundation (under grant SES-1847555), the Social Sciences and Humanities Research Council (under grant #430-2018-00540), as well as the Griswold Center for Economic Policy Studies at Princeton University.

[‖]The material in Section 4.2 was originally included in a previous working version of Brancaccio et al. (2018) ("Geography, Transportation and Endogenous Trade Costs"; please see NBER Working Paper 23581) but the published version of Brancaccio et al. (2020a) no longer addresses search frictions.

1 Introduction

The transportation sector is vital for economic growth and social development. Generating about 1.2 trillion USD annually and employing 12 million workers in the US alone, the sector is responsible for the movement of both people and goods, making today's interconnected world possible.[1] A substantial fraction of transportation operates in a decentralized fashion; for instance, in the case of taxis, trucks, as well as oceanic bulk shipping, a large number of small firms participate in a decentralized market for trips.[2] In such markets, where search frictions are present, a natural question is whether transport agents are suboptimally allocated over space thus distorting transportation flows. And if so, which policies are best at restoring the optimality of the transport network? In recent years, these questions have captured the interest of industry participants and policymakers alike.

In this paper, we address these questions by studying efficiency and optimal policy in decentralized transport markets. We identify the externalities that stem from search frictions and distort transportation flows, as well as the allocation of carriers over space. We characterize analytically the conditions for (constrained) efficiency and derive both the set of efficient pricing rules, as well as optimal taxes and subsidies. We then show that in the case of dry bulk shipping, search frictions are present and lead to substantial welfare loss and misallocation of ships over space.

We base our analysis on a model for decentralized transport markets that is widely used in the literature; the model originates in Lagos (2000) and has been used in Lagos (2003), Buchholz (2020), Castillo (2019) and Rosaia (2020), among others, for taxis, and in Brancaccio et al. (2020a) for oceanic bulk shipping (henceforth BKP). There is a network of locations at different distances to each other.

[1] U.S. Department of Transportation, Bureau of Transportation Statistics, Transportation Economic Trends and Occupation Employment Statistics of the Bureau of Labor Statistics.

[2] The taxi industry generates 25 billion USD in revenues in the US alone. In many countries, the taxi industry consists of a large number of independent drivers who search for customers on the streets. For instance, in New York City at least 40% of taxis must be operated by individual owners by mandate (while the remaining 60% of medallions, the mini-fleets, are operated by approximately 70 fleet companies), and dispatchers are not permitted (Frechette et al. 2019).
The trucking industry currently generating 800 billion USD in revenues in the US. Trucking is also "extremely fragmented. [...] [In the US there are] 3.68 million Class 8 trucks in operation [...] The top ten full truckload-focused carriers only account for around 5% of total truckload market revenue. According to the American Trucking Association, there are over 892,000 for-hire common carriers registered with the Federal Motor Carrier Safety Administration. 97% of those companies operate 20 or fewer trucks, and 91% operate less than 6 trucks" (source: https://resources.coyote.com/coyote-curve/us-truckload-market). Truckers book their trips through a network of brokers, many of whom are small.
Finally, segments of oceanic shipping including tankers and dry bulk carriers are often termed the "taxis of the oceans" and operate in a similar fashion, as we discuss in detail in Section 4. The largest shipowners account for less than 3-5% of the world fleet depending on the type of carrier and contracts for trips are signed via a network of competitive brokers.

In each location, available carriers (e.g. ships, taxis) and customers (e.g. exporters, passengers) meet randomly. Carriers that find a customer transport them to their desired destination for a price, and restart there. Carriers that do not find a customer decide whether to wait at their current location or travel empty elsewhere to search. Customers that find a carrier, obtain a valuation from arriving at their destination, while customers that do not, wait another period. Finally, every period, potential customers decide whether to start searching for a carrier, as well as their destination, thus replenishing the customer pool seeking transportation. We do not impose restrictions on the price setting mechanism, nor the structure of the demand system, in order to nest different modes of transportation (in taxis prices are regulated, while in shipping they are negotiated).

Studying efficiency in this setup is not straightforward due to the dynamic nature of decision-making and the spatial network; yet, we are able to obtain analytical results. In particular, we show that the market equilibrium can be written as a planner's problem where any externality is taken as given. This allows us to pinpoint the different externalities that are present in this setting and derive conditions for each one to be internalized.

We show that search frictions create two types of externalities in transport markets. First, as is well-known in other settings, they generate thin/thick market externalities: when choosing whether to search, carriers and customers affect the finding probabilities faced by other agents both in the same and in the opposite side of the market. If agents' search decisions do not internalize this effect, the overall *number* of agents searching may be distorted away from the efficient one.

Thin/thick market externalities are internalized in equilibrium if and only if the private returns from searching are equal to the social returns. This amounts to the so-called "Hosios (1990) conditions" on surplus sharing: these conditions, which are known to characterize efficiency in search models with homogeneous agents, require the share of the surplus which is appropriated by agents on each side of the market to be equal to the elasticity of the matching function with respect to the same side.

Second, search frictions generate what we call "pooling externalities." Once a carrier drops off a customer at their destination, it needs to restart its search there. As a result, customers' destination decisions have an impact on the distribution of carriers over space. If customers fail to internalize it, the *composition* of realized trips will be distorted and carriers will be misallocated over space, leading to suboptimal transport networks.

Customers internalize pooling externalities in equilibrium if and only if prices are such that carriers receive the same surplus regardless of the customer they match with. This condition for efficiency replicates the no-arbitrage condition obtained in a frictionless world, where competition among carriers ensures that prices coincide with the opportunity cost of each trip, until in equilibrium carriers are indifferent among different customers. Absent this condition, the price paid by customers for a trip may not reflect its social value and the share of destinations with high social value may be too low in equilibrium.

The two efficiency conditions combined characterize the efficient pricing rule, which is useful if a central authority is able to set prices, as in the case of taxicabs. In many markets, however, the planner is not able to directly control prices, but he may be able to impose taxes or subsidies. We show that, when prices are set via Nash bargaining, the planner can achieve efficiency using these instruments and we derive their optimal values.

We apply these results to study empirically the dry bulk shipping industry, often termed the "ocean taxis", where a large number of small homogeneous ship(owner)s meet exporters on a trip by trip basis, through a disperse network of brokers, to arrange the international transportation of raw materials. A number of features of dry bulk shipping, such as information frictions and port infrastructure, can hinder the matching of ships and exporters. We begin by leveraging a rich dataset of vessel movements and bulk shipping prices to document the presence of search frictions. To do so, we propose a novel test to argue that these frictions lead to unrealized potential trade. The test is based on weather shocks at sea that exogenously shift ship arrivals at port: in a frictionless world, in regions with more ships than exporters, the change in the number of ships should not affect matching, since the short side of the market is always matched. Here, instead, matches are indeed affected by weather shocks. We further corroborate the presence of frictions by documenting substantial price dispersion and wasteful ship movements.

We next turn to our welfare analysis. Using the estimates from BKP, combined with aggregate trade data, we first test whether the observed equilibrium is efficient by checking whether our derived conditions for efficiency hold in the data. Perhaps not surprisingly, we find that neither condition is satisfied, suggesting that the market does not operate efficiently.

We next compare (i) the market equilibrium; (ii) the constrained efficient outcome, i.e., the market allocation under the efficient prices; (iii) the first-best, i.e., the allocation in a world without search frictions. We find that total welfare in the market equilibrium allocation is 6% lower than the constrained efficient allocation and 14% lower than the frictionless benchmark. Moreover, trade volume and net trade value are substantially higher under constrained efficiency, as well as under the first-best. This suggests that the externalities have a substantial impact on world trade.

Our results show that under the optimal policy, the market is able to achieve about 44% of the first-best welfare gains. If the first-best allocation is interpreted as a benchmark achievable by a platform that eradicates search frictions (like Uber/Lyft), our results imply that policymakers can improve efficiency substantially through simple policies, such as taxes or subsidies, without resorting to some form of centralization. This is important because centralization may be infeasible in practice, or may come with market power if provided by private firms.

We next delve into the different role of the two externalities. Thin/thick market externalities decrease trade volume substantially: at the observed equilibrium, the entry of an additional exporter has a substantial positive externality on matching rates, but prices remain too high to achieve the socially efficient number of exporters. Thus to correct the thin/thick market externalities the planner heavily subsidizes exporters.

In contrast, pooling externalities have an impact on trade value, as they distort the composition of export destinations and favor destinations with low social value. So which routes get subsidized? First, routes with high export value. In addition, however, the planner aims to optimize the entire dynamic path of ships' travels. To do so, the planner subsidizes routes whose destinations are themselves high-value exporters, and which further export to high export value locations, and so on. For instance, the highest subsidy is awarded to trips towards the West Coast of North America (WCNA), as well as Australia, which (i) are high-value importers; (ii) high-value exporters, thus offering high continuation values to ships that arrive there; (iii) export to regions that themselves are high-value exporters, so that ships that leave WCNA or Australia travel to locations where their options are good. In other words, we find that the planner subsidizes locations that are central in the trade network.

Finally, although the efficient prices and optimal taxes that restore efficiency have known expressions, they may not be feasible to implement in practice, either because the planner does not have access to all instruments; or because they may be too complex or computationally challenging. We thus consider simple policies that are designed to mimic the optimal taxes but may be more easily implementable. We find that a destination-specific tax (customs tax) performs relatively well. In contrast, a tax that is a function of distance achieves no welfare gains, suggesting that such a pricing scheme (indeed used in taxis), is far from efficient. Explicitly targeting both origin and destination is essential in order to correct for the different sources of externalities.

Related Literature

This paper broadly relates to several strands of literature: transportation and trade; search and matching; and industry dynamics.

First, our paper contributes to a large, diverse, and rapidly growing literature on transportation. The study of optimal transport networks dates back to Koopmans (1949), while Fajgelbaum and Schaal (2020) and Allen and Arkolakis (2019) have recently revisited this question within a general equilibrium spatial trade framework. Different from our focus, they consider optimal routing and optimal investment in infrastructure.

Within the context of urban transportation, a series of recent empirical papers have studied issues related to efficiency. Frechette et al. (2019) and Buchholz (2020) both study search frictions and regulation frictions in NYC taxicabs. In particular, Buchholz (2020) numerically implements tariff pricing changes in order to explore whether welfare improvements can be achieved. Frechette et al. (2019) investigate the welfare impact of changes in the number of active medallions, as well as the introduction of an "Uber-like" platform. Shapiro (2018) and Liu et al. (2019) explore the welfare improvements from different centralizing formats; while several papers study platform pricing (e.g. Bian, 2020, Ma et al., 2018, Castillo, 2019).[3] We contribute to this literature by (i) formally examining the sources of inefficiencies in these markets, (ii) providing explicit conditions for efficiency and (iii) deriving optimal policy.

Moreover, since our empirical application involves international oceanic transportation, we relate to a literature studying transportation in the context of international trade; e.g. Hummels and Skiba (2004),

[3] In addition, Ghili and Kumar (2020) investigate demand and supply imbalances in ride-sharing platforms; Ostrovsky and Schwarz (2018) focus on carpooling and self-driving cars; Kreindler (2020) studies optimal congestion pricing; Cao et al. (2018) explore competition in bike-sharing platforms; Rosaia (2020) studies platform competition in the NYC ride-hail industry.

Asturias (2018), Brooks et al. (2018), Cosar and Demir (2018), Holmes and Singer (2018), Wong (2018), Ducruet et al. (2019), Lee et al. (2020).[4] In our prior work, BKP, we explore the role of the transportation sector in world trade and spell out the impact of endogenous trade costs. Although we rely on the model setup and empirical strategy employed there, our focus here is different, as this paper provides a formal treatment of efficiency in decentralized transport markets and an empirical analysis of search frictions and welfare loss in shipping as a case study.[5]

In addition, our paper relates to the literature on efficiency of search models in the spirit of Diamond (1982), Mortensen (1982) and Pissarides (1985). Hosios (1990) considers efficiency in markets with random search and Nash bargaining. He shows that these markets are generically inefficient and derives the well-known "Hosios condition" that restores efficiency. Acemoglu and Shimer (1999) and Acemoglu (2001) show that with random search and heterogeneous agents the Hosios condition does not guarantee efficiency, and that labor market policies, such as unemployment benefits or minimum wages, can potentially improve welfare. Our paper extends the existing literature by deriving explicit conditions for efficiency with random search and heterogeneous agents on one side of the market. In particular, our main theorem shows that efficiency is restored if two conditions hold: first, the Hosios (1990) conditions that ensures that the number of matches in every market is optimal;[6] second, a no arbitrage/indifference condition that ensures that the composition of matches in every market is optimal. This latter condition is novel. In addition, we derive theoretically the set of policy instruments (both efficient pricing rules, and taxes/subsidies) that can restore efficiency.

Finally, we relate to the literature on industry dynamics (Hopenhayn, 1992, Ericson and Pakes, 1995), while our empirical methodology borrows from the literature on the estimation of dynamic setups (e.g.

[4] We also relate to a literature in international trade studying the role of frictions, such as Antras and Costinot (2011), Eaton et al. (2016), and Krolikowski and McCallum (2018) who consider search frictions between importers and exporters and Allen (2014) who investigates information frictions.

[5] Let us outline other similarities and differences of the two papers. Both papers use the same framework, which, as already mentioned, originates in Lagos (2000) and is currently being widely used in spatial research (for the sake of completeness– compared to BKP here we allow for a more general customer demand function, as well as any price mechanism to nest several modes of transport). The present paper however provides a theoretical contribution in deriving efficiency conditions and optimal policies. Moving to the empirics, the contribution here is to (i) provide a formal test for frictions; (ii) analyze welfare loss in shipping. In contrast, the empirical contribution of BKP centered on estimating trade elasticities, providing a methodology to estimate this framework, and performing a series of counterfactuals capturing the impact of transportation on world trade (e.g. a slow-down in China, or the opening of the Northwest passage).

[6] Extending the literature, we also show how to restore efficiency in the case of decreasing returns to scale in the matching technology. In addition, compared to Hosios (1990), we consider a more general setup with a network of interconnected markets and without the Nash bargaining assumption on pricing.

Rust, 1987, Bajari et al., 2007, Pakes et al., 2007; applications include Ryan, 2012, Collard-Wexler, 2013, Kalouptsidi, 2014, 2018). Related to this paper, a small literature lying in the intersection of search and industry dynamics, has explored trading frictions in decentralized markets (e.g. Gavazza, 2011, 2016 for real assets and Brancaccio et al., 2020c for over-the-counter financial markets).

The rest of the paper is structured as follows: Section 2 presents the model. Section 3 provides the efficiency and optimal policy results. Section 4 describes the dry bulk shipping industry and the data used, presents evidence for search frictions and outlines the estimation of the model. Section 5 presents our welfare analysis. Section 6 concludes.

2 Model

We present a model of decentralized transport markets that focuses on the interaction between carriers (e.g. ships, taxis, trucks) and customers (e.g. exporters, passengers). As mentioned, this framework has been used fairly widely in the literature; here we follow the presentation of BKP. We allow the model to be flexible in terms of customer demand and the pricing mechanism in order to nest different modes of transport.

2.1 Environment

Time is discrete and the horizon is infinite. There are I locations, $i \in \{1, 2, ..., I\}$. There are two types of agents: customers and carriers. Both are risk neutral and have discount factor β. Variables with superscript s refer to carriers and e to customers, in line with our empirical exercise of ships and exporters.

There is a measure S of homogeneous carriers in the economy.[7] At the beginning of every period, a carrier is either in some region i, or traveling full or empty, from some location i to some location j. Carriers at i can either search or remain inactive. The per-period payoff of staying inactive is set equal to 0 at each location, while searching carriers incur a per-period cost c_i^s. Carriers traveling from i to j incur a per period traveling cost c_{ij}^s. The duration of a trip between location i and location j is stochastic: a

[7]A constraint on the fleet size is consistent with most applications of interest, and can be due to either regulatory constraints (e.g. fixed number of medallions) or time to build.

traveling carrier arrives at j in the current period with probability d_{ij}, so that the average duration of the trip is $1/d_{ij}$.[8]

Customers can only be delivered to their destination by carriers and each carrier can carry at most one customer. Following the search and matching literature, we model the number of matches that take place every period in region i, m_i, using a matching function, whereby

$$m_i = m_i(s_i, e_i) \leq \min\{s_i, e_i\}$$

where s_i is the measure of available carriers in region i and e_i is the number of available customers in region i. $m_i(s_i, e_i)$ is increasing and concave in both arguments. We allow for the possibility that $m_i(s_i, e_i) < \min\{s_i, e_i\}$ creating the potential for unrealized trade: two agents searching in the same location might fail to meet, due to impediments such as information frictions or physical constraints. As Petrongolo and Pissarides (2001) note, "[...] the matching function [...] enables the modeling of frictions [...] with a minimum of added complexity. Frictions derive from information imperfections about potential trading partners, [...] slow mobility, congestion from large numbers, and other similar factors."

Since search is random, the probability according to which customers in i meet a carrier is $\lambda_i^e = m_i(s_i, e_i)/e_i$, which is the same for all customers. Similarly the probability according to which carriers in i meet a customer is $\lambda_i^s = m_i(s_i, e_i)/s_i$.

When a carrier and a customer meet, if they both accept to match, the customer pays a price τ_{ij} upfront and the carrier begins its trip immediately to j. This price is determined differently in different transport markets: for instance, prices are fixed by regulation in taxicabs, while prices are bilaterally negotiated in bulk shipping and trucking; our results nest these different cases.

Carriers that remain unmatched decide whether to stay in their current region or travel empty to a different region and search there. Customers that remain unmatched wait in their current region. Inactive carriers restart the following period in the same region.

Finally, every period, at each location i, a large pool of potential customers decide whether to enter and search for a carrier, in order to be transported to a destination $j \neq i$, subject to an entry cost κ_{ij}.

[8] It is straightforward to have deterministic trip durations instead. Our specification, however, preserves tractability and allows for some variability e.g. due to weather/traffic shocks, without affecting the steady state properties of the model.

Denote by e_{ij} the endogenous measure of customers in i who search for transportation to j. The total measure of customers searching at location i is $e_i = \sum_{j \neq i} e_{ij}$, while G_{ij} is the share of demand routed from i to j, i.e.,

$$\forall i, j : G_{ij} \equiv e_{ij}/e_i.$$

Once they have entered, customers pay a per-period waiting cost c^e_{ij}.[9]

Upon matching with a carrier, customers obtain a valuation from being transported from origin i to destination j. We model customer valuations via the function, $w : \mathbb{R}^{I \times I}_+ \to \mathbb{R}^{I \times I}_+$, where $w_{ij}(q)$ is the valuation of the marginal customer on route ij, and q is the matrix with typical element q_{ij} denoting the quantity transported every period (i.e. the measure of accepted matches) on route ij. This can be thought of as an inverse demand curve for transportation services, before customer entry and waiting costs. For example, consider customers with heterogeneous valuations for transportation (e.g. passengers looking for taxis with different value of time): when q_{ij} matches are formed on route ij, $w_{ij}(q)$ describes the valuation of the q_{ij}-th (i.e. the marginal) consumer entering route ij.[10] As a simpler case, if valuations are homogeneous so that all customers obtain w_{ij} on route ij, the marginal customer naturally also obtains w_{ij}.

2.2 Behavior and equilibrium

We consider the steady state of our industry model. In a steady state equilibrium, customers and carriers respond optimally to their expectations of the endogenous market variables, which are consistent with agents' behavior (rational expectations) and constant over time. We begin by describing the optimal behavior of carriers and customers facing a *given* tuple $\tau, \lambda^s, \lambda^e, G$. We then endogenize these variables to achieve market clearing.

Carrier optimality Let V^s_{ij} denote the value of a carrier that begins the period traveling from i to j (empty or loaded), V^s_i the value of a carrier that begins the period in location i, and U^s_i the value of a carrier that remained unmatched at i at the end of the period. In everything that follows, we suppress

[9]By convention we set $c^e_{ii} = 0$, $\kappa_{ii} = 0$, and $c^s_{ii} = 0$.

[10]In addition, valuations might depend on quantities through general equilibrium effects; for instance in a general equilibrium trade model, traded goods' prices depend on traded quantities.

the dependence of the value functions on the state of the economy, given our focus on a steady state; in Appendix E we consider out of steady state dynamics. Given prices and meeting rates we have,

$$V_{ij}^{s} = -c_{ij}^{s} + d_{ij}\beta V_{j}^{s} + (1 - d_{ij})\,\beta V_{ij}^{s} \tag{1}$$

In words, a carrier that is traveling from i to j: pays the per period cost of traveling c_{ij}^{s}; with probability d_{ij} it arrives at destination j where it begins with value V_{j}^{s}; with the remaining probability $1 - d_{ij}$, the carrier does not arrive and keeps traveling.

A carrier that starts the period in region i obtains:

$$V_{i}^{s} = \max\left\{-c_{i}^{s} + \lambda_{i}^{s}\sum_{j\neq i} G_{ij}\max\left\{\tau_{ij} + V_{ij}^{s}, U_{i}^{s}\right\} + (1-\lambda_{i}^{s})\,U_{i}^{s}, \beta V_{i}^{s}\right\}.$$

In words, if the carrier decides to search, it pays the per period wait cost c_{i}^{s}; with probability $\lambda_{i}^{s}G_{ij}$ it meets a customer heading to destination j, in which case it accepts to match and transport the customer if and only if the value from doing so, inclusive of the price τ_{ij}, is higher than its outside option, U_{i}^{s}, otherwise it receives the outside option U_{i}^{s}. With probability $1-\lambda_{i}^{s}$, the carrier does not meet a customer and receives value U_{i}^{s}. If the carrier remains inactive, it obtains a flow payoff of zero and restarts the following period at the same location.

Defining the carrier meeting surplus as,

$$\Delta_{ij}^{s} = \max\left\{\tau_{ij} + V_{ij}^{s} - U_{i}^{s}, 0\right\} \tag{2}$$

the carrier's value V_{i}^{s} can be written as follows,

$$V_{i}^{s} = \max\left\{-c_{i}^{s} + \lambda_{i}^{s}\sum_{j\neq i} G_{ij}\Delta_{ij}^{s} + U_{i}^{s}, \beta V_{i}^{s}\right\}. \tag{3}$$

Next, if the carrier remains unmatched, it chooses where to search: it can either keep waiting at i or travel empty to another location. The unmatched carrier value function is equal to:

$$U_{i}^{s} = \max_{j} V_{ij}^{s} \tag{4}$$

where we set $V_{ii}^s \equiv \beta V_i^s$. In words, if the carrier stays in region i, at the beginning of the next period it will be waiting at i; otherwise if the carrier chooses another region $j \neq i$ it begins its trip towards j.

Having defined all the carrier value functions, we now characterize their optimal behavior in terms of the three decisions they make (whether to search in the beginning of the period, whether to accept a match and where to search if unmatched at the end of the period). First, carriers search only when it is profitable to do so, so that from equation (3),

$$s_i > 0 \to V_i^s = -c_i^s + \lambda_i^s \sum_{j \neq i} G_{ij} \Delta_{ij}^s + U_i^s. \tag{5}$$

Second, the following holds[11]

$$q_{ij} < \lambda_i^s s_i G_{ij} \to \Delta_{ij}^s = 0 \tag{6}$$

$$q_{ij} > 0 \to \Delta_{ij}^s = \tau_{ij} + V_{ij}^s - U_i^s \tag{7}$$

In words, if not all meetings result into matches, then it must be that the carrier surplus is zero, since at least some of the carriers preferred to remain unmatched. In addition, if there is realized trade on route ij, it must be that the carrier surplus is positive and thus the carrier accepts the match.

Third, carriers choose where to search when unmatched. Denote by b_{ij} the measure of carriers who decide to relocate empty from i to j (and let b_{ii} be the measure that decides to remain in i); optimality requires that b_{ij} is positive only if option j achieves the maximum value across all possible choices:

$$b_{ij} > 0 \to U_i^s = V_{ij}^s. \tag{8}$$

Finally, it must be the case that, whenever the measure of inactive carriers is greater than zero, there is some location where some carriers would rather not search at all. Since $q_{ij} + b_{ij}$ carriers depart from i towards j every period, traveling for $1/d_{ij}$ periods on average, the total measure of active carriers in steady state is given by $\sum_{ij} (q_{ij} + b_{ij}) / d_{ij}$ (setting $d_{ii} = 1$). Hence this condition can be written as,

[11]Note that we allow for the number of meetings (i.e., $m_{ij} \equiv G_{ij} m_i = s_i \lambda_i^s G_{ij}$) to be higher than the number of realized matches (i.e., q_{ij}) since agents can reject a match upon meeting. However, in equilibrium generically no rejections occur, since a customer would not enter the market, only to have a match with a carrier rejected later. Therefore $q_{ij} = m_{ij}$.

$$\sum_{ij} \frac{q_{ij} + b_{ij}}{d_{ij}} < S \to \exists i : V_i^s = 0. \tag{9}$$

Customer optimality We now turn to the value functions of customers; we begin with existing customers and then consider customer entry. If a customer meets a carrier they can either agree to form a match, in which case the customer pays price τ_{ij} and is transported to the desired destination, receiving its valuation, or the customer can revert to its outside option and stay unmatched. Hence the meeting surplus of the marginal customer with valuation $w_{ij}(q)$ is given by,

$$\Delta_{ij}^e = \max\left\{w_{ij}(q) - \tau_{ij} - \beta U_{ij}^e, 0\right\}, \tag{10}$$

where U_{ij}^e is its value of searching for a carrier in i with destination j:

$$U_{ij}^e = -c_{ij}^e + \lambda_i^e \left(\Delta_{ij}^e + \beta U_{ij}^e\right) + \left(1 - \lambda_i^e\right) \beta U_{ij}^e = -c_{ij}^e + \lambda_i^e \Delta_{ij}^e + \beta U_{ij}^e. \tag{11}$$

In words, the customer pays the cost c_{ij}^e while waiting; then with probability λ_i^e it meets a carrier and receives the meeting surplus on top of its outside option, while with the remaining probability it remains unmatched and receives its outside option.

Similarly to carriers, optimality requires that the following holds,

$$q_{ij} < \lambda_i^e e_{ij} \to \Delta_{ij}^e = 0 \tag{12}$$

$$q_{ij} > 0 \to \Delta_{ij}^e = w_{ij}(q) - \tau_{ij} - \beta U_{ij}^e \tag{13}$$

Finally, the measure of customers entering on each route ij is pinned down by a free entry condition for the marginal customer:

$$U_{ij}^e - \kappa_{ij} \leq 0, \text{ with equality if } e_{ij} > 0. \tag{14}$$

We adopt the convention that customers in i choosing i do not enter, and normalize the payoff in that case to zero.

Feasible allocations An allocation for the transportation economy consists of a tuple (s, E, q, b) where $s = [s_1, \ldots, s_I]$ denotes the measure of carriers waiting in each region, $E \in \mathbb{R}_+^{I\times I}$, with typical element e_{ij}, denotes the measure of customers waiting for transport on each route ij, $q \in \mathbb{R}_+^{I\times I}$ denotes the measure of new matches formed on each route, and $b \in \mathbb{R}_+^{I\times I}$ denotes the measure of carriers departing empty on each route. Equivalently, we will sometimes denote an allocation by (s, e, G, q, b), where $e = [e_1, \ldots, e_I] = \left[\sum_j e_{1j}, \ldots, \sum_j e_{Ij}\right]$ denotes the measure of customers waiting in each region. This will be useful when we want to emphasize the implications of search frictions on transport flows captured by G.

An allocation is feasible if it satisfies: (i) the set of constraints defining a steady state, equations (15) and (16) below; (ii) the total fleet constraint, equation (17) below; and (iii) the constraints on the transported quantities imposed by the meeting technology, equation (18) below. Thus, the following feasibility constraints must hold:

$$\sum_j (q_{ij} + b_{ij}) = \sum_j (q_{ji} + b_{ji}), \ \forall i \tag{15}$$

$$s_i = \sum_j (q_{ij} + b_{ij}), \ \forall i \tag{16}$$

$$\sum_{ij} \frac{q_{ij} + b_{ij}}{d_{ij}} \leq S \tag{17}$$

$$q_{ij} \leq m_i (s_i, e_i) G_{ij}, \ \forall ij \tag{18}$$

The first set of constraints requires that the measure of carriers departing at any given location i equals the measure of arrivals, so that flows into i are equal to the flows out of i. Equation (16) requires that the measure of carriers searching at each location (s_i) must equal those that will be matched ($\sum_j q_{ij}$), those that will be unmatched and choose to remain (b_{ii}) and those that will be unmatched and decide to travel elsewhere empty ($\sum_{j\neq i} b_{ij}$). Equation (17) imposes the fleet constraint. Finally, constraints (18) require that the number of matches does not exceed the number of meetings between carriers and customers.

Equilibrium We now define the equilibrium of the transportation economy, which is a tuple (s, E, q, b, τ) consisting of an allocation (s, E, q, b) and prices, $\tau \in \mathbb{R}_+^{I\times I}$.

Definition 1. An outcome (s, E, q, b, τ) is a steady state equilibrium under prices τ if:

1. (s, E, q, b) satisfies the feasibility constraints (15)-(18).

2. (s, q, b) satisfies the carrier optimality conditions (1)-(9) given τ, λ^s and G.

3. E, q satisfies the customer optimality and free entry conditions (10)-(14) given τ, λ^e.

4. The perceived meeting probabilities are consistent with the true ones, i.e., for all i, j, $\lambda_i^s = m_i(s_i, e_i)/s_i$, $\lambda_i^e = m_i(s_i, e_i)/e_i$ and $G_{ij} = e_{ij}/e_i$.

(s, E, q, b) is an equilibrium allocation if there exists a price matrix τ such that (s, E, q, b, τ) is an equilibrium outcome.

This definition characterizes the entire set of equilibria that can arise under different assumptions on the pricing mechanism. For instance, both exogenous prices in the case of taxicabs, as well as bilateral bargaining in the case of trucking and bulk shipping, select a subset of the equilibria defined in Definition 1.

3 Efficiency

In this section we present our efficiency results. We first identify the sources of externalities in this market by comparing the social planner's problem to the market equilibrium allocation. Next, we provide our main theorem that states the conditions for efficiency. We also discuss the efficient pricing rules. Finally, we derive taxes and subsidies that restore efficiency, when prices are set via Nash bargaining.

3.1 Externalities

We begin by characterizing the set of equilibrium allocations as agents become patient, i.e. as $\beta \to 1$, and compare them to the social planner's solution. Focusing on the case of patient agents simplifies the dynamic problem at hand without sacrificing its essential features. Patient (in the limit) agents only care about their average payoff in the steady state and not about the transition dynamics. In Appendix E we demonstrate that our efficiency results hold with discounting, as well as out of steady state.

Definition 2. (s, E, q, b, τ) is a limit equilibrium outcome if there exists a sequence $(s^n, E^n, q^n, b^n, \tau^n, \beta^n)_{n \geq 0}$ such that: (i) for each n, $(s^n, E^n, q^n, b^n, \tau^n)$ is an equilibrium outcome for the economy populated by agents with discount factor β^n; and (ii) as $\beta^n \to 1$, $(s^n, E^n, q^n, b^n, \tau^n) \to (s, E, q, b, \tau)$. (s, E, q, b) is a limit equilibrium allocation if there exists a price matrix τ such that (s, E, q, b, τ) is a limit equilibrium outcome.

Turning to the planner, when agents do not discount future payoffs, the (constrained) efficient steady state allocation is a solution to the following problem,

$$\max_{s,E,q,b\geq 0} W(q) - \sum_{ij} q_{ij}\kappa_{ij} - \sum_{ij} (q_{ij} + b_{ij}) \frac{c^s_{ij}}{d_{ij}} - \sum_i s_i c^s_i - \sum_{ij} e_{ij} c^e_{ij} \tag{19}$$

s.t. feasibility constraints (15)-(17)

$$\forall i, j : q_{ij} \leq m_i (s_i, e_i) \frac{e_{ij}}{e_i} \tag{20}$$

where $W : q \mapsto \mathbb{R}_+$ is customer welfare, whose derivative is given by the marginal customer value $w(q)$.[12]

In words, the social planner maximizes the per-period welfare corresponding to (s, E, q, b): every period, q_{ij} customers depart on each route ij, generating gross customer value equal to $W(q)$; matched customers are replaced by a pool of new entrants of equal measure who pay the one-time entry cost κ_{ij}; $q_{ij} + b_{ij}$ carriers begin traveling on route ij for $1/d_{ij}$ periods on average, paying a per-period traveling cost c^s_{ij}; at every location i, s_i carriers pay the waiting cost c^s_i, and e_i customers incur the waiting cost c^e_{ij}. The planner is subject to the set of steady state feasibility constraints (15)-(17) and (20) which simply replaces the definition of $G_{ij} = e_{ij}/e_i$ in (18). Note that since we focus on constrained efficiency, the planner is subject to the same frictions as the market.

Comparing the socially optimal allocation to the market one is not straightforward, since neither one has a closed-form expression. Indeed, the market equilibrium allocation solves a nonlinear system of equalities and inequalities, as described in Definition 1 (agent optimality conditions and value functions, feasibility constraints and rational expectations constraints), while the efficient allocation solves the planner's constrained optimization Problem (19) above. The following theorem establishes that the market allocation can be found by solving an optimization problem that is similar, in form, to the planner's problem.

[12]Formally, $W : q \mapsto \mathbb{R}_+$ is defined as any function whose gradient is $w(q)$.

Theorem 1. *If (s, E, q, b) is a limit equilibrium allocation then (i) it solves*

$$\max_{s,E,q,b\geq 0} W(q) - \sum_{ij} q_{ij}\kappa_{ij} - \sum_{ij}(q_{ij}+b_{ij})\frac{c^s_{ij}}{d_{ij}} - \sum_i s_i c^s_i - \sum_{ij} e_{ij}c^e_{ij} \tag{21}$$

s.t. feasibility constraints (15)-(17)

$$\forall i,j:\ q_{ij} \leq \lambda^s_i s_i G_{ij} \tag{22}$$

$$\forall i,j:\ q_{ij} \leq \lambda^e_i e_{ij}. \tag{23}$$

for given λ^s, λ^e, G and (ii) λ^s, λ^e, G satisfy $\lambda^s_i = m_i(s_i, e_i)/s_i, \lambda^e_i = m_i(s_i, e_i)/e_i$ and $G_{ij} = e_{ij}/e_i$.

Theorem 1 states that any equilibrium allocations must solve Problem (21), the "market problem". Put differently, if there exists a price matrix τ that sustains (s, E, q, b) as an equilibrium allocation, then the allocation is a solution to Problem (21).[13]

The proof of Theorem 1, provided in Appendix A, rests heavily on duality. In particular, the dual variables of the market Problem (21) are linked to the carrier and customer value functions. This, in turn allows us to show that the carrier optimality conditions, equations (1)-(9), and the customer optimality conditions, (10)-(14), are equivalent in the limit to the first order conditions of the market Problem (21).[14]

Inspecting the market problem (21) we note that as in the planner Problem (19), the objective function is equal to total welfare. Moreover, both the market and the planner face the steady state constraints (15)-(16), and the total fleet constraint (17). However, when it comes to the matching constraints, Problems (19) and (21) differ. Indeed, the social planner faces constraint (20), which treats the meeting rates λ^s, λ^e and the destination shares G as endogenous objects that are functions of s, e; in contrast, constraints (22) and (23) in the market Problem (21) treat these objects as exogenous constants. In other words, the market Problem (21) does not recognize that in fact $\lambda^s_i = m_i(s_i, e_i)/s_i$, $\lambda^e_i = m_i(s_i, e_i)/e_i$ and

[13] It can be shown that Theorem 1 fully characterizes equilibrium allocations, so that it is an "if and only if" statement rather than "if", when agents are patient, i.e. they do not discount future payoffs. However, taking the limit of equilibria as $\beta \to 1$ involves a selection of these outcomes. For this reason, there might be allocations satisfying the conditions of Theorem 1 that cannot be obtained as limits of equilibria with discounting. This selection might arise when Problem 21 has multiple solutions, which can be the case only if there are situations where agents are indifferent among different alternatives. "Smoothing out" agents' choices by means of idiosyncratic shocks, as we do in the empirical analysis, resolves this multiplicity. In this case, Theorem 1 becomes a full characterization of limit equilibrium allocations.

[14] In part, the technical difficulty of proving the result is related to taking limits as the discount factor goes to one, because the value functions per se may diverge. Once this is resolved, we can compare the market optimality conditions to the first order conditions of the market Problem (21). Detailed arguments are found in the Appendix A.

$G_{ij} = e_{ij}/e_i$, i.e., that by changing s, E, q, b in the optimization, the matching rates $\lambda_i^s, \lambda_i^e, G_{ij}$ are also bound to change.

Theorem 1 pinpoints the unique potential sources of inefficiency in this economy, some of which may not have been obvious ex ante. It shows that there are two externalities, one related to the matching rates, λ^s, λ^e and one to the destination shares, G.[15]

First, when choosing whether to join the search pool, agents may not internalize the effect their entry has on the matching probabilities faced by other agents (λ^s and λ^e). Indeed, an extra carrier makes it easier for customers to find a carrier and harder for other carriers to find a customer; and similarly for an extra customer. These are known as "thin/thick market externalities" in the literature.

Second, when choosing their destination, customers may not internalize the effect that this choice has on the distribution of carriers over space: a carrier will have to take the customer to his destination, and restart its search there. The customer only cares about its private surplus of the trip, whereas the planner also cares about relocating a carrier from the origin of the trip to its destination. The random matching process creates what we call "pooling externalities": as customers with different destinations are *pooled* together in the matching process, prices may fail to fully capture the social value of a match between a carrier and a customer. This affects the customers' destination decisions, which in turn distort transport patterns and misallocate carriers over space.

We now discuss these externalities formally. Define the social value of triplet s, e, G by

$$V^p(s,e,G) \equiv \max_{q,b\geq 0} W(q) - \sum_{ij} q_{ij}\kappa_{ij} - \sum_{ij}(q_{ij}+b_{ij})\frac{c_{ij}^s}{d_{ij}} - \sum_i s_i c_i^s - \sum_i e_i \sum_j G_{ij} c_{ij}^e \tag{24}$$

$$\text{s.t. feasibility constraints (15)-(18)}$$

This problem essentially solves for the carriers' optimal relocation decisions (b) and the decision of whether to accept a match or not (q), while taking as given the entry decisions of carriers (s) and customers

[15] As we discuss later, the inefficiency stemming from inefficient search decisions in a market (which govern λ_i^s and λ_i^e) is well-known in the search literature and might not come as a surprise to readers. However the inefficiency related to destination decisions (which determine G_{ij}) is not as obvious and, to our knowledge, has not been previously identified.

(e), as well as customers' destination decisions (G). The social planner Problem (19) is equivalent to,[16]

$$\max_{s,e,G\geq 0} V^p\left(s,e,G\right), \quad \text{s.t. } \sum_j G_{ij}{=}1, \forall \text{i and } \sum_i s_i \leq \text{S} \tag{25}$$

Intuitively, since the only source of inefficiency results from agents' search behavior captured by the triplet (s, e, G), it is useful to "optimize out" the other variables (i.e. q, b) in order to focus on the impact of the main variables of interest, s, e, G.

Definition 3. At a triplet (s, e, G):

- Carriers internalize thin/thick market externalities if

$$s \in \arg\max_{s'\geq 0} V^p\left(s', e, G\right) \text{ s.t. } \sum_i s_i \leq S. \tag{26}$$

- Customers internalize thin/thick market externalities if

$$e \in \arg\max_{e'\geq 0} V^p\left(s, e', G\right). \tag{27}$$

- Customers internalize pooling externalities if

$$G \in \arg\max_{G'\geq 0} V^p\left(s, e, G'\right) \text{ s.t. } \sum_j G_{ij} = 1\, \forall i. \tag{28}$$

3.2 Efficiency Conditions

Our next theorem states three conditions for the externalities to be internalized in equilibrium. For every $i \in I$, we denote by $\eta_i^s = d \ln m_i\left(s_i, e_i\right)/d \ln s_i$ and $\eta_i^e = d \ln m_i\left(s_i, e_i\right)/d \ln e_i$, the elasticities of the matching function with respect to the measure of carriers and customers searching at i, respectively. To avoid delving into corner solutions arising in trivial cases, we assume that the equilibrium is such that there is a positive measure of customers and carriers searching at each location ($s_i, e_i > 0\ \forall i$) and that $\sum_i s_i < S$.[17] Let $\bar{\Delta}^s$ and $\bar{\Delta}^e$ denote the carrier and customer limit surpluses associated with the limit

[16] Note that for every feasible solution s, e, G of this problem there exists a pair $q, b \geq 0$ such that the resulting allocation is steady state feasible: simply set $q = 0$, $b_{ij} = 0$ for $i \neq j$ and $b_{ii} = s_i\ \forall i$.

[17] If $s_i = 0$ or $e_i = 0$ for some i, then the efficiency conditions must hold only at locations with positive number of carriers and customers. The second is a non-triviality assumption which precludes having no carriers traveling at all, since $\sum_i s_i = S$

equilibrium outcome, (s, e, G, q, b, τ), when $\beta \to 1$. In addition, let $\bar{\Delta}_{ij} \equiv \bar{\Delta}^s_{ij} + \bar{\Delta}^e_{ij}$, $\forall i, j$ denote the limit social surplus. For a formal definition, see Appendix A.1.

Theorem 2. *Let (s, e, G, q, b, τ) be a limit equilibrium outcome. Suppose that Problem (24) admits a unique optimal solution.*[18] *Then:*

(i) Carriers internalize thin/thick market externalities if and only if[19]

$$\forall i \in I : \frac{\sum_j G_{ij} \bar{\Delta}^s_{ij}}{\sum_j G_{ij} \bar{\Delta}_{ij}} = \eta^s_i. \tag{29}$$

(ii) Customers internalize thin/thick market externalities if and only if

$$\forall i \in I : \frac{\sum_j G_{ij} \bar{\Delta}^e_{ij}}{\sum_j G_{ij} \bar{\Delta}_{ij}} = \eta^e_i. \tag{30}$$

(iii) Customers internalize pooling externalities if and only if

$$\bar{\Delta}^s_{ij} = \max_{k \neq i} \bar{\Delta}^s_{ik} \tag{31}$$

for every ij such that $G_{ij} > 0$.

(iv) (s, e, G, q, b) is efficient only if conditions (i)-(iii) hold jointly.

The proof, provided in Appendix A.3, first establishes that the function $V^p(s, e, G)$ is concave. Therefore the supergradients with respect to each of the arguments (s, e, G) are well-defined for every such triplet and in fact V^p is differentiable almost everywhere in its domain. Then we demonstrate, through the use of the dual variables associated with a limit equilibrium allocation, that the resulting first order conditions coincide with the conditions internalizing the three externalities.

Conditions (i) and (ii) of Theorem 2 recast the familiar Hosios (1990) conditions requiring that the share of the surplus appropriated by agents on each side of the market equals the elasticity of the matching function with respect to the measure of agents on that side of the market.They have a similar flavor as the

implies that $(q_{ij}, b_{ij}) = (0, 0)$ for every ij such that $d_{ij} < 1$.

[18] This is a technical condition which is generally satisfied, for example, when the function W is strictly concave. Note that this does not imply that we are requiring a unique equilibrium.

[19] Formally, this condition is necessary only when $V^p(s, e, G)$ is differentiable in s, which is the case almost everywhere. A similar disclaimer applies to statement (ii), (iii) and (iv), where necessity relies on differentiability with respect to e, G and (s, e, G), respectively.

standard Coasian conditions in the presence of externalities, where the private value of an action must be equal to its social value. Indeed, we can rewrite equation (29) as

$$\lambda_i^s \sum_j G_{ij} \bar{\Delta}_{ij}^s = \frac{dm_i(s_i, e_i)}{ds_i} \sum_j G_{ij} \bar{\Delta}_{ij}.$$

The left-hand-side captures the per-period expected private return of a carrier entering market i, which equals the expected carrier surplus from matching ($\sum_j G_{ij} \bar{\Delta}_{ij}^s$) multiplied by its matching probability (λ_i^s). The right-hand-side captures the per-period expected social return from an additional carrier entering i, which equals the expected social surplus from an additional match ($\sum_j G_{ij} \bar{\Delta}_{ij}$) multiplied by the marginal increase in the number of matches ($\frac{dm_i(s_i,e_i)}{ds_i}$).

Condition (iii) of Theorem 2 deals with the pooling externalities and it requires that in each origin i carriers are indifferent across all possible destinations j. To understand how this condition is derived, note that since the customers make the destination decision, they are the ones creating the externality. Therefore, the planner wants to align their private incentives to the social ones. To this end, he would like to set any differences across destinations in the consumers' surplus $\bar{\Delta}_{ij}^e$, equal to differences across destinations in the social surplus $\bar{\Delta}_{ij}$. To do so, he sets differences in the carrier surplus, $\bar{\Delta}_{ij}^s$, across destinations equal to zero.

Pooling externalities distort customers' travel patterns and misallocate carriers over space, leading to suboptimal transport networks. This happens because customers might fail to internalize fully the impact of relocating a carrier from the trip's origin to its destination. To gain intuition on this condition, consider a frictionless environment where competition among carriers ensures that prices coincide with their opportunity cost; in that case, as in Condition (iii), carriers are indifferent among serving different types of customers in equilibrium. In our setup customer destination-specific markets are missing: when heterogeneous customers are pooled together, carriers cannot compete among themselves to serve a particular type of customer; this grants carriers market power, creating a wedge between prices and the carriers' opportunity cost of a trip.[20]

[20]Similarly, in a world with search frictions, but where carriers can direct their search to a specific customer type (i.e. a model of directed search which is efficient, see Moen, 1997), in equilibrium a no-arbitrage condition would make carriers indifferent between searching across destinations; reminiscent of Condition (iii) of Theorem 2. In this case, there is a separate market for each customer destination; carriers enter different markets, until in equilibrium they are (ex ante) indifferent across different choices. Markets of more desirable destinations entail longer waiting times for carriers and vice versa.

Moreover, conditions (i) and (ii) imply that a necessary condition for efficiency is that the matching function exhibits constant returns to scale; indeed, if we add equations (29) and (30), the left-hand-side is equal to one, and thus the elasticities must add to one as well. Corollary 3 in Appendix A.5 demonstrates that under non-constant returns to scale in matching, efficiency can still be restored, via a tax or subsidy which however creates a wedge between the price paid by the customer and the one received by the carrier.

Finally, in Appendix E we relax the assumptions of no discounting and steady state. In this case, we cannot prove an analogous Theorem 1; but having identified the sources of externalities from the no discounting case (see Theorem 1) we are able to show that the conditions of Theorem 2 still hold, along the path, at each time period. The proofs are more complicated, as we need to work with the planner's and market's discounted stream of payoffs, but no additional restrictions are required.

3.3 Efficient prices

We next use Theorem 2 to derive the efficient prices. More specifically, Condition (iv) of Theorem 2 provides a characterization of the efficient pricing rule:

Corollary 1. *Let a limit equilibrium outcome* (s, e, G, q, b, τ) *be efficient. Then we have* $\eta_i^s = 1 - \eta_i^e$ *and the equilibrium prices are such that, for every* ij *such that* $G_{ij} > 0$*:*

$$\tau_{ij} = w_{ij}(q) - \kappa_{ij} - \eta_i^e \sum_j G_{ij} \bar{\Delta}_{ij} - \left(\bar{\Delta}_{ij} - \sum_j G_{ij} \bar{\Delta}_{ij} \right). \tag{32}$$

To gain some intuition for this pricing rule, we can show that (32) can be rewritten as follows:

$$\forall i, j : \, (1 - \eta_i^s) \, \bar{\Delta}_{ij}^s = \eta_i^s \left[\bar{\Delta}_{ij}^e - \left(\bar{\Delta}_{ij} - \sum_j G_{ij} \bar{\Delta}_{ij} \right) \right], \tag{33}$$

where the terms $\bar{\Delta}_{ij}^s$ and $\bar{\Delta}_{ij}^e$ depend on the price τ_{ij}. Relationship (33) is reminiscent of a surplus sharing condition under Nash bargaining (i.e. the equilibrium condition that determines prices when agents Nash bargain in a decentralized fashion), where however we have (i) replaced the bargaining coefficients with the respective matching function elasticities (this amounts to satisfying the Hosios condition under Nash bargaining, see Section 3.4); and (ii) adjusted the outside option of the customer by the deviation of route

ij's social surplus from the average social surplus across destinations. By adjusting the outside option of the customer, we ensure that customers fully internalize the social value of their destination in their decision-making. If the customer has chosen a destination whose social surplus is higher than the mean from origin i, he should enjoy a higher outside option (and thus a lower price), and vice versa.

Corollary 1 suggests that if a central authority can post prices, they should choose them according to (32).[21] For instance, in the case of taxicabs, prices are regulated and are roughly set equal to a tariff plus a fee proportional to distance. Corollary 1 indicates that this pricing rule is unlikely to be efficient, since the efficient prices should be origin-destination specific. In other markets prices cannot be fully regulated, so in the next section we consider optimal policy when prices are bilaterally negotiated, as is often the case in decentralized markets.

3.4 Optimal policy under Nash bargaining

In this section we consider the problem of a planner who cannot directly control prices, but can impose taxes/subsidies. We show that the planner can indeed achieve efficiency using such instruments and we derive their optimal values. We focus on a specific price mechanism, that of Nash bargaining, which is commonly employed to capture bilateral negotiations. In that case, prices are determined by the usual surplus sharing condition,

$$(1-\gamma_i)\,\bar{\Delta}^s_{ij} = \gamma_i \bar{\Delta}^e_{ij} \tag{34}$$

where γ_i is the carrier bargaining coefficient at i. Suppose that the planner can impose a one-time tax/subsidy h^q on loaded trips and a per-period tax/subsidy h^s on searching carriers, and h^e on searching customers (see Appendix A.6 for a full description of the setup). Corollary 2 derives the tax scheme h that resolves the two externalities:

Corollary 2. *Let (s, e, G, q, b, τ) be a limit equilibrium outcome under taxes h and Nash bargaining. Then:*

[21] That said, equation (32) does not provide the prices in closed-form, since the right-hand-side also depends on them. Hence, to recover the efficient prices one needs to compute the equilibrium under the pricing rule defined by (32) via for instance a fixed point algorithm. We propose such an algorithm in Appendix C.3 where we provide the algorithm we use in our empirical analysis to perform counterfactuals. In our experience the algorithm is extremely robust and has always provided a unique solution.

(i) Thin/thick market externalities are internalized if and only if for every i

$$\gamma_i \sum_j G_{ij}\bar{\Delta}_{ij} - \left(\frac{h_i^s}{\lambda_i^s} + \gamma_i \sum_j G_{ij} h_{ij}^q\right) = \eta_i^s \sum_j G_{ij}\bar{\Delta}_{ij} \tag{35}$$

and similarly,

$$(1-\gamma_i)\sum_j G_{ij}\bar{\Delta}_{ij} - \left(\frac{h_i^e}{\lambda_i^e} + (1-\gamma_i)\sum_j G_{ij} h_{ij}^q\right) = \eta_i^e \sum_j G_{ij}\bar{\Delta}_{ij}. \tag{36}$$

(ii) Pooling externalities are internalized if and only if for all ij

$$h_{ij}^q - \sum_j G_{ij} h_{ij}^q \leq -\frac{\gamma_i}{1-\gamma_i}\left(\bar{\Delta}_{ij} - \sum_j G_{ij}\bar{\Delta}_{ij}\right) \tag{37}$$

with equality if $G_{ij} > 0$.

Before discussing the result, note that if the planner does not impose any taxes, so that $h = 0$, the conditions required to internalize the thin/thick market externalities (35) and (36) boil down to

$$\gamma_i = \frac{d\ln m_i(s_i, e_i)}{d\ln s_i} \equiv \eta_i^s, \quad \text{and} \quad 1-\gamma_i = \frac{d\ln m_i(s_i, e_i)}{d\ln e_i} \equiv \eta_i^e, \tag{38}$$

which are the well-known Hosios (1990) conditions. In addition, notice that, regardless of who pays the matching tax h_{ij}^q, we can think of $\gamma_i h_{ij}^q$ and $(1-\gamma_i)\, h_{ij}^q$ as, respectively, the incidence of this tax on carriers and customers since Nash bargaining implies that the agents split the gross surplus from the match according to γ_i.[22] Moreover, a waiting carrier in expectation pays h_i^s/λ_i^s and a waiting customer pays h_i^e/λ_i^e.

Similarly to Condition (29) of Theorem 2, Condition (35) sets the tax incidence of the carrier so that the private value of an additional carrier searching in i must equal its social value. In particular, he sets the share of the total surplus accruing to the carrier (which is equal to his bargaining coefficient times the total surplus) minus his tax incidence, equal to the contribution of the extra carrier to total surplus.

Condition (37) determines the trip taxes that resolve the pooling externalities. At first glance, it

[22]To see this note that the social surplus of a trip is now defined by $\bar{\Delta}_{ij} = \bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e + h_{ij}^q$ (as either $\bar{\Delta}_{ij}^s$ or $\bar{\Delta}_{ij}^e$ include $-h_{ij}^q$, $\bar{\Delta}_{ij}$ does not depend on h_{ij}^q) and therefore $\bar{\Delta}_{ij}^s = \gamma_i\left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e\right) = \gamma_i\left(\bar{\Delta}_{ij} - h_{ij}^q\right)$. Note also that the definition of efficiency remains the same, except that we include the planner's revenue in the welfare.

has an intuitive explanation: destinations with a surplus above average should receive a subsidy above average and vice versa. Next, note that the system of equations in (37) is not full rank; for simplicity we can set the planner revenue in region i, $\sum_j G_{ij} h^q_{ij}$, equal to zero.[23] Multiplying both sides by $-(1-\gamma_i)$, it is easy to see that Condition (37) requires that the subsidy on route ij that falls on the customer, $(1-\gamma_i)(-h^q_{ij})$, is equal to the deviation of the carrier surplus, $\gamma_i \bar{\Delta}_{ij}$ from the average carrier surplus from i, $\gamma_i \sum_j G_{ij} \bar{\Delta}_{ij}$. Therefore, routes where the carrier surplus is high (low) are subsidized (taxed). By setting the customer tax/subsidy equal to the deviation of the carrier surplus, the planner forces the customer to fully internalize the impact of his destination decision on the carrier surplus.

Finally, note that if the planner can only use the search taxes h^s, h^e, he can correct the thin/thick market externalities.[24] Similarly if he can tax only matches but not search of any side of the market, then he can correct the pooling externalities (using equation (37) as discussed above). The planner can correct all externalities by taxing matches and either waiting carriers or waiting customers.[25]

4 Empirical application: dry bulk shipping

In this section we describe our empirical application using data from the dry bulk shipping industry. We begin in Section 4.1 with a brief description of the industry and the available data. In Section 4.2 we discuss search frictions in this market. In Section 4.3 we briefly discuss estimation and results. Throughout the following sections, unless otherwise noted, we split ports into 15 geographical regions, depicted in Figure 7 of Appendix D. The material in Section 4.2 used to be included in a previous working version of BKP (NBER Working Paper 23581).

4.1 Industry description and data

Dry bulk shipping involves vessels designed to carry a homogeneous unpacked dry cargo for individual shippers. Dry bulk shipping is the prevalent mode of transportation for international trade in commodities,

[23] Condition (37) defines a linear system of equations in terms of the $I-1$ trip taxes h^q_{ij} for each location i. This system has multiple solutions as its rank equals $I-2$. Thus, to obtain a unique solution we would have to impose a linear constraint. Imposing the constraint $\sum_j G_{ij} h^q_{ij} = 0$ is natural as it implies that the budget is balanced in each location.

[24] He can do so by setting $h^e_i/\lambda^e_i = (1-\gamma_i) \sum_j G_{ij} \bar{\Delta}_{ij} - \eta^e_i \sum_j G_{ij} \bar{\Delta}_{ij}$ and $h^s_i/\lambda^s_i + h^e_i/\lambda^e_i = (1-\eta^e_i - \eta^s_i) \sum_j G_{ij} \bar{\Delta}_{ij}$.

[25] If he taxes matches and waiting carriers, he sets $(1-\gamma_i) h^q_{ij} = (1-\gamma_i) \bar{\Delta}_{ij} + \sum_j G_{ij} \bar{\Delta}_{ij} - \bar{\Delta}_{ij} - \eta^e_i \sum_j G_{ij} \bar{\Delta}_{ij}$, if $G_{ij} > 0$ and $h^s_i/\lambda^s_i + \sum_j G_{ij} h^q_{ij} = (1-\eta^e_i - \eta^s_i) \sum_j G_{ij} \bar{\Delta}_{ij}$.

such as iron ore, steel, coal, bauxite, phosphates, but also grain, sugar, chemicals, lumber and wood chips; it accounts for about half of total seaborne trade in tons (UNCTAD) and 45% of the total world fleet, which includes also containerships and oil tankers.[26]

Bulk carriers operate much like taxi cabs: a specific cargo is transported individually by a specific ship, for a trip between a single origin and a single destination. Ships carry the cargo of one customer at a time, who fills up the entire ship. In this paper, we focus on spot contracts and in particular the so-called "trip-charters", in which the shipowner is paid with a per day rate.[27]

Dry bulk shipping is a decentralized market with many small shipowners (see Kalouptsidi, 2014) and exporters, who find each other using brokers. "The broker's task is to discover what cargoes or ships are available....[M]ost owners and charterers use one or more brokers ....often in tense competition with [each] other" (Stopford, 2009).

There are four size categories of dry bulk carriers: Handysize (10,000–40,000 DWT), Handymax (40,000–60,000 DWT), Panamax (60,000–100,000 DWT) and Capesize (larger than 100,000 DWT). Beyond this, shipping services are largely perceived as homogeneous. In his lifetime, a shipowner will contract with hundreds of different exporters, carry a multitude of different products and visit numerous countries. We discuss the industry further at the end of Section 4.2.

We combine the following data sets. First, we use a data set of shipping contracts, from 2010 to 2016, collected by Clarksons Research. An observation is a transaction between a shipowner and a charterer for a specific trip and reports the vessel, the charterer, the contract signing date, the loading and unloading dates, the price in dollars per day, as well as some information on the origin and destination. Second, we use AIS (Automatic Identification System) data from exactEarth Ltd for the ships in the Clarksons data set in the same time period. We obtain their position (longitude and latitude), speed, and level of draft (the vertical distance between the waterline and the bottom of the ship's hull), at regular intervals of at most six minutes. The draft is a crucial variable, as it allows us to determine whether a ship is loaded or not at any point in time. Since the data is virtually the same as in BKP, we refer the interested reader

[26]Bulk ships are different from containerships, which carry cargo (mostly manufactured goods) from many different cargo owners in container boxes, along fixed itineraries according to a timetable. It is not technologically possible to substitute bulk with container shipping.

[27]Trip-charters are the most common type of contract. Long-term contracts ("time-charters"), however, do exist: on average, about 10% of contracts signed are long-term. Interestingly, though, ships in long-term contracts, are often "relet" in a series of spot contracts, suggesting that arbitrage is possible.

there for summary statistics (see Table 1 in BKP) and for details on the construction of the final dataset. Finally, we augment the two ship data sets, with international trade data from Comtrade on export value and volume by country pair for bulk commodities.

Some salient facts An important feature of this market revealed by the AIS data, is that most countries are either large net importers or large net exporters. For instance, Australia, Brazil and Northwest America, the world's biggest exporters of commodities, are rich in minerals, grain, coal, etc. At the same time, China and India, the world's biggest importers, require raw materials to grow further. As a result, commodities flow out of the former, towards the latter. The trade imbalances have implications for both ship ballast (i.e. unloaded travel) behavior and shipping prices. Indeed, at any point in time, 42% of ships are traveling without cargo. At the same time, prices are largely asymmetric and depend on the destination's trade imbalance: all else equal, the prospect of having to ballast after offloading is associated with higher shipping prices (see results in Table 9).

4.2 Search frictions in dry bulk shipping

A number of features of dry bulk shipping, such as information frictions and port infrastructure, can hinder the matching of ships and exporters. In this section we argue that these frictions indeed lead to unrealized potential trade. Consider a geographical region, such as a country or a set of neighboring countries, where there are s ships available to pick up cargo and e exporters searching for a ship to transport their cargo. We define search frictions by the inequality:

$$m < \min\{s, e\} \tag{39}$$

where m is the number of matched ships and exporters. In other words, under frictions there is potential trade that remains unrealized; in contrast, in a frictionless world, the entire short side of the market gets matched, so that $m = \min\{s, e\}$. When inequality (39) holds, matches are often modeled via a matching function, $m = m(s, e)$, as is done in Section 2 above, and also extensively in the labor literature.

We present three facts consistent with frictions, as defined by (39). In particular, we (i) provide a direct test for inequality (39); (ii) we document wastefulness in ship loadings; (iii) we document substantial price dispersion.

First, we provide a simple test that directly verifies the presence of search frictions. If we observed all variables s, e, m, it would be straightforward to test (39); this is essentially what is done in the labor literature, where the co-existence of unemployed workers and vacant firms is interpreted as evidence of frictions. However in our setup, we observe m (i.e. ships leaving loaded) and s, but not e; we thus need to adopt a different approach.

Assume there are more ships than exporters, i.e. $\min(s, e) = e$. We begin with this assumption, because our sample period is one of low shipping demand and severe ship oversupply due to high ship investment between 2005 and 2008 (see Kalouptsidi, 2014). If there are no search frictions, so that $m = \min(s, e) = e$, small exogenous changes in the number of ships should not affect the number of matches. In contrast, if there are search frictions, an exogenous change in the number of ships changes the number of matches, through the matching function $m = m(s, e)$. We approximate an exogenous change in the number of ships, with unpredictable ocean weather conditions. The intuition is that wind affects the speed at which ships travel and therefore exogenously shifts the supply of ships at port. We therefore explore whether exogenously changing the number of ships in regions with a lot more ships than exporters affects the realized number of matches.[28] Since we do not observe exporters directly, to select periods in which there are more ships than exporters, for each region we consider weeks when there are at least twice as many ships waiting in port as matches. Table 1 presents the results: matches are affected by weather conditions in all but one region, consistent with the presence of search frictions. It is worth noting that although search frictions have been prevalent in the recent literature on transport markets, to our knowledge we are the first to propose a formal test for their presence.

Second, we document simultaneous arrivals and departures of empty ships. Indeed, the first two panels of Figure 1 display the weekly number of ships that arrive empty and load, as well as the number of ships that leave empty, in two net exporting countries: Norway and Chile. In Norway, several ships arrive empty and load, but almost no ship departs empty. In Chile, however, the picture is quite different: it frequently happens that an empty ship arrives and picks up cargo, while at the same time another ship departs empty. This is suggestive of wastefulness in Chile: why does the ship that depart empty, not pick

[28]We collect global data on daily sea weather from the ERA-Interim archive, from the European Centre for Medium-Range Weather Forecasts (CMWF). This allows us to construct weekly data on the wind speed (in each direction) across all oceans (see Table 1 for details on constructing the instrument).

	N	Joint Significance	$\frac{s}{m}$
North America West Coast	193	0	2.706
North America East Coast	200	0	3.172
Central America	199	0.001	3.451
South America West Coast	198	0	2.913
South America East Coast	200	0	4.083
West Africa	200	0.001	5.862
Mediterranean	200	0	4.244
North Europe	200	0	3.577
South Africa	200	0	2.862
Middle East	200	0	3.86
India	200	0.34	8.58
South East Asia	200	0	3.334
China	200	0.038	6.194
Australia	187	0	2.457
Japan-Korea	200	0	5.311

Table 1: Test for search frictions. Regressions of the number of matches in each region on the unpredictable component of weather conditions in the surrounding seas. For each region we use weeks in which there are at least twice as many ships as matches. The first column reports the number of observations; the second column joint significance; and the third column the average ratio of ships over matches in each region during these weeks. To proxy for the unpredictable component of weather, we partition the globe into cells of $9° \times 9°$, and for each cell we collect data on the speed of the horizontal (E/W) and vertical (N/S) component of wind, as well as wave period and height. To control for seasonality, we residualize the weather measurements for each cell on a quarter fixed effect. The potential regressors include one and two weeks lagged values of all the weather measurements for cells in the sea. Finally, we follow Belloni et al. (2012) to select the relevant instruments in each region.

up the cargo, instead of having another ship arrive from elsewhere to pick it up?

This pattern is observed in many countries. Indeed, the third panel of Figure 1 depicts the histogram of the bi-weekly ratios of outgoing empty ships over incoming empty and loading ships for net exporting countries. In the absence of frictions, one would expect this ratio to be close to zero. However, we see that the average ratio is well above zero. Moreover, this pattern is quite robust in a number of dimensions.[29]

Third, again inspired by the labor literature, we investigate dispersion in prices. In markets with no frictions, the law of one price holds, so that there is a single price for the same service. This does

[29]This figure is robust to alternative market definitions, time periods and when constructed separately for each vessel size. Capesize vessels exhibit somewhat larger mass towards zero, consistent with the somewhat higher concentration of ships and charterers, as well as the ships' ability to approach fewer ports. The figure is also similar if done by port rather than country. To control for repairs we remove stops longer than 6 weeks. Finally, we only consider as "ships arriving empty" the ships arriving empty and sailing full towards another region, and we consider as "ships leaving empty" ships sailing empty toward a different country; so movements to nearby ports are excluded. This definition also implies that refueling cannot explain the fact either- though there are very small differences in fuel prices across space anyway (less than 10%).

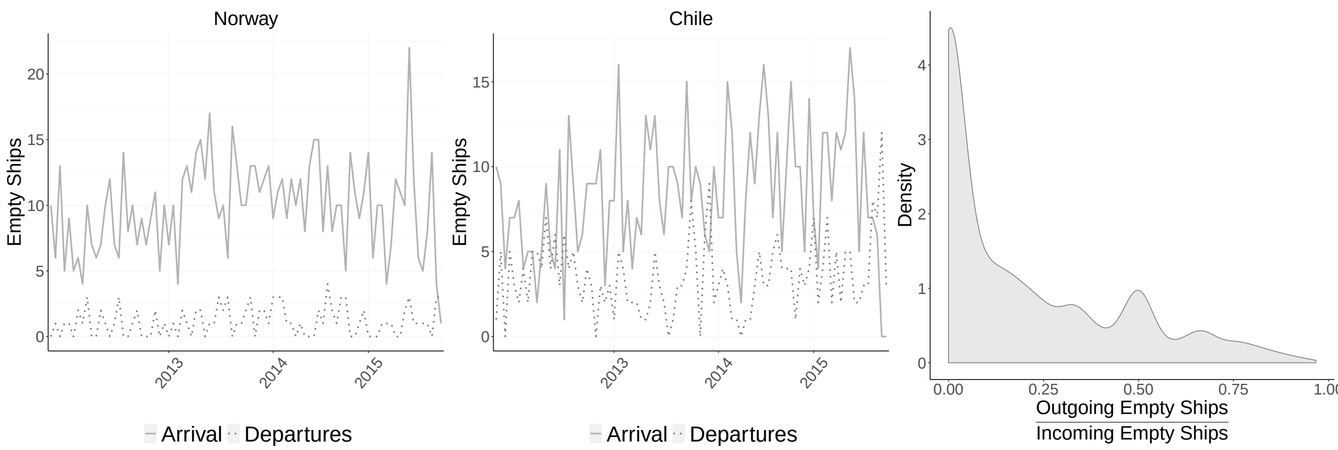


Figure 1: Simultaneous arrivals and departures of empty ships: The first two panels depict the flow of ships arriving empty and loading, and ships leaving empty in two-week intervals in Norway and Chile. The last panel shows the histogram of the ratio of outgoing empty over incoming empty and loading ships across all net exporting countries.

not hold in labor markets, where there is large wage dispersion among workers who are observationally identical.[30] Table 9 in Appendix D shows that there is substantial price dispersion in shipping contracts. More specifically, at best we can account for about 70% of price variation, controlling for ship size, as well as quarter, origin and destination fixed effects. Moreover, the coefficient of variation of prices within a given quarter, origin and destination triplet is about 30% (23%) on average (median). In the most popular trip, from Australia to China, the *weekly* coefficient of variation is on average 34% and ranges from 15% to 65% across weeks.

In addition, the type of product carried affects the price paid and overall more valuable goods lead to higher contracted prices, as shown in the same table. In the absence of frictions, if there are more ships than exporters, as is the case during our sample period, we would expect prices to be bid down to the ships' opportunity cost.[31] In contrast, in markets with frictions and bilateral bargaining, since ships now have market power, the price depends on the exporter's valuation and higher value exporters pay more.

As in labor markets, a multitude of factors can lead to frictions (i.e. unrealized matches) in shipping. First, the decentralized and unconcentrated nature of the market and the mere existence of brokers, suggest that information frictions are present. The meeting process involves a disperse network of brokers;

[30]This fact has generated an influential literature on frictional wage inequality. See for instance Mortensen (2003) and references therein.

[31]In a frictionless market with more ships than exporters and homogeneous ships, in equilibrium the price from an origin to a destination would be such that ships are indifferent between transporting the cargo and staying unmatched.

oftentimes more than two brokers intervene to close a deal, suggesting that the ship's and the exporter's brokers do not always find each other, and that an "intermediate broker" was necessary to bring the two together (Panayides, 2016). In interviews, brokers claimed to receive 5,000-7,000 emails per day; sorting through these emails is reminiscent of an unemployed worker sorting through hundreds of vacancy postings. Port infrastructure, congestion and regulations may also hinder matching.

Are bulk ships similar to taxis? It may seem surprising at first that we make a parallel between the bulk shipping industry and the taxi industry. Yet, it is remarkable how similar the operation of ships (as well as trucks) is to that of taxicabs. In fact, shipowners themselves call bulk ships "the ocean taxis", as a bulk ship "does not operate on a fixed schedule but merely trades in all parts of the world in search of suitable cargoes" (Plakantonaki, 2010, which is the internal manual of a large shipping firm). Much like taxis, bulk ships transport one cargo at a time and are homogeneous; they largely operate on a voyage basis; they sign contracts very close to loading; and they operate in a decentralized environment without dispatchers. Let us discuss two assumptions that are particularly important: ship homogeneity and random search.

Ships do not specialize neither geographically, nor in terms of products: during the period of our data ships deliver cargo to 13 out of 15 regions on average and carry at least 2 of the 3 main products (coal, ore and grain). Moreover, neither shipowner characteristics, nor shipowner fixed effects have any explanatory power in price regressions, as shown in Table 4 in Appendix B.1. Furthermore, we fail to reject that ships' ballast destination decisions are the same across ships, from a given origin (See Appendix B.1).[32]

Next, we argue that random search is a reasonable approximation in shipping. Meetings occur through an unconcentrated network of brokers. The contracting process begins when a shipowner's broker "is contacted by a competitive broker by email or phone or yahoo/MSN messenger for a specific cargo" (Plakantonaki, 2010). In Appendix B.2, we copy an extract from Plakantonaki (2010) with an example of a representative email notifying brokers of an available cargo. Roughly, the email has the following format: "Please indicate or offer a vessel to be chartered for CORUS to load a full cargo of bulk coal, size

[32]If heterogeneity were an important driver of ships' ballasting decisions, we would expect ships to choose diverse destinations from a given origin. Yet ballast choices are similar across ships from a given origin (the CR_3 measure for the chosen destinations, i.e. the concentration ratio of the top 3 destinations, is higher than 70% in most regions). Moreover, home-ports are not an important consideration for shipowners, as the crew flies to their home country every 6-8 months.

50,000 +/- 5%; load in Poland and discharge in Immingham; vessel must appear between 4/23 and 4/30". As mentioned above, the broker receives thousands such emails a day. Continuing from Plakantonaki (2010), the broker "considers whether he has a ship available, *usually in the vicinity*, which can safely arrive at the load port during the requested laycan". Indeed in our data contracts are signed on average only 6 days prior to loading and ships are in the vicinity of the loading area.

Next we examine the random search assumption more formally. In particular, in Appendix B.2 we investigate a standard implication of directed search: whether matching rates differ across destinations from a given origin. Directed search would imply that carriers choose to search in a specific "market", i.e. a market for customers heading to a specific destination. Under directed search, profitable markets attract more carriers, thereby reducing their matching probabilities compared to less profitable markets. However, we find that in a given origin, i, ships' matching probabilities do not differ across destinations j, while prices exhibit high variance; see Appendix B.2 for details. This finding suggests that differences in the attractiveness of different destinations is reflected in prices, but not in matching probabilities, as would be the case in directed search.[33]

4.3 Estimation and selected results

The estimation strategy, which is presented in detail in Appendix C, proceeds in two steps. In the first step we recover the matching functions $m_i(s_i, e_i)$ for all i, as well as the ship travel and wait costs c^s_{ij}, c^s_i, for all i, j. For both objects we directly rely on the estimates of BKP. Let us emphasize that our estimation strategy avoids parametric restrictions on the matching function; this is important, since as shown in Theorem 2, in frictional markets, the conditions for constrained efficiency depend crucially on the elasticity of the matching function with respect to the search input.

The second step of the estimation consists of recovering exporter demand for transportation services. To do so, we assume that prices are determined via Nash bargaining, with γ_i denoting with the ship's bargaining coefficient in region i. Importantly, we allow the bargaining coefficient to vary by region for greater flexibility, given the importance of that parameter regarding the thin/thick market externalities (indeed, this is the main reason we are re-estimating the model rather than using all BKP estimates

[33]We also test whether the distribution of trips' destinations for loaded trips varies across ships, after controlling for the trip's origin. As detailed in Appendix B.2, we reject the null that (loaded) ships travel to systematically different destinations.

directly). We also assume for simplicity that $w_{ij}(q) = w_{ij}$ for all ij. In sum, the parameters we estimate are: the exporter valuations w_{ij}, wait costs c^e_{ij}, and entry costs κ_{ij} for all i, j; and the ship bargaining coefficients γ_i for all i. As detailed in Appendix C.1, we rely on data on export values from Comtrade to measure w_{ij} and we estimate the remaining parameters from shipping prices and trade flows.

Figure 6 in Appendix C.1 presents the estimated search frictions, computed as the average percentage of weekly "unrealized" matches; i.e. $(\min\{s_i, e_i\} - m_i) \,/\, \min\{s_i, e_i\}$ in every region. Search frictions are heterogeneous over space and may be somewhat sizable, with up to 20% of potential matches "unrealized" weekly in regions like South and Central America and Europe. On average, 13.5% of potential matches are "unrealized".[34]

Estimated search frictions are positively correlated with the observed within-region price dispersion (0.47), another indicator of search frictions. Moreover, frictions are negatively correlated with the Herfindahl-Hirschman Index of charterers reported in the Clarksons contract data(-0.31); this suggests that when the clientele is disperse, frictions are higher. Finally, when we estimate the matching function separately for Capesize (biggest size) and Handysize (smallest size) vessels; as expected, for Capesize, where the market is thinner, search frictions are lower.

The exporter parameters are presented in Table 8 in Appendix D. The exporter wait costs, c^e_i, are equal to about 3% of the exporters' valuation on average, but there is substantial heterogeneity over space; the estimated costs are highest in Central and South America, as well as parts of Africa. These parameters capture inventory expenditures, delay costs, risks of damage or theft etc. Consistent with this interpretation, exporter wait costs are positively correlated with the recovered wait costs for ships (0.34), and are negatively correlated with the World Bank index of quality of port infrastructure (-0.50). Finally, the estimates for the bargaining coefficients suggest that exporters get a somewhat larger share of the surplus in almost all regions.

[34]It is worth noting that this does not imply that in the absence of search frictions there would be 13.5% more matches, as we would need to take into account the optimal response of ships and exporters. This is simply a measure of the severity of search frictions in different regions.

5 Efficiency in dry bulk shipping

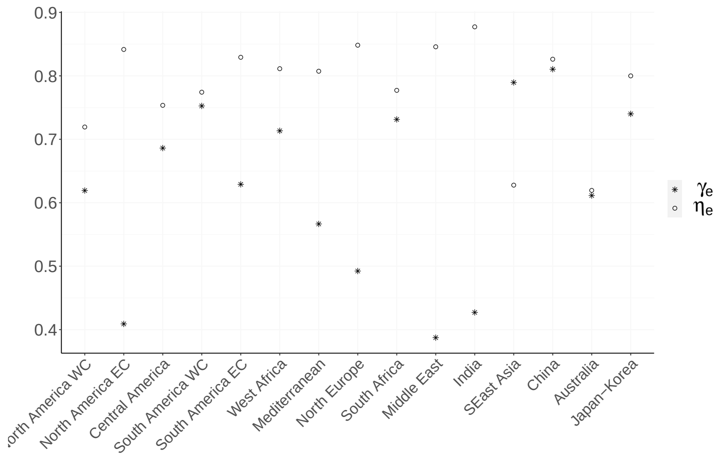
In this section we present our welfare results. In Section 5.1 we check whether the efficiency conditions hold in the case of bulk shipping. In Section 5.2 we present our main welfare analysis and in Section 5.3 we discuss policy implementation.

5.1 Is dry bulk shipping efficient?

Efficiency requires that the following conditions are met: (i) the elasticity of the matching function with respect to each input must equal the corresponding bargaining coefficient (thin/thick market externalities); (ii) the ship surplus must equalize across destinations (pooling externalities). We test each of these conditions in the data.

Figure 2 examines whether the thin/thick market externalities are internalized. For each region, the left panel presents the average estimated matching function elasticity with respect to exporters, η^e, as well as the estimated exporter bargaining coefficient. For several regions, as shown in the right panel, we reject that the two are equal, so that the "Hosios conditions" (29) and (30) are not satisfied. Although the "knife-edge" nature of these conditions implies that this finding is not particularly surprising, it is worth noting that the difference between the elasticity and the bargaining coefficient is often large. Moreover, the exporter bargaining coefficient tends to be lower than the matching function elasticity, suggesting that the planner would like to see an increase in the share of the surplus accruing to the exporter.

Figure 3 checks whether the pooling externalities are internalized. For each region i, it plots the coefficient of variation of the ship surplus from matching with exporters headed to different destinations $j \neq i$. When pooling externalities are internalized, this coefficient of variation should be equal to zero, since the ship is indifferent across destinations. Figure 3 demonstrates that this is not the case. In all regions the coefficient of variation is significantly different from zero, and larger than 20%, while in several regions it is substantially higher. It is worth noting that since the ship surplus only depends on the ship cost parameters and observed matching rates, this test does not rely on the Nash bargaining assumption on pricing. We conclude that the market has not internalized neither externality.



	t-stat
North America WC	4.900
North America EC	10.155
Central America	3.002
South America WC	3.497
South America EC	4.080
West Africa	1.169
Mediterranean	6.129
North Europe	7.756
South Africa	1.649
Middle East	6.936
India	8.200
South East Asia	0.685
China	1.290
Australia	1.932
Japan-Korea	2.500

Figure 2: The left panel compares the exporter bargaining coefficient γ_i^e and the average elasticity of the matching function with respect to exporters, estimated nonparametrically. The right panel presents the t-statistic for the null that the exporter bargaining coefficient γ_i^e coincides with the average elasticity of the matching function with respect to exporters.

5.2 Welfare loss

We now come to our main welfare analysis. We begin by a comparison of (i) the market equilibrium; (ii) the constrained efficient outcome we analyzed in Section 3; (iii) the frictionless equilibrium (first-best), i.e., the outcome in a world without search frictions, so that $m = \min\{s, e\}$. To compute the constrained efficient outcome, we compute the equilibrium under the efficient prices given in equation (32) of Corollary 1; in Appendix C.3 we provide the algorithm used which we found to be very robust and always delivers a unique solution.[35] In terms of policy relevance, one can think of (ii) as what can be achieved by policymakers who are not able to affect the meeting process or the search environment. In contrast, (iii) loosely corresponds to a centralized market; one can think of it as a meeting platform, like Uber, which however does not exercise market power.[36] This three-way comparison allows us to assess both the overall impact of frictions on welfare, as well as the impact of the two externalities under study.

The results are shown in Table 2. As reported in the first three columns, total welfare in the market equilibrium allocation is 6% lower than the constrained efficient allocation and 14% lower than the frictionless equilibrium. Moreover, externalities coming from search frictions have a substantial impact on

[35] Alternatively we can impose the optimal tax/subsidies derived in Corollary 2. The resulting allocation is the same.
[36] Other work has indeed modeled platforms as the eradication of search frictions; e.g. Frechette et al. (2019); Buchholz (2020).

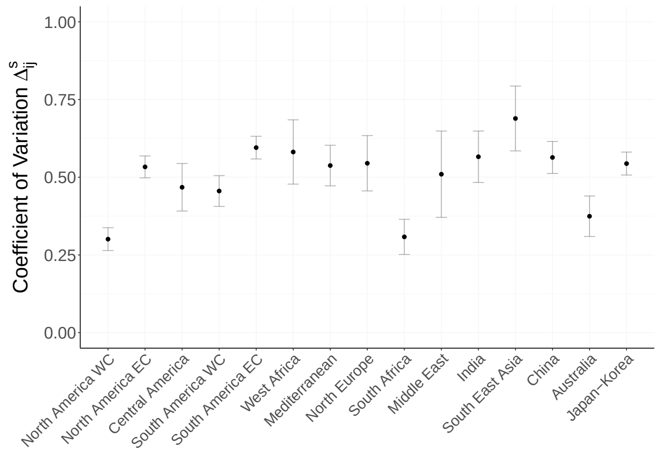


Figure 3: For each region i, we plot the coefficient of variation (standard deviation over mean) of ship surplus for all destinations $j \neq i$. When pooling externalities are internalized, the coefficient of variation should be zero.

world trade, both in terms of value and volume. Indeed, trade volume is 13% higher under constrained efficiency and 36% higher under the first-best, while net trade value (i.e. $w_{ij} - \kappa_{ij}$) is 12% higher under constrained efficiency and 43% higher under the first-best. Moreover, ships would ballast 10% and 0.6% less under constrained efficiency and the first-best respectively; this suggests that although most of ballast traveling is attributed to the natural imbalance in the endowment of commodities, wasteful traveling does exist in the market equilibrium. Finally, ships wait less under constrained efficiency and in the frictionless world (9% and 23% respectively).

These results relay an important message: under the optimal policy, the market is able to achieve about 44% of the first-best welfare gains, which, following the literature, we interpret as centralization. This is important as, in contrast to policies like taxes/subsidies, centralization may not be feasible in practice, or may come with substantial market power if provided by private firms. Indeed, platforms that reduce search frictions between agents are emerging in a multitude of markets (for instance, Uber/Lyft in the taxi market, Uber Freight and other entrants in the trucking industry, but also Airbnb in the rental housing market and peer-to-peer lending in financial markets). Yet these platforms are likely exerting market power rather than acting as benevolent planners, so that the 14% welfare gain in the first-best allocation is likely a crude upper bound on the gains from centralization. Hence, the constrained efficient allocation, which at the very least achieves almost half of overall welfare gains, may well be the desirable outcome, and it is attainable by policy.

We now discuss the different role of the two externalities. In the third and fourth columns of Table 2, we compute the welfare loss when only pooling externalities or only thin/thick market externalities are internalized. To do so, we impose the relevant tax derived in Section 3.4 (see Footnotes 24 and 25) and compute the equilibrium. The welfare gains are 3.3% when thin/thick market externalities are internalized and 5.1% when pooling externalities are internalized. This suggests that both externalities introduce substantial distortions in the market equilibrium, with pooling externalities having a larger impact.

Table 2 also reveals that the two externalities have a qualitatively different impact on the economy. The thin/thick market externalities have a large impact on the volume of trade, as they essentially distort the number of searching agents and therefore the total *number* of matches formed. Indeed, correcting the thin/thick market externalities has a bigger impact on trade volume (which rises by 19%) than correcting both externalities (13% increase). Recall that based on our estimates, the elasticity of the matching function with respect to exporters, η_i^e , is larger than the exporter bargaining coefficient in most regions (see Figure 2). Therefore, an additional exporter has a substantial positive externality on matching rates, but shipping prices are too high to appropriately encourage the entry of exporters; the planner corrects this imbalance by lowering prices and increasing exporter entry. In contrast, pooling externalities have a large impact on trade value, as they distort the composition of exports by favoring destinations with low social value. As shown in Table 2, correcting the pooling externalities leads to a large increase in trade *value* (by 13%), as destinations with high social value are subsidized.

Finally, we highlight the policy implications of pooling externalities. As discussed in Section 3, pooling externalities distort the allocation of ships over space. The table in the right panel of Figure 4 explores what types of routes are subsidized by regressing the change in prices when pooling externalities are internalized on route characteristics. First, the table reveals that routes ij with high w_{ij} are subsidized. Indeed, note that the Nash bargained price depends positively on w_{ij}, so that high value exporters pay more to transport their goods. This is because search frictions allow ships to “take advantage” of high value exporters and charge them a higher price, which in turn inefficiently reduces exporter entry of high value exporters. Thus, the planner precisely promotes trips with high w_{ij}.

	Frictionless	Constrained Efficient	Pooling	Thin/Thick
Welfare	14.32 %	6.33 %	5.14 %	3.29%
Trade	36.50 %	13.55 %	-13.62 %	19.36 %
Trade value (net)	42.71 %	11.69 %	13.61 %	6.48%
Ballast miles	-0.60 %	-9.68%	-6.43%	-12.12%
Waiting ships	-22.97 %	-9.48%	5.19%	-8.47%

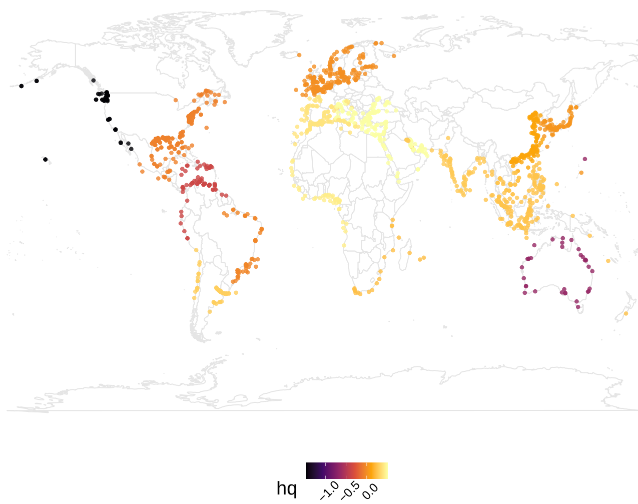
Table 2: Welfare Loss Analysis. The first column presents the frictionless allocation, i.e. the market equilibrium in the absence of search frictions when the matching function is m=min$\{s, e\}$. The second column presents the constrained efficient allocation, i.e. the market equilibrium under the efficient prices. The third and fourth columns present the market equilibrium when only the pooling and only the thin/thick market externalities respectively are internalized. All columns present the percent difference compared to the market equilibrium.

In addition, however, the efficient price for route ij also reflects the value of the ship's subsequent trips, which, as discussed above, the exporter does not fully internalize. Indeed, the planner subsidizes routes ij whose destination j is itself a high-value exporter. Naturally, the recursion continues, as the planner optimizes the entire dynamic path of ships' travels. So next, the planner also subsidizes routes ij for which location j's exports are directed to locations that themselves have high value exports, and so on. One way to capture these intricate connections between locations is to consider their centrality. As shown in the second row of the table, the planner subsidizes regions that are "central"; i.e. regions with high value import and high value export destination, and which further export to high value locations, and so on.[37]

So which locations get taxed/subsidized by the planner? The left panel of Figure 4 plots the average import tax for each region (i.e., $\sum_i q_{ij}h^q_{ij}/\sum_i q_{ij}$), while Figure 8 in Appendix D plots exporters' revenues, w_{ij}, as well as each region's average import tax. The highest subsidy is awarded to trips towards the West Coast of North America (WCNA), as well as Australia. These regions (i) are high-value importers (w_{ij} is high when j corresponds to these regions); (ii) they are also high value exporters, thus offering high continuation values to ships that arrive there; (iii) they export to regions that themselves are high-value exporters, so that ships that leave WCNA or Australia travel to locations where their options are good. To illustrate (iii), compare Australia to North Europe (NE). As shown in Figure 8 in Appendix D, the value of imports and exports are are roughly equal in the two regions. However, while NE exports to low

[37]More specifically, we consider the betweenness centrality based on the network of exporters' values, w_{ij}.

value exporters, such as India, Australia exports to Central America and WCNA, some of the world's highest-value exporters. Hence Australia is more "central" than NE and thus subsidized more heavily. In contrast, the highest taxes are levied on West Africa and India, as these are both low value importers and provide poor reloading options to ships; they also export to regions with low value exports.[38]



	$\log(\tau^p_{ij}) - \log(\tau_{ij})$
Value of exports on route ij (log)	-0.049^{**}
	(0.02)
Betweeness centrality of j (log)	-0.028^{**}
	(0.01)

Figure 4: Pooling Externalities. The left panel plots the the average import tax for each region, $\sum_i q_{ij} h^q_{ij} / \sum_i q_{ij}$. The right panel regresses the change in prices when pooling externalities are internalized (i.e. $\log\left(\tau^p_{ij}\right) - \log\left(\tau_{ij}\right)$, where τ^p_{ij} is the price that internalizes the pooling externalities only) on exporter revenue on the route, as well as region j's betweenness centrality (based on w_{ij}).

5.3 Policy implementation

Although the prices and optimal taxes that restore (constrained) efficiency have known expressions, they may not be feasible to implement in practice, either because the planner does not have access to all instruments; or because the expressions may be too complex or computationally challenging. As an example, the planner may not be able to set prices. Moreover, he may be able to tax trips, but not searching agents; indeed, it may be difficult to tax hailing passengers and searching exporters, or waiting taxis/ships. Finally, the matrix h^q may be very large, in which case the planner might prefer a simpler tax scheme.

In this section we consider simple policies which may be more easily implementable. In particular, we

[38]While imports in some regions are taxed, it is worth noting that exporters from all countries end up gaining in our results. In addition, as mentioned in Section 3.4, we restrict the planner budget to be zero at each origin, so that the taxes redistribute exporters only across destinations within each origin.

consider the following taxes: (i) an origin-specific tax on matches which can be interpreted as a flat tax on exports; (ii) a destination-specific tax on matches which can be interpreted as a customs tax; (iii) a linear in distance tax, resembling the taxi price schedule.

Table 3 reports the maximum welfare gains under these tax schemes. The destination-specific tax works best, as it achieves welfare gains of 2.8%. The origin-specific tax delivers only 0.9% welfare gains. This is consistent with our finding that pooling externalities account for a larger portion of the overall welfare loss and to resolve them it is crucial to impose different taxes across destinations. Finally, taxes that are a function of distance achieve no welfare gains. This suggests that the pricing scheme used in taxis is far from efficient and cannot alleviate either externality. This finding is not surprising given the optimal taxes derived in Section 3.4, that explicitly target the origin and the destination to correct the thin/thick and pooling externalities respectively.

Optimal tax h^q_{ij}	Export tax h^q_j	Customs tax h^q_i	Distance αd_{ij}
6.33%	2.82%	0.9%	0%

Table 3: Simple policy instruments. This table reports the maximum welfare gains attained via: an origin-specific tax on matches (second column); a destination-specific tax on matches (third column); a linear in distance tax (fourth column).

6 Conclusion

Transport markets are the engine of economic activity. Yet, little is known about their efficiency properties. In this paper we contribute to a nascent but growing literature studying optimal transport networks, by focusing on decentralized transport markets, such as taxis, trucks and bulk shipping. We show that in these markets search frictions distort the transportation network and the dynamic allocation of carriers over space and derive explicit and intuitive conditions for efficiency, which lead naturally to efficient pricing rules. We also derive optimal taxes and subsidies for a social planner that cannot set prices directly. Then, using data from dry bulk shipping, we demonstrate that search frictions lead to a sizeable social loss and substantial misallocation of ships over space. The planner subsidizes locations that are central in the trade network in order to maximize social welfare and restore the optimal allocation of ships over the globe. By doing so, the optimal policy achieves almost half of the first-best welfare gains, suggesting that it may constitute a good alternative to centralized platforms.

References

Acemoglu, D. (2001): "Good Jobs versus Bad Jobs," *Journal of Labor Economics*, 19, 1–21.

Acemoglu, D. and R. Shimer (1999): "Holdups and Efficiency with Search Frictions," *International Economic Review*, 40, 827–849.

Allen, T. (2014): "Information Frictions in Trade," *Econometrica*, 82, 2041–2083.

Allen, T. and C. Arkolakis (2019): "The Welfare Effects of Transportation Infrastructure Improvements," *NBER working paper*, 25487.

Antras, P. and A. Costinot (2011): "Intermediated Trade," *Quarterly Journal of Economics*, 126, 1319–1374.

Asplund, E. (1968): "Fréchet Differentiability of Convex Functions," *Acta Mathematica*, 121, 31–47.

Asturias, J. (2018): "Endogenous Transportation Costs," *mimeo, School of Foreign Service in Qatar, Georgetown University.*

Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012): "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain," *Econometrica*, 80, 2369–2429.

Bertsekas, D. P. (2009): *Convex Optimization Theory*, Athena Scientific Belmont.

Bian, B. (2020): "Search Frictions, Network Effects and Spatial Competition: Taxis versus Uber," *mimeo, Penn State University.*

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2018): "Geography, search frictions and endogenous trade costs," *National Bureau of Economic Research.*

——— (2020a): "Geography, Transportation, and Endogenous Trade Costs," *Econometrica*, 88, 657–691.

——— (2020b): "A Guide to Estimating Matching Functions in Spatial Models," *International Journal of Industrial Organization, (Special Issue EARIE 2018)*, 70.

Brancaccio, G., D. Li, and N. Schuerhoff (2020c): "Learning by Trading: The Case of the US Market for Municipal Bonds," *mimeo, Cornell University.*

Brooks, L., N. Gendron-Carrier, and G. Rua (2018): "The Local Impact of Containerization," *mimeo, University of Toronto.*

Buchholz, N. (2020): "Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry," *mimeo, Princeton University.*

Cao, G., G. Z. Jin, X. Weng, and L.-A. Zhou (2018): "Market Expanding or Market Stealing? Competition with Network Effects in Bike-Sharing," *NBER working paper*, 24938.

Castillo, J. C. (2019): "Who Benefits from Surge Pricing?" *mimeo, Stanford University.*

Collard-Wexler, A. (2013): "Demand Fluctuations in the Ready-Mix Concrete Industry," *Econometrica*, 81, 1003–1037.

Cosar, A. K. and B. Demir (2018): "Shipping inside the Box: Containerization and Trade," *Journal of International Economics*, 114, 331–345.

Diamond, P. A. (1982): "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 49, 217–227.

Ducruet, C., R. Juhasz, D. K. Nagy, and C. Steinwender (2019): "All Aboard: The Aggregate Effects of Port Development," *mimeo, Columbia University.*

Eaton, J., D. Jinkins, J. Tybout, and D. Xu (2016): "Two-sided Search in International Markets," *mimeo, Penn State University.*

Ericson, R. and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53–82.

Fajgelbaum, P. D. and E. Schaal (2020): "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 88.

Frechette, G. R., A. Lizzeri, and T. Salz (2019): "Frictions in a Competitive, Regulated Market Evidence from Taxis," *American Economic Review*, 109, 2954–2992.

Galichon, A. (2018): *Optimal Transport Methods in Economics*, Princeton University Press.

Gavazza, A. (2011): "The Role of Trading Frictions in Real Asset Markets," *American Economic Review*, 101, 1106–1143.

——— (2016): "An Empirical Equilibrium Model Of a Decentralized Asset Market," *Econometrica*, 84, 1755–1798.

Ghili, S. and V. Kumar (2020): "Spatial Distribution of Supply and the Role of Market Thickness: Theory and Evidence from Ride Sharing," *mimeo, Yale University.*

Holmes, T. J. and E. Singer (2018): "Indivisibilities in Distribution," *NBER working paper*, 24525.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Hosios, A. J. (1990): "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies*, 57, 279–298.

Hummels, D. and S. Skiba (2004): "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112, 1384–1402.

Imbens, G. W. and W. K. Newey (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77, 1481–1512.

Kalouptsidi, M. (2014): "Time to Build and Fluctuations in Bulk Shipping," *American Economic Review*, 104, 564–608.

——— (2018): "Detection and Impact of Industrial Subsidies, the Case of Chinese Shipbuilding," *Review of Economic Studies*, 85, 1111–1158.

Koopmans, T. C. (1949): "Optimum Utilization of the Transportation System," *Econometrica*, 17, 136–146.

Kreindler, G. (2020): "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," *Mimeo, Harvard University.*

Krolikowski, P. M. and A. H. McCallum (2018): "The Role of Trade Costs in the Surge of Trade Imbalances," *Federal Reserve Bank of Cleveland, Working Paper no. 16-35R.*

Lagos, R. (2000): "An Alternative Approach to Search Frictions," *Journal of Political Economy*, 108, 851–873.

——— (2003): "An Analysis of the Market for Taxicab Rides in New York City," *International Economic Review*, 44, 423–434.

Lee, J., S.-J. Wei, and J. Xu (2020): "The Welfare Cost of the Current Account Imbalance: A New Channel," *mimeo, Columbia University.*

Liu, T. X., Z. Wan, and C. Yang (2019): "The Efficiency of A Dynamic Decentralized Two-sided Matching Market," *mimeo, University of Rochester.*

Ma, H., F. Fang, and D. C. Parkes (2018): "Spatio-Temporal Pricing for Ridesharing Platforms," *arXiv preprint arXiv:1801.04015.*

Marimon, R. and J. Werner (2019): "The Envelope Theorem, Euler and Bellman Equations, without Differentiability," *mimeo, European University Institute.*

Matzkin, R. L. (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339–1375.

Moen, E. R. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*, 105, 385–411.

Mortensen, D. T. (1982): "The Matching Process as a Noncooperative Bargaining Game," in *The Economics of Information and Uncertainty*, ed. by J. J. McCall, University of Chicago Press.

——— (2003): *Wage Dispersion: Why Are Similar Workers Paid Differently?*, Cambridge, MA: MIT Press.

Ostrovsky, M. and M. Schwarz (2018): "Carpooling and the Economics of Self-Driving Cars," *Mimeo, Stanford University*, 25487.

Pakes, A., M. Ostrovsky, and S. T. Berry (2007): "Simple Estimators for the Parameters of Discrete Dynamics Games," *The RAND Journal of Economics*, 38, 373–399.

Panayides, P. M. (2016): *Principles of Chartering*, CreateSpace Independent Publishing Platform, 2nd ed.

Petrongolo, B. and C. A. Pissarides (2001): "Looking Into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

Pissarides, C. A. (1985): "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages," *American Economic Review*, 75, 676–690.

Plakantonaki, C., ed. (2010): *An Introduction to the World of Shipping Through the Eyes of a Ship Management Company*, Thenamaris (Ships Management) Inc.

Rosaia, N. (2020): "Competing Platforms and Transport Equilibrium: Evidence from New York City," *mimeo, Harvard University.*

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

Ryan, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 80, 1019–1062.

Shapiro, M. H. (2018): "Density of Demand and the Benefit of Uber," *mimeo, Singapore Management University.*

Stopford, M. (2009): *Maritime Economics*, London and New York: Routledge, 3rd ed.

Wong, W. F. (2018): "The Round Trip Effect: Endogenous Transport Costs and International Trade," *mimeo,University of Oregon.*

Online Appendix

A Proofs

A.1 Preliminaries: limit equilibrium outcomes and associated dual variables

In this section we show that every limit equilibrium outcome can be associated with a set of dual variables corresponding to the no-discounting limits of the agents' value functions. These variables will be instrumental in the proofs of Theorems 1 and 2 below.

Let (s, E, q, b, τ) be a limit equilibrium outcome and $(s^n, E^n, q^n, b^n, \tau^n, \beta^n)_{n \geq 0}$ be a corresponding sequence of equilibrium outcomes and discount factors, as defined in Definition 2. For each n, let $V^{s,n}, U^{s,n}, \Delta^{s,n}$ and $U^{e,n}, \Delta^{e,n}$ be the corresponding value functions and meeting surpluses for carriers and customers, respectively. Fix an arbitrary reference location i^*.

Definition 4. $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ is a tuple of equilibrium dual variables associated with the limit equilibrium outcome (s, E, q, b, τ) if there exist exists a sequence $(n_k)_{k \geq 0} \subseteq \mathbb{N}$ such that we can define the limits

$$\phi_i = \lim_{k \to \infty} V_i^{s,n_k} - V_{i^*}^{s,n_k}$$

$$\psi_i = \lim_{k \to \infty} U_i^{s,n_k} - V_{i^*}^{s,n_k}$$

$$\upsilon = \lim_{k \to \infty} \left(1 - \beta^{n_k}\right) V_{i^*}^{s,n_k}$$

$$\bar{\Delta}_{ij}^s = \lim_{k \to \infty} \Delta_{ij}^{s,n_k}.$$

If $\lambda_i^e > 0$ then we can also define the limit

$$\bar{\Delta}_{ij}^e = \lim_{k \to \infty} \Delta_{ij}^{e,n_k}$$

otherwise we simply define $\bar{\Delta}_{ij}^e = \max\left\{w_{ij}\left(q\right) - \tau_{ij} - \kappa_{ij}, 0\right\}$. We show in Lemma 2 below that these limits are indeed well-defined.

Lemma 1. *Let (s, E, q, b, τ) be a limit equilibrium outcome and $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ be a tuple of dual variables associated with it. Then the following conditions hold for every i, j:*

$$\psi_i \geq \phi_j - \frac{c_{ij}^s}{d_{ij}} - \frac{\upsilon}{d_{ij}} \textit{ with equality if } b_{ij} > 0 \tag{40}$$

$$\bar{\Delta}_{ij}^s \geq 0 \textit{ with equality if } q_{ij} < s_i \lambda_i^s G_{ij} \tag{41}$$

$$\phi_i \geq -c_i^s + \lambda_i^s \sum_{j \neq i} G_{ij} \bar{\Delta}_{ij}^s + \psi_i \textit{ with equality if } s_i > 0 \tag{42}$$

$$\upsilon \geq 0 \textit{ with equality if } \sum_{ij} \frac{q_{ij} + b_{ij}}{d_{ij}} < S \tag{43}$$

$$\bar{\Delta}_{ij}^e \geq 0 \textit{ with equality if } q_{ij} < \lambda_i^e e_{ij} \tag{44}$$

$$-c_{ij}^e + \lambda_i^e \bar{\Delta}_{ij}^e = 0 \tag{45}$$

$$\bar{\Delta}_{ij}^s \geq \tau_{ij} + \phi_j - \psi_i - \frac{\upsilon}{d_{ij}} \textit{ with equality if } q_{ij} > 0 \tag{46}$$

$$\bar{\Delta}_{ij}^e \geq w_{ij}(q) - \kappa_{ij} - \tau_{ij} \textit{ with equality if } q_{ij} > 0. \tag{47}$$

Proof. The reader can verify this by taking the no-discounting limits of the equilibrium conditions (2)-(14). For example, the equilibrium conditions (4) and (8) can be written as (taking into account that in steady state $V_{ij}^s = \left(\beta d_{ij} V_j^s - c_{ij}^s\right) / \left(1 - \beta\left(1 - d_{ij}\right)\right)$)

$$U_i^{s,n} > \frac{\beta^n d_{ij} V_j^{s,n} - c_{ij}^s}{1 - \left(1 - d_{ij}\right)\beta^n}, \text{ with equality if } b_{ij}^n > 0.$$

Subtracting $V_{i^*}^{s,n}$ from both sides we obtain,

$$U_i^{s,n} - V_{i^*}^{s,n} > \frac{\beta^n d_{ij}\left(V_j^{s,n} - V_{i^*}^{s,n}\right)}{1 - \left(1 - d_{ij}\right)\beta^n} - \frac{c_{ij}^s}{1 - \left(1 - d_{ij}\right)\beta^n} - \frac{\left(1 - \beta^n\right)V_{i^*}^{s,n}}{1 - \left(1 - d_{ij}\right)\beta^n}, \quad \text{with equality if } b_{ij}^n > 0$$

Taking limits of both sides as $n \to \infty$ yields Condition (40).

As another example, notice that the equilibrium conditions (2), (6) and (7) are equivalent to

$$\Delta_{ij}^{s,n} \geq 0, \text{ with equality if } q_{ij}^n < s_i^n \lambda_i^{s,n} G_{ij}^n$$

and

$$\Delta_{ij}^{s,n} \geq \tau_{ij}^{n} + V_{ij}^{s,n} - U_{i}^{s,n}, \text{ with equality if } q_{ij}^{n} > 0.$$

Taking the limit of the first one gives Condition (41). The second condition can be written as,

$$\Delta_{ij}^{s,n} \geq \tau_{ij}^{n} + V_{ij}^{s,n} - V_{i^*}^{s,n} - \left(U_{i}^{s,n} - V_{i^*}^{s,n}\right), \text{ with equality if } q_{ij}^{n} > 0.$$

Taking the limits of both sides gives Condition (46).

As a last example, re-write Condition (11) as

$$(1-\beta)\, U_{ij}^{e} = -c_{ij}^{e} + \lambda_{i}^{e} \Delta_{ij}^{e}.$$

From this and Conditions (10) and (14) we get that U_{ij}^{e} is bounded by $0 \leq U_{ij}^{e} \leq \kappa_{ij}$. Hence taking the limit of both sides as $\beta \to 1$ gives

$$0 = -c_{ij}^{e} + \lambda_{i}^{e} \bar{\Delta}_{ij}^{e}$$

which is Condition (45).

Analogous arguments establish the remaining conditions. More precisely, Condition (42) is a consequence of the equilibrium conditions (3) and (5); Condition (43) is a consequence of the equilibrium conditions (3) and (9); Condition (44) results from the the equilibrium conditions (10) and (12); Condition (45) is obtained from the equilibrium conditions (11) and (14); and finally, Condition (47) is obtained from the equilibrium conditions (10), (13) and (14). Finally notice that combining conditions (46) and (47) we obtain,

$$\bar{\Delta}_{ij}^{s} + \bar{\Delta}_{ij}^{e} \geq w_{ij}\,(q) - \kappa_{ij} + \phi_{j} - \psi_{i} - \frac{\upsilon}{d_{ij}} \text{ with equality if } q_{ij} > 0. \tag{48}$$

□

Finally, we prove that the limits in Definition 4 are indeed well-defined.

Lemma 2. *The dual variables* $\left(\phi, \psi, \upsilon, \bar{\Delta}^{s}, \bar{\Delta}^{e}\right)$ *always exist.*

Proof. To prove the lemma we show that the sequences $(V_{i}^{s,n} - V_{i^*}^{s,n})_{n\geq 0}$, $(U_{i}^{s,n} - V_{i^*}^{s,n})_{n\geq 0}$, $((1-\beta^{n})\, V_{i^*}^{s,n})_{n\geq 0}$ and $\left(\Delta_{ij}^{s,n}\right)_{n\geq 0}$ are bounded for every i and j, while $\left(\Delta_{ij}^{e,n}\right)_{n\geq 0}$ is also bounded provided that $\lambda_{i}^{e} > 0$.

Taking into account equation (4) we can rewrite the carrier's surplus, $\Delta_{ij}^{s,n}$, defined in equation (2) as $\Delta_{ij}^{s,n} = \max\left\{\tau_{ij}^{n} + V_{ij}^{s,n} - \max_j V_{ij}^{s,n}, 0\right\}$. Hence $\Delta_{ij}^{s,n}$ is bounded above by τ_{ij}^{n} and below by zero. Since τ^n converges to τ, it follows that $\Delta_{ij}^{s,n}$ is bounded. Note that in the steady state, equation (1) becomes: $V_{ij}^{s} = \left(-c_{ij}^{s} + \beta d_{ij} V_j^{s}\right) / \left(1 - \beta\left(1 - d_{ij}\right)\right)$. $\left(1-\beta^n\right) V_{i^*}^{s,n}$ is bounded as an average of bounded prices and the finite set of all possible per-period search and traveling costs.[39] $V_i^{s,n} - V_{i^*}^{s,n}$ is bounded below, since we have

$$
\begin{aligned}
V_i^{s,n} - V_{i^*}^{s,n} &\geq -c_i^s + \lambda_i^{s,n} \sum_{j \neq i} G_{ij}^n \Delta_{ij}^{s,n} + \frac{\beta^n d_{ii^*} V_{i^*}^{s,n} - c_{ii^*}^s}{1 - \beta^n\left(1 - d_{ii^*}\right)} - V_{i^*}^{s,n} \\
&= -c_i^s + \lambda_i^{s,n} \sum_{j \neq i} G_{ij}^n \Delta_{ij}^{s,n} - \frac{c_{ii^*}^s}{1 - \beta^n\left(1 - d_{ii^*}\right)} - \frac{\left(1-\beta^n\right) V_{i^*}^{s,n}}{1 - \beta^n\left(1 - d_{ii^*}\right)}
\end{aligned}
$$

and all sequences on the right-hand-side are bounded. Reversing the roles of i and i^* it follows that $V_i^{s,n} - V_{i^*}^{s,n}$ is bounded above as well.

Finally, if $\lambda_i^e > 0$, then $\lambda_i^{e,n} > 0$ for n large enough. Hence, based on Equations (10) and (11), for n large enough, we have,

$$
\frac{-c_{ij}^e + \lambda_i^{e,n}\left(w_{ij}\left(q^n\right) - \tau_{ij}^n\right)}{1 - \left(1 - \lambda_i^{e,n}\right)\beta^n} \leq U_{ij}^{e,n} \leq \kappa_{ij}.
$$

Since the left-hand-side converges, $U_{ij}^{e,n}$ is bounded. By equation (10), this implies that $\Delta_{ij}^{e,n}$ is bounded as well. Finally note that for every i it holds that

$$
\lim_{k\to\infty}\left(1-\beta^{n_k}\right) V_i^{s,n_k} = \underbrace{\lim_{k\to\infty}\left(1-\beta^{n_k}\right)\left(V_i^{s,n_k} - V_{i^*}^{s,n_k}\right)}_{=0} + \lim_{k\to\infty}\left(1-\beta^{n_k}\right) V_{i^*}^{s,n_k} = \upsilon.
$$

[39]In particular, from equation (3), we have that $-c_i^s + \lambda_i^{s,n}\sum_j G_{ij}^n \Delta_{ij}^{s,n} + V_{ij}^{s,n} \leq V_i^{s,n}$ for all ij, so that

$$
-c_i^s + \lambda_i^{s,n} \sum_j G_{ij}^n \Delta_{ij}^{s,n} + \frac{-c_{ij}^s + d_{ij}\beta^n V_j^{s,n}}{1 - \beta^n\left(1 - d_{ij}\right)} - V_i^{s,n} \leq 0 \tag{49}
$$

Let $k_{ij}^n = -c_i^s + \lambda_i^{s,n}\sum_j G_{ij}^n \Delta_{ij}^{s,n} + \frac{-c_{ij}^s}{1-\beta^n\left(1-d_{ij}\right)}$. This is a bounded sequence. Hence (49) is written as $\frac{d_{ij}\beta^n V_j^{s,n}}{1-\beta^n\left(1-d_{ij}\right)} \leq V_i^{s,n} - k_{ij}^n$ for all $i \neq j$, or $V_j^{s,n} \leq \frac{1-\beta^n\left(1-d_{ij}\right)}{d_{ij}\beta^n} V_i^{s,n} - \frac{1-\beta^n\left(1-d_{ij}\right)}{d_{ij}\beta^n} k_{ij}^n$. Applying the same inequality on V_i^s and rearranging, we obtain

$$
\left(1 - \frac{\left(1-\beta^n+\beta^n d_{ij}\right)\left(1-\beta^n+\beta^n d_{ji}\right)}{d_{ji} d_{ij} \beta^n}\right) V_j^{s,n} \leq -\frac{\left(1-\beta^n+\beta^n d_{ij}\right)}{d_{ij}\beta^n}\frac{\left(1-\beta^n+\beta^n d_{ji}\right)}{d_{ji}\beta^n} k_{ji}^n - \frac{1-\beta^n-\beta^n d_{ij}}{d_{ij}\beta^n} k_{ij}^n
$$

It is easy to see that the right-hand-side is bounded. Moreover, the left-hand-side, after straightforward computations, becomes $\frac{1-\left(1-\beta^n\right)-\beta^n\left(d_{ij}+d_{ji}\right)}{d_{ij}d_{ji}\beta^n}\left(1-\beta^n\right)V_j^{s,n}$. But $\lim_{\beta\to1}\frac{1-\left(1-\beta^n\right)-\beta^n\left(d_{ij}+d_{ji}\right)}{d_{ij}d_{ji}\beta^n} = \frac{\left(d_{ij}+d_{ji}\right)}{d_{ij}d_{ji}}$ and hence $\left(1-\beta^n\right)V_j^{s,n}$ is bounded.

□

A.2 Proof of Theorem 1

Consider Problem (21), and let $\phi, \psi, \upsilon, \bar{\Delta}^s$ and $\bar{\Delta}^e$ be the dual variables associated with constraints (15), (16), (17), (22) and (23), respectively. Since Problem (21) is concave, convex duality implies that $\left(s, E, q, b, \phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ is an optimal dual pair of Problem (21) (that is, (s, E, q, b) is an optimal solution of Problem (21) and $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ are the multipliers associated with the constraints) if and only if it satisfies the Karush-Kuhn-Tucker conditions (see for example Bertsekas, 2009, Prop. 5.2.2, pg. 167). The reader can verify that these can be written as conditions (40)-(45), as well as condition (48). Hence the result follows by taking $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ to be a tuple of equilibrium dual variables associated with s, E, q, b and applying Lemma 1.

A.3 Proof of Theorem 2

First, we show that V^p is concave and characterize its supergradient. We then express the first order conditions of the three optimization problems defining the externalities, (26)-(28), in terms of the supergradient. Finally, we show that a suitable combination of the dual variables associated with the market allocation lies in the supergradient to prove the main result.

Concavity of V^p. Let $u = (s, e, G)$ with $\sum_i s_i \leq S$ and $\sum_j G_{ij} = 1$ for all i. The set $M(u)$ of all pairs $(q, b) \geq 0$ satisfying constraints (15)-(17) given s, e, G has the following convexity property

$$(1 - \lambda) M(u) + \lambda M(u') \subseteq M\left((1 - \lambda) u + \lambda u'\right)$$

for every u, u'. This follows readily from the concavity of the matching function.

Letting $x = (q, b)$, the definition of V^p, equation (24), becomes

$$V^p(u) = \max_{u \in M(u)} f(u, x)$$

where $f(u, x) = f(s, e, G, q, b) = W(q) - \sum_{ij} q_{ij}\kappa_{ij} - \sum_{ij}(q_{ij} + b_{ij})\frac{c^s_{ij}}{d_{ij}} - \sum_i s_i c^s_i - \sum_i e_i \sum_j G_{ij} c^e_{ij}$. It is obvious that $f(u,x)$ is jointly concave in (u,x). To establish concavity of $V^p(u)$ pick $u_1, u_2,\ \lambda \in [0,1]$ and $\epsilon > 0$. Then there exist $x_1 \in M(u_1)$ and $x_2 \in M(u_2)$ such that $f(u_1, x_1) \leq g(u_1) + \epsilon$ and $f(u_2, x_2) \leq g(u_2) + \epsilon$. Then,

$$g(\lambda u_1 + (1-\lambda) u_2) = \inf_{x \in M(\lambda u_1 + (1-\lambda) u_2)} f(\lambda u_1 + (1-\lambda) u_2, x)$$

Since $\lambda x_1 + (1-\lambda) x_2 \in \lambda M(u_1) + (1-\lambda) M(u_2) \subseteq M(\lambda u_1 + (1-\lambda) u_2)$, we have,

$$\inf_{x \in M(\lambda u_1 + (1-\lambda) u_2)} f(\lambda u_1 + (1-\lambda) u_2, x) \leq f(\lambda u_1 + (1-\lambda) u_2, \lambda x_1 + (1-\lambda) x_2)$$

Since $f(\cdot)$ is concave in (x,u) we have,

$$\begin{aligned} g(\lambda u_1 + (1-\lambda) u_2) &\leq f(\lambda u_1 + (1-\lambda) u_2, \lambda x_1 + (1-\lambda) x_2) \\ &\leq \lambda f(u_1, x_1) + (1-\lambda) f(u_2, x_2) \\ &\leq \lambda g(u_1) + (1-\lambda) g(u_2) + \epsilon \end{aligned}$$

Since this is true for all ϵ, concavity is established.

Externalities and supergradient of V^p. We now relate the definitions of externalities, equations (26)-(28), with the first order conditions expressed in terms of the supergradient of V^p.

The supergradient of V^p, $\partial V^p(s,e,G)$, at a triplet s,e,G, is the set of all vectors

$$y = (y(s_i)_{i\in I}, y(e_i)_{i \in I}, y(G_{ij}))_{i,j \in I} \in \mathbb{R}^I \times \mathbb{R}^I \times \mathbb{R}^{I\times I}$$

such that for every triplet s', e', G':

$$V^p(s', e', G') - V^p(s,e,G) \leq \sum_i y(s_i)(s'_i - s_i) + \sum_i y(e_i)(e'_i - e_i) + \sum_{ij} y(G_{ij}) \left(G'_{ij} - G_{ij}\right).$$

Similarly, for every i,j, we denote by $\partial_{s_i} V^p(s,e,G)$, $\partial_{e_i} V^p(s,e,G)$ and $\partial_{G_{ij}} V^p(s,e,G)$ the supergradients of V^p at s,e,G with respect to s_i, e_i and G_{ij}, respectively.[40]

[40] Clarifying the notation, we use the letter y to denote vectors belonging to the space $\mathbb{R}^I \times \mathbb{R}^I \times \mathbb{R}^{I\times I}$, while $\partial_{s_i} V^p(s,e,G)$, $\partial_{e_i} V^p(s,e,G)$ and $\partial_{G_{ij}} V^p(s,e,G)$ are vectors belonging to $\mathbb{R}^I$, $\mathbb{R}^I$ and $\mathbb{R}^{I\times I}$, respectively.

Recall that we are assuming that s and e are in the interior of the feasible set ($s_i, e_i > 0$ for each i and $\sum_i s_i < S$). Hence conditions (26) and (27), which correspond to the thin/thick market externalities, are equivalent to the first order conditions,

$$0 \in \partial_{s_i} V^p (s, e, G) \text{ and, } 0 \in \partial_{e_i} V^p (s, e, G) \ \forall i \tag{50}$$

respectively. Denoting by ν_{ij} and μ_i the multipliers associated with the constraints $G_{ij} \geq 0$ and $\sum_j G_{ij} = 1$, condition (28), which corresponds to the pooling externalities, is equivalent to

$$0 - \nu_{ij} - \mu_i \in \partial_{G_{ij}} V^p (s, e, G) \tag{51}$$

for some $\mu, \nu \in \mathbb{R}^I \times \mathbb{R}_+^{I \times I}$ such that $\nu_{ij} G_{ij} = 0$.

Dual variables and supergradient of V^p We finally prove that the dual variables associated with a limit equilibrium allocation readily generate elements in the supergradient of V^p which are used to express the first order conditions, (50) and (51), and establish the proof of Theorem 2.

Consider a limit equilibrium allocation (s, e, G, q, b), and let $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ be a tuple of equilibrium dual variables associated with it.[41] The vector y with entries

$$y (s_i) = -\phi_i - c_i^s + \frac{dm_i (s_i, e_i)}{ds_i} \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e\right) + \psi_i$$

$$y (e_i) = -c_{ij}^e + \frac{dm_i (s_i, e_i)}{de_i} \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e\right)$$

$$y (G_{ij}) = -e_i c_{ij}^e + m_i (s_i, e_i) \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e\right).$$

belongs to $\partial V^p (s, e, G)$, so that $y \in \partial V^p (s, e, G)$. Indeed, consider the Lagrangian associated with

[41] That is, $\left(\phi, \psi, \upsilon, \bar{\Delta}^s, \bar{\Delta}^e\right)$ satisfies the conditions of Lemma 1. The ordering in the notation does not imply a duality relationship between particular pairs of elements. For example, it is not correct to say that "ϕ is the dual of s".

Problem (24),

$$L\left(q',b',\psi',\phi',\bar{\Delta}',\upsilon'|s,e,G\right) = W\left(q'\right) + \sum_{ij}\left(q'_{ij}+b'_{ij}\right)\left(-\frac{c^s_{ij}}{d_{ij}}+\phi'_j-\psi'_i-\frac{\upsilon}{d_{ij}}\right) - \sum_{ij} q'_{ij}\left(\bar{\Delta}'_{ij}+\kappa_{ij}\right)$$
$$-\sum_i e_i \sum_j G_{ij} c^e_{ij} - \sum_i s_i\left(\phi'_i-\psi'_i+c^s_i\right) + \sum_i m_i\left(s_i,e_i\right)\sum_j G_{ij}\bar{\Delta}'_{ij} + S\upsilon'$$

and the Karush-Kuhn-Tucker (K-K-T) conditions as

$$\psi_i \geq \phi_j - \frac{c^s_{ij}}{d_{ij}} - \frac{\upsilon}{d_{ij}} \text{ with equality if } b_{ij} > 0$$

$$\bar{\Delta}_{ij} \geq 0 \text{ with equality if } q_{ij} < m_i\left(s_i,e_i\right)G_{ij}$$

$$\bar{\Delta}_{ij} \geq \phi_j - \frac{c^s_{ij}}{d_{ij}} - \kappa_{ij} - \psi_i - \frac{\upsilon}{d_{ij}} \text{ with equality if } q_{ij} > 0$$

$$\upsilon \geq 0 \text{ with equality if } \sum_{ij}\frac{q_{ij}+b_{ij}}{d_{ij}} < S$$

which are equivalent to the set of Conditions (40), (41)/(44), (48) and (43), respectively, taking $\bar{\Delta}_{ij} = \bar{\Delta}^s_{ij} + \bar{\Delta}^e_{ij}$. Since the problem is concave, the K-K-T conditions are necessary and sufficient for optimality. Hence letting $\left(\phi,\psi,\upsilon,\bar{\Delta}^s,\bar{\Delta}^e\right)$ be a tuple of equilibrium dual variables associated with (s,e,G,q,b), it follows that $\left(q,b,\psi,\phi,\bar{\Delta}^s+\bar{\Delta}^e,\upsilon\right)$ is an optimal dual pair for Problem (24). From the assumptions of Theorem 2 , it follows that (q,b) is the unique optimal solution of Problem (24). Hence the result follows from Theorem 2 of Marimon and Werner (2019).

We use the vector y given above to express the first order conditions (50) and (51) as follows,

$$\forall i : y\left(s_i\right) = 0 \tag{52}$$

$$\forall i : y\left(e_i\right) = 0 \tag{53}$$

$$\forall i,j : y\left(G_{ij}\right) + \mu_i \leq 0 \text{ with equality if } G_{ij} > 0 \tag{54}$$

Note that these conditions are sufficient for conditions (26), (27) and (28), respectively, and they are necessary whenever $V^p\left(s,e,G\right)$ is differentiable. Comparing with condition (42), condition (52) is equivalent to

$$\forall i: -c_i^s + \frac{dm_i(s_i, e_i)}{ds_i} \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e \right) + \psi_i = \phi_i = -c_i^s + \lambda_i^s \sum_j G_{ij} \bar{\Delta}_{ij}^s + \psi_i$$
$$\Leftrightarrow \eta_i^s \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e \right) = \sum_j G_{ij} \bar{\Delta}_{ij}^s,$$

Similarly, comparing with condition (45), condition (53) is equivalent to

$$\forall i: -c_{ij}^e + \frac{dm_i(s_i, e_i)}{de_i} \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e \right) = 0 = -c_{ij}^e + \lambda_i^e \sum_j G_{ij} \bar{\Delta}_{ij}^e$$
$$\Leftrightarrow \sum_j G_{ij} \bar{\Delta}_{ij}^e = \eta_i^e \sum_j G_{ij} \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e \right).$$

Condition (54) requires

$$-e_i c_{ij}^e + m_i(s_i, e_i) \left(\bar{\Delta}_{ij}^s + \bar{\Delta}_{ij}^e \right) + \mu_i \leq 0 \text{ with equality if } G_{ij} > 0. \tag{55}$$

Since $-e_i c_{ij}^e + m_i(s_i, e_i) \bar{\Delta}_{ij}^e = 0$ from Condition (45), this is equivalent to Condition (iii) in the Theorem 2. This completes the proof of Theorem 2.

A.4 Proof of Corollary 1

Suppose that (s, e, G, q, b, τ) is efficient. Conditions (i) and (ii) of Theorem 2 imply that $\eta_i^s = 1 - \eta_i^e$ for all i. For every ij such that $G_{ij} > 0$, Conditions (i) and (iii) of Theorem 2 imply $\bar{\Delta}_{ij}^s = (1 - \eta_i^e) \sum_j G_{ij} \bar{\Delta}_{ij}$. By Condition (47) we have $\bar{\Delta}_{ij}^e = w_{ij}(q) - \kappa_{ij} - \tau_{ij}$. Substituting $\bar{\Delta}_{ij}^s = \bar{\Delta}_{ij} - \bar{\Delta}_{ij}^e = \bar{\Delta}_{ij} - w_{ij}(q) + \kappa_{ij} + \tau_{ij}$ yields Condition (32).

A.5 Efficiency Without Constant Returns to Scale

In the discussion following Theorem 2 we noted that, unless all matching functions display constant returns to scale, efficiency cannot be achieved. In this section we allow the planner to charge a price to customers, τ^e, that is different from the price paid to carriers, τ^s. The price wedge $\tau^e - \tau^s$ can be interpreted as a tax/subsidy. We show that in this setting the planner can achieve efficiency even when the matching functions do not exhibit constant returns to scale.

The definition of equilibrium can be extended to this case in a straightforward manner: $(s, e, G, q, b, \tau^e, \tau^s)$ is an equilibrium outcome, if carriers behave optimally given τ^s, λ^s and G; customers behave optimally given τ^e and λ^e; the feasibility constraints are satisfied; and λ^s, λ^e and G are consistent with the allocation.

The social planner's problem is the one stated in (25). Notice that the tax revenues do not appear in the social welfare since they are a transfer from the agents to the planner (one can imagine that the tax revenues are paid back to agents by means of a lump sum transfer). On the other hand, the limit social surplus resulting from a match now takes the planner's revenue into account:

$$\bar{\Delta}_{ij} = \bar{\Delta}^s_{ij} + \bar{\Delta}^e_{ij} + \tau^e_{ij} - \tau^s_{ij}.$$

With these modifications, we can proceed along the lines of the proof of Theorem 2 to show the following:

Theorem 3. *Let $(s, e, G, q, b, \tau^e, \tau^s)$ be a limit equilibrium outcome. Suppose that Problem (24) admits a unique optimal solution. Then:*

(i) Carriers internalize thin/thick market externalities if and only if

$$\forall i \in I: \frac{\sum_j G_{ij} \bar{\Delta}^s_{ij}}{\sum_j G_{ij} \bar{\Delta}_{ij}} = \eta^s_i. \tag{56}$$

(ii) Customers internalize thin/thick market externalities if and only if

$$\forall i \in I: \frac{\sum_j G_{ij} \bar{\Delta}^e_{ij}}{\sum_j G_{ij} \bar{\Delta}_{ij}} = \eta^e_i. \tag{57}$$

(iii) Customers internalize pooling externalities if and only if

$$\bar{\Delta}^s_{ij} + \tau^e_{ij} - \tau^s_{ij} = \max_{k \neq i} \left(\bar{\Delta}^s_{ik} + \tau^e_{ik} - \tau^s_{ik} \right) \tag{58}$$

for all ij such that $G_{ij} > 0$.

Hence, the characterization of efficiency in the economy with a price wedge is the same as the one in the main text except for the last condition, which requires that, at each location i, the sum of the carrier surplus and the planner's revenue is constant across destinations– in other words, customers must fully internalize the differences in the matching surpluses across different destinations. We can then use this result to characterize the optimal pricing rules. Let $(s, e, G, q, b, \tau^e, \tau^s)$ be an efficient limit equilibrium outcome and suppose that (s, e, G, q, b) is efficient. For simplicity, consider the case where $G_{ij} > 0$ for all ij. Condition (57) can be written as

$$\sum_j G_{ij} \left(\bar{\Delta}^s_{ij} + \tau^e_{ij} - \tau^s_{ij} \right) = \sum_j G_{ij} \left(\bar{\Delta}_{ij} - \bar{\Delta}^e_{ij} \right) = (1 - \eta^e_i) \sum_j G_{ij} \bar{\Delta}_{ij}.$$

Hence Condition (58) implies that

$$\bar{\Delta}_{ij} - \bar{\Delta}^e_{ij} = \bar{\Delta}^s_{ij} + \tau^e_{ij} - \tau^s_{ij} = (1 - \eta^e_i) \sum_j G_{ij} \bar{\Delta}_{ij}$$

for every ij. Substituting $\bar{\Delta}^e_{ij} = w_{ij}(q) - \tau^e_{ij} - \kappa_{ij}$ into this equation we find

$$\tau^e_{ij} = w_{ij}(q) - \kappa_{ij} - \bar{\Delta}_{ij} + (1 - \eta^e_i) \sum_j G_{ij} \bar{\Delta}_{ij} \tag{59}$$

which is the pricing rule in equation (32). The average price wedge can be derived by summing conditions (56) and (57), and recalling that $\tau^e_{ij} - \tau^s_{ij} = \bar{\Delta}_{ij} - \bar{\Delta}^e_{ij} - \bar{\Delta}^s_{ij}$:

$$\sum_j G_{ij} \left(\tau^e_{ij} - \tau^s_{ij} \right) = (1 - \eta^e_i - \eta^s_i) \sum_j G_{ij} \bar{\Delta}_{ij}. \tag{60}$$

This expression can be interpreted as saying that the average price wedge at each location is proportional to the "degree of decreasing returns to scale". Under constant returns to scale the wedge is zero: consistently with our main results, efficiency in this case can be achieved by setting a unique price on every route. If the matching function has decreasing returns to scale then the price wedge is positive, imposing a tax on matches at that location, capturing the social cost of making additional matches harder to form because of decreasing returns. On the contrary, matches are subsidized when the matching functions have increasing returns.

Conversely, it is easy to see that equations (59) and (60) imply equations (56)-(58).

We state the conclusions of this section below:

Corollary 3. *Let* $(s, e, G, q, b, \tau^e, \tau^s)$ *be a limit equilibrium outcome. Then* s, e, G, q, b *is efficient if and only if for all* i, j

$$\tau_{ij}^e = w_{ij}(q) - \kappa_{ij} - \bar{\Delta}_{ij} + (1 - \eta_i^e) \sum_j G_{ij} \bar{\Delta}_{ij}$$

and

$$\sum_j G_{ij} \left(\tau_{ij}^e - \tau_{ij}^s\right) = (1 - \eta_i^e - \eta_i^s) \sum_j G_{ij} \bar{\Delta}_{ij}$$

A.6 Proof of Corollary 2

Before proceeding with the proof, we briefly describe how incentives and total welfare are affected by a vector of taxes/subsidies $h = (h^q, h^s, h^e)$. The dynamic problem for customers is the same as in Section 2.2 except that now customers waiting at location i pay the amount h_i^e every period (on top of the private waiting cost c_{ij}^e); carriers searching at location i pay the amount h_i^s every period (on top of their private waiting cost c_i^s); and every match on route ij is taxed by the amount h_{ij}^q. It does not matter which side pays the trip tax (customers or carriers), so suppose that it is paid by customers (see footnote 22). Therefore, the only expressions that change compared to Section 2.2 are the carriers' value of searching:

$$V_i^s = \max \left\{ -c_i^s - h_i^s + \lambda_i^s \sum_{j \neq i} G_{ij} \Delta_{ij}^s + U_i^s, \beta V_i^s \right\}$$

and the customers' value of waiting and meeting surplus:

$$U_{ij}^e = -c_{ij}^e - h_i^e + \lambda_i^e \Delta_{ij}^e + \beta U_{ij}^e$$

$$\Delta_{ij}^e = \max \left\{ w_{ij}(q) - \tau_{ij} - h_{ij}^q - \beta U_{ij}^e, 0 \right\}.$$

We can extend the definition of equilibrium to accommodate Nash bargaining and taxes in a straightforward manner: (s, e, G, q, b, τ) is an equilibrium outcome under taxes h and Nash bargaining, if carriers and customers behave optimally given h, τ, λ^s, λ^e and G; the feasibility constraints are satisfied; λ^s, λ^e and G are consistent with the allocation; and finally, prices are determined by the usual surplus sharing condition (34).

The social planner's problem is the one stated in (25). Notice that the tax revenues do not appear in the social welfare since they are a transfer from the agents to the social planner (one can imagine that the tax revenues are paid back to agents by means of a lump sum transfer). On the other hand, the limit social surplus resulting from a match now takes into account the planner's revenue as well:

$$\bar{\Delta}_{ij} = \bar{\Delta}^s_{ij} + \bar{\Delta}^e_{ij} + h^q_{ij}.$$

Now for the proof of Corollary 2, let q, s, e, G be a limit equilibrium allocation in the economy with taxes. Proceeding as in the proof of Theorem 2, one can show that[42]

- Thin/thick market externalities are internalized if and only if for all i,

$$\sum_j G_{ij}\bar{\Delta}^s_{ij} = \eta^s_i \sum_j G_{ij}\bar{\Delta}_{ij} + \frac{h^s_i}{\lambda^s_i} \tag{61}$$

  and

$$\sum_j G_{ij}\bar{\Delta}^e_{ij} = \eta^e_i \sum_j G_{ij}\bar{\Delta}_{ij} + \frac{h^e_i}{\lambda^e_i}$$

- Pooling externalities are internalized if and only if for all i, j,

$$\bar{\Delta}^s_{ij} + h^q_{ij} \leq L_i \quad \text{with equality if } G_{ij} > 0$$

[42] In particular, we proceed by showing that the limits of the equilibrium conditions are a slight modification of conditions (40)-(47) in Section A.1

$$w_{ij}(q) - \frac{c^s_{ij}}{d_{ij}} - \kappa_{ij} + \phi_j - \psi_i - h^q_{ij} - \bar{\Delta}^e_{ij} - \bar{\Delta}^s_{ij} - \frac{v}{d_{ij}} \leq 0 \text{ with equality if } q_{ij} > 0$$

$$-\frac{c^s_{ij}}{d_{ij}} + \phi_j - \psi_i - \frac{v}{d_{ij}} \leq 0 \text{ with equality if } b_{ij} > 0$$

$$\bar{\Delta}^e_{ij}, \bar{\Delta}^s_{ij} \geq 0 \text{ with equality if } q_{ij} \leq m_i(s_i, e_i)\, G_{ij}$$

$$v \geq 0 \text{ with equality if } \sum \frac{q_{ij} + b_{ij}}{d_{ij}} < S$$

$$\phi_i = -c^s_i - h^s_i + \psi_i + \lambda^s_i \sum_j G_{ij}\bar{\Delta}^s_{ij}$$

$$-c^e_{ij} - h^e_i + \lambda^e_i \bar{\Delta}^e_{ij} \leq 0$$

$$w_{ij}(q) - \tau^e_{ij} - \kappa_{ij} - \bar{\Delta}^e_{ij} \leq 0 \text{ with equality if } q_{ij} > 0$$

Moreover, Nash bargaining requires that $(1-\gamma_i)\,\bar{\Delta}^s_{ij} = \gamma_i \bar{\Delta}^e_{ij}$. On the other hand, the conditions for different externalities to be internalized are unchanged. Comparing the efficiency conditions with the new equilibrium conditions as in the proof of Theorem 2 we get the result shown below.

where L_i is an arbitrary constant.

Using the definition $\bar{\Delta}_{ij} = \bar{\Delta}^e_{ij} + \bar{\Delta}^s_{ij} + h^q_{ij}$ and the Nash bargaining condition $(1-\gamma_i)\,\bar{\Delta}^s_{ij} = \gamma_i \bar{\Delta}^e_{ij}$ it follows that $\bar{\Delta}_{ij} = \frac{1}{\gamma_i}\bar{\Delta}^s_{ij} + h^q_{ij}$ or $\bar{\Delta}^s_{ij} = \gamma_i \bar{\Delta}_{ij} - \gamma_i h^q_{ij}$. Substituting $\bar{\Delta}^s_{ij}$ into (61) we obtain the Condition (35). We proceed similarly for customers to obtain Condition (36).

Next, we turn to the relationship $\bar{\Delta}^s_{ij} + h^q_{ij} \leq L_i$ with equality if $G_{ij} > 0$. The constant L_i is related to the Lagrange multiplier associated with the constraint $\sum_j G_{ij} = 1$. Consider all j such that $G_{ij} > 0$. Then $\bar{\Delta}^s_{ij} + h^q_{ij} = L_i$. Multiply by G_{ij} and sum over j to obtain, $L_i = \sum_j G_{ij} \bar{\Delta}^s{}_{ij} + \sum_j G_{ij} h^q_{ij}$. Clearly the sums can be extended to all j since the terms with $G_{ij} = 0$ do not contribute to the sum. Thus pooling externalities are internalized if and only if

$$\bar{\Delta}^s_{ij} + h^q_{ij} \leq \sum_j G_{ij} \bar{\Delta}^s{}_{ij} + \sum_j G_{ij} h^q_{ij}$$

with equality if $G_{ij} > 0$. We now express $\bar{\Delta}^s_{ij}$ in terms of $\bar{\Delta}_{ij}$ using the surplus sharing condition which yields $\bar{\Delta}^s_{ij} = \gamma_i \left(\bar{\Delta}_{ij} - h^q_{ij}\right)$:

$$\gamma_i \left(\bar{\Delta}_{ij} - h^q_{ij}\right) + h^q_{ij} \leq \gamma_i \sum_j G_{ij} \left(\bar{\Delta}_{ij} - h^q_{ij}\right) + \sum_j G_{ij} h^q_{ij}$$

$$\gamma_i \bar{\Delta}_{ij} + (1-\gamma_i)\, h^q_{ij} \leq \gamma_i \sum_j G_{ij} \bar{\Delta}_{ij} + (1-\gamma_i) \sum_j G_{ij} h^q_{ij}$$

which proves (37).

B Homogeneity and random search in bulk shipping

B.1 Homogeneity

In this section we present additional descriptive evidence supporting that bulk ships are homogeneous. We first explore shipping prices and then ship ballast choices. Table 4 regresses shipping prices on ship and shipowner characteristics and fixed effects; we demonstrate that none of them are significant.

Next, it is worth noting that the fit of the ballast discrete choice model is very good (see BKP); this already suggests that the factors capturing a region's attractiveness considered in our model, can predict

	log(price per day)			
	I	II	III	IV
$I\{\text{orig.} = \text{home country}\}$			0.004	
			(0.019)	
$I\{\text{dest.} = \text{home country}\}$			−0.012	
			(0.015)	
ln (Number Employees)				0.008
				(0.007)
ln (Operating Revenues)				0.003
				(0.005)
Time FE	Qtr×Yr	Qtr×Yr	Qtr×Yr	Qtr×Yr
Shipowner FE	No	Yes	No	No
Ship characteristics	Yes	Yes	Yes	Yes
Region FE	Orig. & Dest.	Orig. & Dest.	Orig. & Dest.	Orig. & Dest.
Observations	7,263	7,263	7,973	7,973
Adj. R^2	0.530	0.540	0.537	0.537

*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression of shipping prices on shipowner characteristcs and fixed effects (Table SI in Supplement to BKP). Shipping prices, ships' characteristics (age and size), and the identity of the shipowner are obtained from Clarksons. Information on shipowner characteristics is obtained from ORBIS. In particular, we match the shipowners in Clarksons to ORBIS; we do so for two reasons: (i) ORBIS allows us to have reliable firm identities, as shipowners may appear under different names in the contract data; (ii) ORBIS reports additional firm characteristics (e.g. number of employees, revenue, headquarters). Here we identify the shipowner with the global ultimate owner (GUO); results are robust to controlling for the identity of the domestic owner (DUO) and the shipowner as reported in Clarksons. Finally, the data used span the period 2010-2016.

behavior. Moreover, as mentioned in Footnote 32, ships' ballast choices are not dispersed from a given origin (i.e. ships tend to ballast to the same regions). To explore this more formally, in Table 5 we test whether ships' ballast destination decisions are heterogeneous across ships. To do so, consider origin i, and denote by $j(i)$ the most common destination for ballast trips originating from i. For any ship traveling empty from i, we denote n_i the number of the ship's ballast trips starting in i, $(j_t)_{t=1}^{n_i}$ the destination for each one of these trips, and we define the variable y_t to be equal to one if in trip t the ship heads to $j(i)$ and zero otherwise (i.e. $y_t = \mathbb{I}\{j_t = j(i)\}$). Under the assumption that there is no heterogeneity, the variable y_t follows a binomial distribution with the same probability for every ship; in particular $y_t \sim Bin\left(\mathbb{P}_{i,j(i)}, n_i\right)$, where $\mathbb{P}_{i,j(i)}$ is the overall share of ballast trips in i that head to destination $j(i)$. We test this assumption for every ship focusing on the most frequent origin for each ship. Table 5 reports summary statistics for the distribution of p-values across ships and shows that the test fails to reject the null for the overwhelming majority of ships.

	p-value				N
	Mean	Median	p25	p75	
All ships	0.508	0.501	0.177	0.826	5,180
$n_i > 2$	0.495	0.463	0.167	0.757	4,777
$n_i > 5$	0.427	0.348	0.105	0.715	2,700

Table 5: Test for heterogeneity in ships ballast trips. For each origin i we focus on the most common destination $j(i)$. For each ship leaving empty from i we define the variable y_t, which takes value one if the ship heads to $j(i)$ for trip t. We test the null that y_t follows the same binomial distribution for all ships. The table presents the distribution of p-values for this test across ships.

B.2 Random Search

In this section we investigate further the assumption of random search in the model of Section 2. As discussed in Section 4.2, we present an extract from Plakantonaki (2010), which is the internal manual of a large shipping firm, that describes the contracting process. Figure 5 shows an example of an email that the shipowner's broker receives in his inbox (again, this is one out of several thousand emails in a day).[43]

Plakantonaki (2010) writes: "This order means: Please indicate or offer a vessel to be chartered for the account of the charterer CORUS to load a full cargo of bulk coal, but minimum 47,500 mt and maximum 52,500 mt , i.e. 50,000 +/- 5%. The loading port will be in Poland and the discharge port is Immingham.

[43]We have corroborated in interviews with brokers that this email is indeed representative.

ACC CORUS (charterer)
50000/5 MT COAL (cargo size and type)
POLAND TO IMMINGHAM (load/discharge port/range)
23/30 APRIL (laycan)
3 DAYS SHINC, 24 HRS TTIME (loading time)
15000 MT SHINC, 24 HRS TTIME (discharge time)
3.75 TTL (total deductible commission)
PLS INDICATE (OR OFFER) (request for owners' interest)

Figure 5: Sample broker email, from Plakantonaki (2010)

The laydays for loading will commence on the 23rd of April and the canceling date will be on the 30th of April. For loading the laytime is 3 days Sundays Holidays Included, allowing up to 24 hours for delays prior to counting the 3 days allowed. [...] For discharging the time will be calculated on the basis of 15,000 t discharged per day, usually of 24 consecutive hours, Sundays Holidays Included, allowing up to 24 hours for delays prior counting of the time, if a berth is not available. The total commission payable by the owners to the charterer and the competitive brokers will be 3.75% on freight, deadfreight and demurrage earned. [...]

As soon as the in-house broker receives this information he considers whether he has a ship available, usually in the vicinity, which can safely arrive at the load port during the requested laycan. In the dry cargo example above, the ship should arrive in Poland ready to load on or after the 23rd of April and before the 30th of April.

Assuming for example that the M/V Prince, a Supramax vessel of about 58,000 mt summer deadweight is expected to complete its discharge in nearby Russian St. Petersburg around the 22nd of April, it can reach Polish ports on the 23rd of April or the 24th depending on the port."

A formal test for random search Finally, we turn to a more formal test of whether search in bulk shipping is random (or undirected). We contrast this with the case of directed search (see e.g. Moen, 1997), where carriers choose to search in a specific "market", i.e. a market for customers heading to a specific destination. Under directed search, profitable markets attract more carriers, thereby reducing their matching probabilities compared to less profitable markets. We can directly test this implication of directed search by checking whether in a given origin, i, ships' waiting time is different across destinations j. We use 15 regions, so for a given region there are (up to) 14 possible destinations; therefore there are

$\binom{14}{2} = 91$ such equalities to test for every origin i. Using a simple F-test we are only able to reject the null of no difference for 16% of the equalities.

In addition, we examine the coefficient of variation of matching probabilities within a given origin. Weighted by trade shares, the average coefficient of variation is just 8%. In contrast, the coefficient of variation of trip prices from a given origin is substantially higher and equal to 46%, suggesting that differences in the attractiveness of different types of customers is reflected in prices, but not in matching probabilities, as would be the case in directed search.

Finally, one might worry that shipowners attach different values to different destinations, due to unobserved heterogeneity, and direct their search toward trips heading to their favorite destination. To rule this out, we test whether there is persistent heterogeneity in trips' destinations for loaded trips across ships. In particular, we apply the framework described in Appendix B.1 to ships' *loaded* trips. The results are described in Table 6, and show that for the overwhelming majority of ships we fail reject the assumption that ships loaded trip destinations are the same across ships.

	p-value				N
	Mean	Median	p25	p75	
All ships	0.445	0.381	0.109	0.708	5,251
$n_i > 2$	0.434	0.354	0.103	0.693	4,988
$n_i > 5$	0.385	0.280	0.058	0.691	3,399

Table 6: Test for heterogeneity in ships loaded trips. For each origin i we focus on the most common destination $j(i)$. For each ship leaving empty from i we define the variable y_t, which takes value one if the ship heads to $j(i)$ for trip t. We test the null that y_t follows the same binomial distribution for all ships. The table presents the distribution of p-values for this test across ships and origins.

C Estimation and computation details

C.1 Model estimation and results

In this section we discuss the estimation of the model. We make four changes that render the model presented in Section 2 amenable to empirical analysis. First, we impose a specific pricing mechanism, Nash bargaining, with γ_i the ship bargaining coefficient in market i. Second, we add randomness to the discrete choice problem for ships of where to ballast, by adding idiosyncratic shocks to equation (4), so that it becomes,

$$U_i^s = \max_j V_{ij}^s + \sigma \epsilon_{ij} \tag{62}$$

where ϵ_{ij} are drawn i.i.d. from the Type I extreme value distribution with standard deviation σ. Third, we consider the version of the model with $\beta < 1$. In Appendix E we demonstrate that our efficiency results hold in this empirical model with discounting and idiosyncratic shocks. Fourth, we also add randomness to the exporters' problem (14), so that they solve the following discrete choice problem of whether and where to export,

$$\max_j \left\{ U_{ij}^e - \kappa_{ij} + \epsilon_{ij} \right\}$$

with ϵ_{ij} drawn i.i.d. from the Type I extreme value distribution; we normalize $U_{ii}^e - \kappa_{ii} = 0$ and interpret this as the option of not exporting at all. We also assume for simplicity that $w_{ij}(q) = w_{ij}$ for all ij.

The main parameters of interest are: the matching functions $m_i(s_i, e_i)$ for all i, the ship travel and wait costs c_{ij}^s, c_i^s, for all i, j, as well as the standard deviation of the logit shocks σ; the exporter valuations w_{ij}, the exporter waiting costs c_i^e (to gain power, we assume that c_{ij}^e do not vary over j), and entry costs κ_{ij} for all i, j; and the bargaining coefficients γ_i for all i. The available data consist of the matches m_i and ships s_i for all i, the ship ballast choice probabilities P_{ij}, for all ij, the average prices τ_{ij} on all routes ij, the exporter entry probabilities P_{ij}^e, for all ij as well as total trade values by country pair (Comtrade). Below we describe the estimation of each object in turn.

As explained in the main text, the estimation strategy follows BKP; the main reason we re-estimate some of the parameters is to allow the bargaining coefficients γ_i to vary by region, given the importance of these parameters (in BKP there is a scalar bargaining parameter across the globe). Moreover, we introduce trade data to obtain w_{ij}; this permits us to include the exporter waiting costs c_i^e which are not present in BKP.

Matching function estimation We briefly outline the approach adopted to estimate the matching function in BKP. The estimation draws from the literature on nonparametric identification (Matzkin, 2003) and non-separable instrumental variable techniques (e.g. Imbens and Newey, 2009); see also Brancaccio et al. (2020b) for a guide on the implementation of this approach in this and other settings.

To illustrate, assume that s and e are independent. We assume that $m(s,e)$ is continuous and strictly increasing in e, that it exhibits constant returns to scale (CRS), so that $m(as,ae) = am(s,e)$ for all $a > 0$, and that there is a known point $\{\bar{s},\bar{e},\bar{m}\}$, such that $\bar{m} = m(\bar{s},\bar{e})$. Suppose we have a sample $\{s_{it}, m_{it}\}_{t=0}^{T}$ for each market i. The unknowns of interest are the I matching functions $m_i(\cdot)$ and the exporters e_{it}, for all i,t; henceforth, we suppress the i subscript to ease notation. Let $F_{m|s}$ denote the distribution of matches conditional on ships, and F_e the distribution of exporters, e. Then at a given point $\{s_t, e_t, m_t\}$ we have:

$$\begin{aligned} F_{m|s=s_t}(m_t|s=s_t) &= \Pr(m(s,e) \le m_t|s=s_t) = \Pr\left(e \le m^{-1}(s,m_t)\,|s=s_t\right) \\ &= \Pr\left(e \le m^{-1}(s_t,m_t)\right) = F_e(e_t) \end{aligned}$$

This equation, along with the CRS assumption, allows us to recover the distribution $F_e(e)$, for all e: using the known point $\{\bar{s},\bar{e},\bar{m}\}$ and letting $a = e/\bar{e}$, for all e,

$$F_e(a\bar{e}) = F_{m|s=a\bar{s}}(m(a\bar{s},a\bar{e})\,|s=a\bar{s}) = F_{m|s=a\bar{s}}(a\bar{m}|s=a\bar{s})$$

We use this and vary a to trace out $\hat{F}_e(e)$, relying on a kernel density estimator for the conditional distribution $\hat{F}_{m|s=a\bar{s}}(a\bar{m}|s=a\bar{s})$. As shown in BKP the results are robust to alternative restrictions to CRS, such as a parametric assumption on e.[44]

Since it is unlikely that s and e are independent, we employ an instrument, which consists of the ocean weather conditions (unpredictable wind at sea) that shift the arrival of ships at a port without affecting the number of exporters (also employed in the search frictions test, see Section 4.2).[45] Table 7 presents the first stage estimates.

Figure 6 reports our estimates for search frictions, along with confidence intervals constructed from 200 bootstrap samples.

[44] We choose the known point, $\{\bar{s},\bar{e},\bar{m}\}$, to be of the form $1 = m(\bar{s},1)$, so that one exporter is always matched when there are $\bar{s}$ ships. We set $\bar{s}$ iteratively, to be the lowest value such that $m_t \le e_t$, for all t, thus obtaining a conservative bound on search frictions.

[45] Assume that an instrument z exists such that $s = h(z,\eta)$, with z independent of e, η. The approach now has two steps. In the first step, we recover η using the relationship $s = h(z,\eta)$. In the second step, we repeat the above conditioning on both s (as before) and η, $F_{m|s=s_t,\eta}(m_t|s=s_t,\eta) = F_{e|\eta}(e_t|\eta)$. We recover the unknowns of interest e and $m(\cdot)$, by integrating both sides over η.

	F-stat
North America West Coast	21.132
North America East Coast	18.429
Central America	17.877
South America West Coast	18.671
South America East Coast	16.889
West Africa	16.333
Mediterranean	46.072
North Europe	28.651
South Africa	13.153
Middle East	68.037
India	29.521
South East Asia	34.909
China	28.642
Australia	35.977
Japan-Korea	32.794

Table 7: First Stage, Matching Function Estimation. Regressions of the number of ships in each region on the unpredictable component of weather conditions in the surrounding seas. The table reports the F-statistic. For the construction of the instrument, see Table 1.

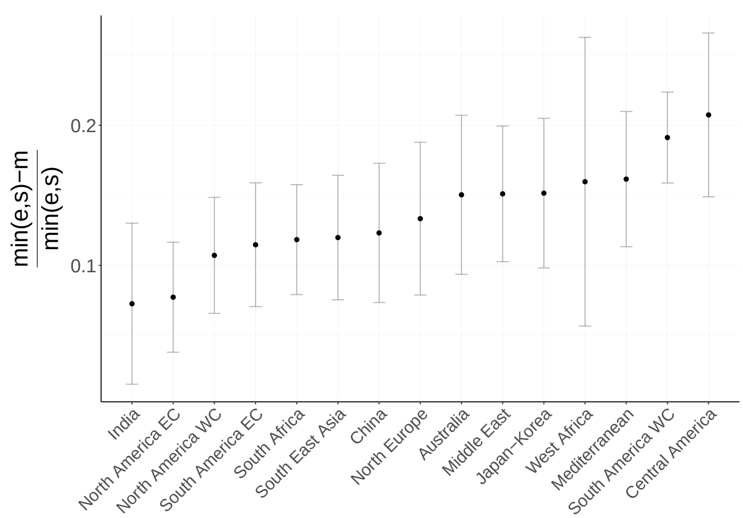


Figure 6: Search Frictions. Average weekly share of unrealized matches, with confidence intervals from 200 bootstrap samples.

Ship parameters We use the estimates for the ship parameters $\left\{c_{ij}^s, c_i^s, \sigma\right\}$ from BKP. To estimate these parameters, we used a Nested Fixed Point Algorithm (Rust, 1987): at every guess of the parameters

$\left\{c_{ij}^s, c_i^s, \sigma\right\}$ for all i, j, we employ a fixed point algorithm to solve for the ship value functions V_i^s, V_{ij}^s, U_i^s, for all i, j from equations (1), (3), (4), using the observed average prices for each route ij and the observed meeting probability λ_i^s (which is set equal to the average m_i/s_i). We then match the ship ballast choices predicted by our model and given by the logit choice probabilities,

$$P_{ij} = \frac{\exp\left(V_{ij}^s/\sigma\right)}{\sum_l \exp\left(V_{il}^s/\sigma\right)} \tag{63}$$

to the observed ballast choices. We do so by maximizing over the parameters via Maximum Likelihood. See BKP for the results, as well as further details on identification and estimation.

Exporter parameters and bargaining coefficients We are left with four sets of parameters: the exporter valuations w_{ij}, the waiting costs c_i^e, the bargaining coefficients γ_i; and the exporter entry costs κ_{ij}, for all i, j.

The valuations w_{ij} are the revenues of exporters in i from selling their commodities to destination j. We compute them using aggregate trade data from Comtrade, which reports product-level export values and quantities by country pair. We focus on bulk commodities and compute the average value of a cargo of commodities exported from each region i to each j, which forms our direct estimate for w_{ij}; details are provided in the next section.

Next, we turn to c_i^e and γ_i, for all i, which we estimate from observed shipping prices. Nash bargaining implies the surplus sharing condition,

$$(1-\gamma_i)\left(\tau_{ij} + V_{ij}^s - U_i^s\right) = \gamma_i\left[w_{ij} - \tau_{ij} - U_{ij}^e\right]$$

where if we substitute the exporter value U_{ij}^e from its steady state value, $U_{ij}^e = \left(-c_i^e + \lambda_i^e\left(w_{ij} - \tau_{ij}\right)\right) / \left(1 - \beta\left(1 - \lambda_i^e\right)\right)$, we obtain,

$$\tau_{ij} = \gamma_i \frac{c_i^e + \left(\left(1-\beta\right)\left(1-\lambda_i^e\right)\right) w_{ij}}{1 - \beta\left(1-\lambda_i^e\right) - \gamma_i \lambda_i^e} + \frac{\left(1-\gamma_i\right)\left(1-\beta\left(1-\lambda_i^e\right)\right)}{1-\beta\left(1-\lambda_i^e\right) - \gamma_i\lambda_i^e}\left(U_i^s - V_{ij}^s\right)$$

In this equation, the only unknowns are γ_i and c_i^e, for all i; indeed, note that λ_i^e is known from the matching function (set equal to m_i/e_i); U_i^s, V_{ij}^s are known once the ship cost parameters are known; w_{ij} is obtained from Comtrade data as described above; and β is calibrated to 0.995. We thus estimate γ_i and

c_i^e via non-linear least squares. Identification results from variation over the regions i, j: intuitively, the identification of the bargaining coefficient γ_i relies on the correlation of prices τ_{ij} and values w_{ij} across destinations j, while the inventory cost c_i^e matches the overall level of prices at origin i. To gain power, we restrict c_i^e to be constant within a continent.

Finally, exporter entry costs κ_{ij} are estimated using the exporter entry probabilities, which are given by

$$P_{ij}^e = \frac{\exp\left(U_{ij}^e - \kappa_{ij}\right)}{1 + \sum_{l \neq i} \exp\left(U_{il}^e - \kappa_{il}\right)} \tag{64}$$

for $j \neq i$, where $P_{ii}^e = 1/\left(1 + \sum_{l \neq i} \exp\left(U_{il}^e - \kappa_{il}\right)\right)$ is interpreted as the option of not exporting at all. Then,

$$\ln P_{ij}^e - \ln P_{ii}^e = U_{ij}^e - \kappa_{ij} = \frac{-c_i^e + \lambda_i^e\left(w_{ij} - \tau_{ij}\right)}{1 - \beta\left(1 - \lambda_i^e\right)} - \kappa_{ij}$$

where κ_{ij} is the only unknown.[46] The results are presented in Table 8.

C.2 Exporter valuations

We construct exporter valuations, w_{ij}, from product-level data on export value and quantity by country-pair, obtained from Comtrade. We select bulk commodities among all possible 4-digit HS product codes. The list includes cereals (except rice and barley); oil seeds (which consists of mostly soybeans); cocoa beans; salt and cement; ores; mineral fuels (except petroleum coke); fertilizers; fuel wood and wood pulp; metals; cermets and articles thereof.

To compute the average value of a cargo exported from region i to j, we first compute the average "price" of a ton exported by dividing total export value by total export quantity from i to j. Then, we multiply this price by the average ship tonnage capacity in our sample.[47]

Finally, although most countries belong to one of our regions (depicted in Figure 7), the USA and Canada each belong to two regions (according to the coast). We thus need to split the Comtrade data for the USA and Canada into east and west coast export values. To do so, we employ data on state-

[46] To recover P_{ii}^e, the share of the "outside good", corresponding to the choice of not exporting, we use the total production of the relevant commodities for each region i.

[47] This is robust to using the average ship tonnage capacity on route ij.

level exports from the US Census, as well as on province-level exports from the Canadian International Merchandise Trade Database. In particular, we assign every state (province) to either the east or the west coast and compute, for every product, the share of the total value of trade in that commodity that is exported by east and west coast states (provinces). Then, we compute the total value and quantity of trade for the region East Coast of North America (West Coast of North America) by summing over products the share of the value of east (west) coast trade by the total value of the country's trade for the USA and Canada. Implicitly, this approach assumes that export values from these two regions are only different due to the composition of products, not their prices.

C.3 Algorithm to compute the efficient allocation

Here, we describe the algorithm employed to compute the steady state of our model. In order to simulate both the market equilibrium and the efficient allocation we approximate the matching function that we obtained non-parametrically with a Cobb Douglas. In particular, for each region we impose $m_{it} = A_i s_{it}^{1-\alpha_i} e_{it}^{\alpha_i}$, and select the parameters (A_i, α_i) through non-linear least-squares using the non-parametrically estimated exporters.

The algorithm proceeds as follows:

1. Make an initial guess for $\{U^{e,0}, \tau^0, s^0, E^0\}$.

2. At each iteration k, inherit $\left\{U^{e,k-1}, \tau^{k-1}, s^{k-1}, E^{k-1}\right\}$. Let G^{k-1}, e^{k-1}, and q^{k-1} denote the associated destination shares, searching exporters, and matches respectively.[48] Moreover, let $\lambda^{e,k-1}$ and $\lambda^{s,k-1}$ denote the associated matching rates. We update our guess according to the following steps:

   (a) First, in an inner loop we compute the ship optimal policy and value function implied by the matching rates $\lambda^{s,k-1}$, prices τ^{k-1} , and destination shares G^{k-1}. In particular, after initializing $V^{s,0}$, repeat the following steps until convergence

      i. At iteration h, compute the value of traveling $V_{ij}^{s,h}$ from $V_{ij}^{s,h} = \frac{-c_{ij}^s + d_{ij}\beta V_i^{s,h-1}}{1-\beta(1-d_{ij})}$.

[48]That is, $e_i^{k-1} = \sum_j e_{ij}^{k-1}$, $G_{ij}^{k-1} = \frac{e_{ij}^{k-1}}{e_i^{k-1}}$, and $q_i^{k-1} = m\left(s_i^{k-1}, e_i^{k-1}\right)$

ii. Compute the value $U_i^{s,h}$ from:

$$U_i^{s,h} = \sigma \ln \left(\exp \frac{\beta V_i^{s,h-1}}{\sigma} + \sum_{j \neq i} \exp \frac{V_{ij}^{s,h}}{\sigma} \right) + \sigma \gamma^{euler}$$

where γ^{euler} is the Euler constant.[49]

iii. Update $V_i^{s,h}$ from $V_i^{s,h} = -c_i^s + \left(1 - \lambda_i^{s,k-1}\right) U_i^{s,h} + \lambda_i^{s,k-1} \sum_j G_{ij}^{k-1} \left(V_{ij}^{s,h} + \tau_{ij}^{k-1}\right)$

iv. Upon convergence, we set $V_{ij}^{s,k} = V_{ij}^{s,\infty}$, $V_i^{s,k} = V_i^{s,\infty}$, $U_i^{s,k} = U_i^{s,\infty}$, and compute the ship optimal choice probabilities based on $P_{ij}^k = \exp\left(\frac{V_{ij}^{s,k}}{\sigma}\right) / \left(\sum_{l \neq i} \exp\left(\frac{V_{il}^{s,k}}{\sigma}\right) + \exp\left(\frac{\beta V_i^{s,k}}{\sigma}\right)\right)$ for $i \neq j$ and $P_{ii}^k = \exp\left(\frac{\beta V_i^{s,k}}{\sigma}\right) / \left(\sum_{l \neq i} \exp\left(\frac{V_{il}^{s,k}}{\sigma}\right) + \exp\left(\frac{\beta V_i^{s,k}}{\sigma}\right)\right)$ for $i = j$.

(b) To update the efficient prices τ^k compute the total surplus from matching as,[50]

$$\Delta_{ij}^k = w_{ij} - \delta \beta U_{ij}^{e,k-1} + V_{ij}^{s,k} - U_i^{s,k},$$

and compute the efficient prices based on,

$$\tau_{ij}^k = w_{ij} - \delta \beta U_{ij}^{e,k-1} - \Delta_{ij}^k + \alpha_i \sum_j G_{ij}^{k-1} \Delta_{ij}^k,$$

where α_i denotes the elasticity of the matching function with respect to the number of exporters. Similarly, compute prices under Nash bargaining using the surplus sharing condition.

(c) Update the exporter value function $U^{e,k}$ based on the efficient prices τ^k and matching rates $\lambda^{e,k-1}$ setting $U_{ij}^{e,k} = \frac{-c_i^e + \lambda_i^{e,k-1}\left(w_{ij} - \tau_{ij}^k\right)}{1 - \beta\delta\left(1 - \lambda_i^{e,k-1}\right)}$.

(d) Finally, update the number of ships and exporters searching $\left\{s^k, E^k\right\}$ according to

$$s_i^k = \sum_j P_{ij}^k \left(s_j^{k-1} - q_j^{k-1}\right) + \sum_j q_{ij}^{k-1},$$

[49] This is the closed form expression for the expectation of the maximum over multiple choices, and is obtained by integrating U_i^s over the distribution of ϵ.

[50] Following BKP we assume that unmatched exporters survive with probability δ so that their effective discount factor is $\beta\delta$ (this is also true in the estimation procedure even though it was omitted there for notational simplicity). We calibrate $\delta = 0.99$. This makes no difference in our theoretical analysis.

and

$$e_{ij}^{k}=\underbrace{\mathcal{E}_i\frac{\exp\left(U_{ij}^{e,k}-\kappa_{ij}\right)}{1+\sum_{l\neq i}\exp\left(U_{il}^{e,k}-\kappa_{il}\right)}}_{\text{new entrants}}+\delta\underbrace{\left(e_{i}^{k-1}-q_{i}^{k-1}\right)}_{\text{unmatched}}$$

where $\mathcal{E}_i$ is the mass of potential entrants.

3. If $\left\|s^{k}-s^{k-1}\right\|<\epsilon$, $\left\|E^{k}-E^{k-1}\right\|<\epsilon$, $\left\|U^{e,k}-U^{e,k-1}\right\|<\epsilon$, and $\left\|\tau^{k}-\tau^{k-1}\right\|<\epsilon$ stop; otherwise go back to point (a).

	Exporter wait costs c_i^e	Ship bargaining coefficient γ_i	Average exporter value $\bar{w}_i$
North America West Coast	83.49 (10.72)	0.384 (0.018)	13,738
North America East Coast	83.49 (10.72)	0.585 (0.012)	12,192
Central America	302.3 (69.28)	0.344 (0.038)	14,350
South America West Coast	302.3 (69.28)	0.259 (0.017)	20,096
South America East Coast	302.3 (69.28)	0.371 (0.042)	6,971
West Africa	396.9 (512.74)	0.292 (0.078)	4,547
Mediterranean	3.44 (8.60)	0.412 (0.014)	10,508
North Europe	3.44 (8.60)	0.517 (0.014)	14,577
South Africa	396.9 (512.74)	0.24 (0.075)	6,224
Middle East	20.53 (9.04)	0.615 (0.026)	7,160
India	20.53 (9.04)	0.568 (0.022)	6,305
South East Asia	174.01 (44.00)	0.215 (0.036)	4,918
China	282.8 (77.86)	0.194 (0.028)	8,231
Australia	174.01 (44.00)	0.388 (0.037)	12,475
Japan-Korea	282.8 (77.86)	0.265 (0.038)	2,977

Table 8: Average exporter valuation (over destinations), wait costs and bargaining coefficients estimates. All the estimates are in 1,000 USD. To gain power, we restrict exporter wait costs to be constant within a continent. Standard errors computed from 200 bootstrap samples.

D Additional figures and tables

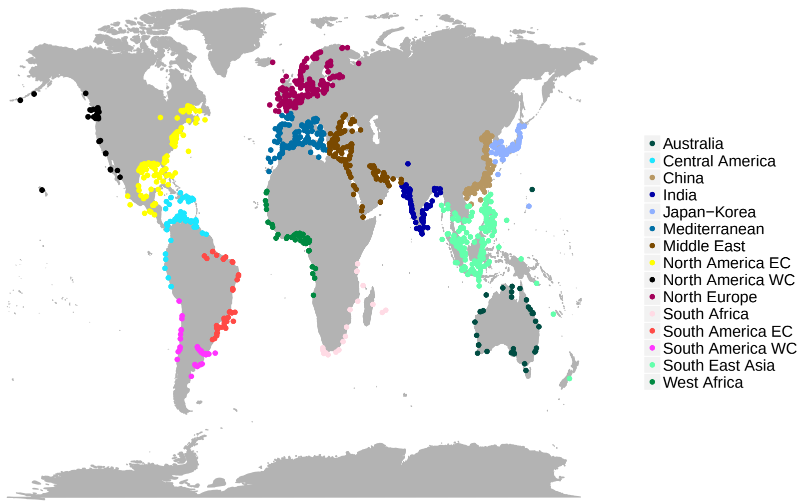


Figure 7: Definition of regions. Each color depicts one of the 15 geographical regions.

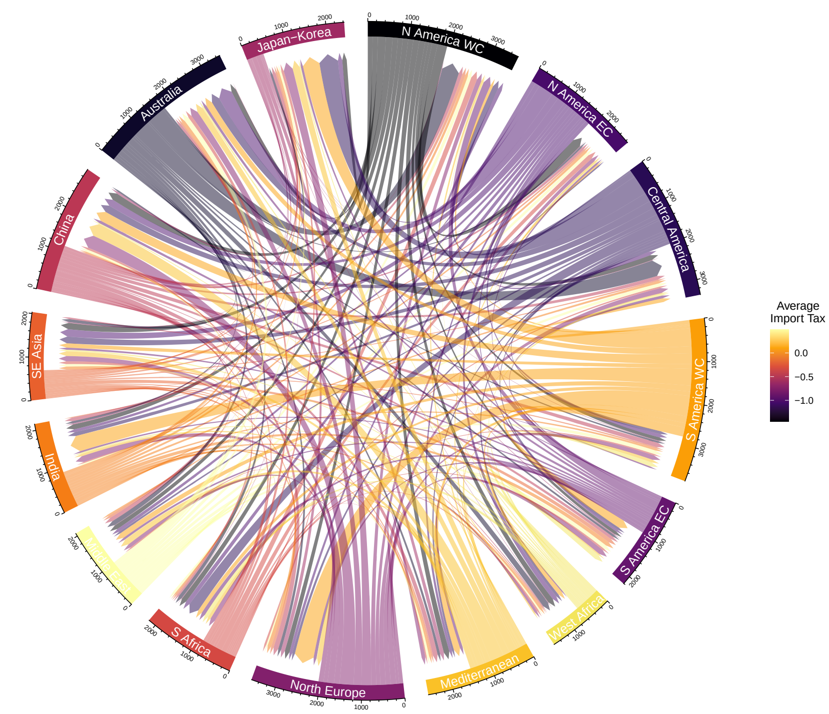


Figure 8: The figure plots the matrix of exporters' revenues w_{ij}. For each region, the outgoing arrows depict the value of export to each of the other regions, as the size of the arrow is proportional to the average exporter revenues w_{ij} along the route. The incoming arrows depict the value of import from each of the other regions. Therefore, the length of the arc corresponding to each region reflects the value of import and exports. In addition, for each region, the color of outgoing arrows (and corresponding arc) reflects the value of the average pooling taxes, h^q_{ij}, for the region.

	I	II	III
	log(price per day)		
Probability of ballast		0.234**	0.556**
		(0.030)	(0.081)
Avg duration of ballast trip (log)		0.166**	0.065**
		(0.014)	(0.032)
Coal			0.088**
			(0.045)
Fertilizer			0.245**
			(0.051)
Grain			0.131**
			(0.048)
Ore			0.124**
			(0.045)
Steel			0.135**
			(0.049)
Constant	10.284**	9.127**	8.915**
	(0.103)	(0.099)	(0.408)
Destination FE	Yes	No	No
Origin FE	Yes	Yes	Yes
Ship type FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Obs	11,014	11,011	1,662
Adjusted R^2	0.694	0.674	0.664

$^{**}p < 0.05, ^{*}p < 0.1$

Table 9: Shipping price regressions (Table II in BKP). The dependent variable is the logged price per day in USD. The independent variables include combinations of: the average frequency of ballast traveling after the contract's destination (Probability of ballast), the average logged duration (in days) of the ballast trip after the contract's destination, as well as ship type, origin, destination and quarter FEs. The product is reported in only 20% of the sample, so the regression in column III has substantially fewer observations. The omitted product category is cement.

E Supplemental Material: Discounting, preference shocks and out of steady state dynamics

In this section we show that the main results of Section 3.2 are valid in a more general setup. In particular, we extend the model of Section 2 to allow for idiosyncratic preference shocks in carriers relocation choice (this is not essential, but it is relevant in our empirical application), as well as out of steady state dynamics, and we derive an efficiency result analogous to that of Theorem 2.

E.1 Model

We begin by laying out the model focusing on the changes made compared to Section 2.

States and transitions In this Appendix we do not consider the steady state equilibrium. Hence, we now state explicitly the dependence of actions and value functions on the relevant state variables and transitions, which were only implicit in the model of Section 2. At the beginning of a given time period, the state of the economy is described by a vector,

$$z = (x, y) \in \mathbb{R}_+^{I\times I} \times \mathbb{R}_+^{I\times I}.$$

The first element of z, $x = (x_{ij})_{i,j\in I}$, corresponds to the supply at every origin i,

- x_{ii} is the measure of carriers waiting at location i
- x_{ij} is the measure of carriers traveling from i to j, either empty or full, for every destination $j \neq i$.

The second element of z, $y = (y_{ij})_{i,j\in I}$, corresponds to demand. For every origin-destination pair ij, y_{ij} is the measure of customers who are waiting on route ij at the beginning of the current period. These are customers that entered in some previous period and have not yet been matched with a carrier.

At a given state z, the choice sets that agents face, as well as the search and matching process are the same as in Section 2. At each origin i, a measure $s_i \leq x_{ii}$ of carriers choose to search for a customer, while the remaining measure $x_{ii} - s_i$ choose to remain inactive. Similarly, a measure $e_{ij} \geq y_{ij}$ of customers search for a carrier on route ij, so that $e_{ij} - y_{ij}$ is the measure of new customers joining the existing search pool.

Once a customer and a carrier meet, they can choose whether to match or remain unmatched. The outcome of this process is a vector (b, q) describing the measure of carriers that start traveling empty (b_{ij}) or full (q_{ij}) on each route ij. The state transitions as a function of the allocation (s, E, q, b) are as follows for all ij:

$$
\begin{aligned}
x_{ii}^{+1}(s, E, q, b|z) &= x_{ii} - s_i + \sum_j d_{ji}(x_{ji} + q_{ji} + b_{ji}) \\
x_{ij}^{+1}(s, E, q, b|z) &= (1 - d_{ij})(x_{ij} + q_{ij} + b_{ij}) \\
y_{ij}^{+1}(s, E, q, b|z) &= e_{ij} - q_{ij}. \qquad (65)
\end{aligned}
$$

The feasibility constraints on the allocation (s, E, q, b) are as follows for all i, j:

$$
\begin{aligned}
x_{ii} \geq s_i, \quad e_{ij} \geq y_{ij}, \quad s_i, e_{ij}, q_{ij}, b_{ij} \geq 0 \qquad (66) \\
\sum_j (q_{ij} + b_{ij}) = s_i, \quad m_i(s_i, e_i) G_{ij} \geq q_{ij}
\end{aligned}
$$

Prices, expectations and allocation rules The pricing rule maps each state into the associated vector of transportation prices on each route, $\boldsymbol{\tau} : z \mapsto \boldsymbol{\tau}(z) = (\boldsymbol{\tau}_{ij}(z))_{i,j\in I}$. As in Section 2, we begin by remaining agnostic regarding the structure of the pricing rule, and later we characterize the pricing rules that are consistent with efficient equilibria and compare them to Nash bargaining.

In state z carriers expect to meet customers at rate $\boldsymbol{\lambda}_i^s(z)$ in location i, and customers expect to meet carriers at rate $\boldsymbol{\lambda}_i^e(z)$, where $\boldsymbol{\lambda}^s : z \mapsto \boldsymbol{\lambda}^s(z) = (\boldsymbol{\lambda}_i^s(z))_{i\in I}$, $\boldsymbol{\lambda}^e : z \mapsto \boldsymbol{\lambda}^e(z) = (\boldsymbol{\lambda}_i^e(z))_{i\in I}$. Agents make optimal choices under rational expectations about the state transitions, the matching probabilities and prices at each state, generating an allocation rule $(\boldsymbol{s}, \boldsymbol{E}, \boldsymbol{q}, \boldsymbol{b}) : z \mapsto (\boldsymbol{s}(z), \boldsymbol{E}(z), \boldsymbol{q}(z), \boldsymbol{b}(z))$, mapping states into feasible allocations. That is, for every state z, $(\boldsymbol{s}(z), \boldsymbol{E}(z), \boldsymbol{q}(z), \boldsymbol{b}(z))$ satisfies (66).

Similarly to Section 2, we will sometimes denote an allocation rule by $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{b})$, where $\boldsymbol{e}_i(z) = \sum_j \boldsymbol{e}_{ij}(z)$ and $\boldsymbol{G}_{ij}(z) = \boldsymbol{e}_{ij}(z)/\boldsymbol{e}_i(z)$, and we will often refer to the first triplet $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G})$ as a search rule.

Preference shocks and carrier optimality Carriers' payoff structure is the same as in Section 2. In addition, we allow for stochasticity in carriers' preferences for destinations. The stochastic component at each origin i is represented by a random vector $\epsilon_i = (\epsilon_{ij})_{i,j\in I}$ that enters the carriers' utility of relocating to different destinations additively, is i.i.d. across carriers and satisfies the conditional independence assumption, $\epsilon_i^{+1}|z^{+1} \perp \epsilon_i, z$. To simplify the exposition, we assume that ϵ_i is independent of z and i, so that $\forall i, z : \epsilon_i \sim \mathbb{P} \in \Delta\mathbb{R}^I$, although this assumption is not needed for the results. We assume that $\mathbb{P}$ has full support and that it admits a continuous density.

The value of a carrier that remained unmatched at origin i at state z depends on the particular realization of the shock. We denote its expectation by

$$U_i^s(z) = \mathrm{E}_{\mathbb{P}} \max_j \left(V_{ij}^s(z) + \epsilon_{ij} \right). \tag{67}$$

The values of a carrier traveling from i to j and a carrier waiting in i are given by:

$$\begin{aligned} V_{ij}^s(z) &= -\, c_{ij}^s + \beta \left[d_{ij} V_j^s\left(z^{+1}\right) + (1 - d_{ij}) V_{ij}^s\left(z^{+1}\right) \right] \\ V_i^s(z) &= \max \left\{ -c_i^s + \boldsymbol{\lambda}_i^s(z) \sum_{j \neq i} G_{ij} \Delta_{ij}^s(z) + U_i^s(z), \beta V_i^s\left(z^{+1}\right) \right\} \end{aligned}$$

as before, where

$$\Delta_{ij}^s(z) = \max \left\{ \boldsymbol{\tau}_{ij}(z) + V_{ij}^s(z) - U_i^s(z), 0 \right\}$$

is the carrier expected surplus of being matched with respect of being unmatched.

Denote by P^b the matrix of carrier relocation choice probabilities associated with $b \in \mathbb{R}^{I \times I}$:

$$P_{ij}^b = b_{ij} / \sum_k b_{ik}.$$

for all ij. Optimality in state z requires that for all ij,

$$P_{ij}^{\boldsymbol{b}(z)} = \mathbb{P} \left[V_{ij}^s(z) + \epsilon_{ij} = \max_k \left(V_{ik}^s(z) + \epsilon_{ik} \right) \right]. \tag{68}$$

The remaining optimality conditions of Section 2 still hold. In particular, carriers search only when it is profitable to do so:

$$\boldsymbol{s}_i(z) > 0 \rightarrow V_i^s(z) = -c_i^s + \boldsymbol{\lambda}_i^s(z) \sum_{j \neq i} G_{ij} \Delta_{ij}^s(z) + U_i^s(z) \tag{69}$$

Moreover, they do not reject any match yielding a strictly positive surplus, and they accept only matches yielding a positive surplus:

$$\begin{aligned} &\boldsymbol{q}_{ij}(z) < \boldsymbol{\lambda}_i^s(z)\, \boldsymbol{s}_i(z)\, G_{ij} \rightarrow \Delta_{ij}^s(z) = 0 \\ &\boldsymbol{q}_{ij}(z) > 0 \rightarrow \Delta_{ij}^s(z) = \boldsymbol{\tau}_{ij}(z) + V_{ij}^s(z) - U_i^s(z). \end{aligned} \tag{70}$$

Customer optimality Customer value functions are the same as in Section 2, but we make the dependence on the state of the economy explicit. In state z, the meeting surplus of the marginal customer (with respect to being unmatched) is given by

$$\Delta_{ij}^e(z) = \max\left\{ w_{ij}(\boldsymbol{q}(z)) - \boldsymbol{\tau}_{ij}(z) - \beta U_{ij}^e\left(z^{+1}\right), 0\right\},$$

where $U_{ij}^e(z)$ is the value of customer with destination j that is searching for a carrier in location i:

$$U_{ij}^e(z) = -c_{ij}^e + \boldsymbol{\lambda}_i^e(z)\, \Delta_{ij}^e(z) + \beta U_{ij}^e\left(z^{+1}\right). \tag{71}$$

Optimality requires that the marginal customer does not reject a match yielding a strictly positive surplus:

$$\boldsymbol{q}_{ij}(z) < \boldsymbol{\lambda}_i^e(z)\,_{ij}(z) \rightarrow \Delta_{ij}^e(z) = 0. \tag{72}$$

The measure of customers searching on each route ij is pinned down by a free entry condition for the marginal customer:

$$U_{ij}^e(z) - \kappa_{ij} \leq 0, \quad \text{with equality if } \boldsymbol{e}_{ij}(z) > y_{ij}. \tag{73}$$

Equilibrium An outcome is a tuple $(\boldsymbol{s}, \boldsymbol{E}, \boldsymbol{q}, \boldsymbol{b}, \boldsymbol{\tau})$ consisting of an allocation rule and a price rule.

Definition 5. An outcome is a Markovian equilibrium if, for every state z:

1. $(\boldsymbol{s}(z), \boldsymbol{E}(z), \boldsymbol{q}(z), \boldsymbol{b}(z))$ satisfies the feasibility constraints (66).

2. $(\boldsymbol{s}(z), \boldsymbol{q}(z), \boldsymbol{b}(z))$ satisfies the carrier optimality conditions (67)-(70) given $\boldsymbol{\tau}(z), \boldsymbol{\lambda}^s(z), z^{+1}$ and $\boldsymbol{G}(z)$.

3. $(\boldsymbol{E}(z), \boldsymbol{q}(z))$ satisfies the customer optimality and free entry conditions (71)-(73) given $\boldsymbol{\tau}(z), \boldsymbol{\lambda}^e(z)$ and z^{+1}.

4. Expectations are consistent with the realized outcomes:

$$\forall i: \boldsymbol{\lambda}_i^s(z) = m_i(\boldsymbol{s}_i(z), \boldsymbol{e}_i(z)) / \boldsymbol{s}_i(z), \; \boldsymbol{\lambda}_i^e(z) = m_i(\boldsymbol{s}_i(z), \boldsymbol{e}_i(z)) / \boldsymbol{e}_i(z)$$
$$z^{+1} = z^{+1}(\boldsymbol{s}(z), \boldsymbol{E}(z), \boldsymbol{q}(z), \boldsymbol{b}(z)).$$

$(\boldsymbol{s}, \boldsymbol{E}, \boldsymbol{q}, \boldsymbol{b})$ is an equilibrium allocation rule if there exists a price rule $\boldsymbol{\tau}$ such that $(\boldsymbol{s}, \boldsymbol{E}, \boldsymbol{q}, \boldsymbol{b}, \boldsymbol{\tau})$ is a Markovian equilibrium.

E.2 Externalities and efficiency

The social planner solves an infinite horizon constrained Markov decision problem in which, conditional on every initial state z, he chooses a dynamic allocation rule maximizing the discounted sum of future social payoffs.

The social welfare $W^p(s, e, G, q, b; z)$ at each state z and for every allocation (s, e, G, q, b) entails the welfare terms encountered in Section 3.1, but involves an additional term that captures the welfare due to carrier preference shocks. This term is given by $\left(\sum_j b_{ij}\right) f\left(P_i^b\right)$ where $f(P_i^b)$ represents the value associated with the best allocation of shocks to destinations at i conditional on the aggregate choice probabilities being given by P_i^b:[51]

$$W^p(s, e, G, q, b; z) \equiv W(q) - \sum_{ij}(x_{ij} + q_{ij} + b_{ij}) c_{ij}^s - \sum_i s_i c_i^s - \sum_i e_i \sum_j G_{ij} c_{ij}^e - \sum_{ij}(e_i G_{ij} - y_{ij}) \kappa_{ij} + \sum_{ij} b_{ij} f\left(P_i^b\right).$$

In what follows, we use the upper bar notation $\bar{a} = (a^t)_{t=0}^{\infty}$ for infinite sequences. When dealing with a sequence of allocations $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)$ and an initial state z^0, unless stated otherwise, it is understood

[51]Formally:

$$f\left(P_i^b\right) = \max_{\pi \in \Pi\left(P_i^b\right)} \mathrm{E}_\pi(\epsilon_j) \tag{74}$$

where the expectation on the right hand side is with respect to a joint realization of the vector $\epsilon = (\epsilon_j)_{j \in I}$ and the destination $j \in I$, and $\Pi\left(P_i^b\right)$ is the set of all probability measures $\pi \in \Delta\left(\mathbb{R}^I \times I\right)$ such that the marginal of π over I is P_i^b and the marginal of π over $\mathbb{R}^I$ is $\mathbb{P}$. For a discussion, the reader is referred to Galichon (2018). For example, if ϵ_i is distributed according to a logit, we have $f\left(P_i^b\right) = -\sum_j P_{ij}^b \ln P_{ij}^b$.

that $\bar{z}$ refers to the sequence of states induced by $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)$ from z^0:

$$z^{t+1} = z^{+1}\left(s^t, e^t, G^t, q^t, b^t; z^t\right).$$

for $t \geq 0$. Moreover, when dealing with a feasible allocation rule $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{b})$ and an initial state z^0, it is understood that $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)$ refers to the sequence of allocations induced by $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{b})$ from z^0:

$$\left(s^t, e^t, G^t, q^t, b^t\right) = \left(\boldsymbol{s}\left(z^t\right), \boldsymbol{e}\left(z^t\right), \boldsymbol{G}\left(z^t\right), \boldsymbol{q}\left(z^t\right), \boldsymbol{b}\left(z^t\right)\right).$$

The planner's dynamic problem at state z^0 is given by

$$\begin{aligned} V^p\left(z^0\right) = & \max_{\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)} \sum_{t=0}^{\infty} \beta^t W^p\left(s^t, e^t, G^t, q^t, b^t; z^t\right) \\ & \text{s.t. } x_{ii}^t \geq s_i^t \\ & e_i^t G_{ij}^t \geq y_{ij}^t \\ & \sum_j \left(q_{ij}^t + b_{ij}^t\right) = s_i^t \\ & m_i\left(s_i^t, e_i^t\right) G_{ij}^t \geq q_{ij}^t \\ & \sum_j G_{ij}^t = 1 \\ & s_i^t, e_i^t, G_{ij}^t \geq 0, \quad \forall i, j, t \end{aligned} \tag{75}$$

Definition 6. An allocation rule $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{b})$ is efficient at a state z^0 if $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)$ solves Problem (75).

Similarly to Section 2, we distinguish three different potential sources of inefficiency. To do so, for each state z^0, let $\mathcal{A}\left(z^0\right)$ be the set of allocation sequences $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right)$ which are feasible from z^0, that is, they satisfy the constraints of Problem (75). Let also

$$\mathcal{SA}\left(z^0\right) = \left\{\left(\bar{s}, \bar{e}, \bar{G}\right) : \exists\left(\bar{q}, \bar{b}\right), \left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right) \in \mathcal{A}\left(z^0\right)\right\}$$

be the set of feasible sequences of search allocations, and

$$\mathcal{SA}\left(z^0|\bar{e},\bar{G}\right)=\left\{\bar{s}:\ \left(\bar{s},\bar{e},\bar{G}\right)\in\mathcal{SA}\left(z^0\right)\right\}$$
$$\mathcal{SA}\left(z^0|\bar{s},\bar{G}\right)=\left\{\bar{e}:\ \left(\bar{s},\bar{e},\bar{G}\right)\in\mathcal{SA}\left(z^0\right)\right\}$$
$$\mathcal{SA}\left(z^0|\bar{s},\bar{e}\right)=\left\{\bar{G}:\ \left(\bar{s},\bar{e},\bar{G}\right)\in\mathcal{SA}\left(z^0\right)\right\}.$$

For every $\left(\bar{s},\bar{e},\bar{G}\right)\in\mathcal{SA}\left(z^0\right)$, we define the maximum dynamic welfare attainable by this sequence by:

$$V^p\left(\bar{s},\bar{e},\bar{G},z^0\right)=\max_{\bar{q},\bar{b}}\sum_{t=0}^{\infty}\beta^t W^p\left(s^t,e^t,G^t,q^t,b^t;z^t\right) \tag{76}$$
$$\text{s.t.}\left(\bar{s},\bar{e},\bar{G},\bar{q},\bar{b}\right)\in\mathcal{A}\left(z^0\right)$$

so that we have

$$V^p\left(z^0\right)=\max_{\left(\bar{s},\bar{e},\bar{G}\right)\in\mathcal{SA}(z^0)}V^p\left(\bar{s},\bar{e},\bar{G},z^0\right).$$

Given an equilibrium allocation rule $(\boldsymbol{s},\boldsymbol{e},\boldsymbol{G},\boldsymbol{q},\boldsymbol{b})$ and an initial state z^0 we say that:

(i) Carriers internalize thin/thick market externalities at z^0 if $\bar{s}$ solves

$$\max_{\bar{s}'\in\mathcal{SA}\left(z^0|\bar{e},\bar{G}\right)}V^p\left(\bar{s}',\bar{e},\bar{G},z^0\right)$$

(ii) Customers internalize thin/thick market externalities at z^0 if $\bar{e}$ solves

$$\max_{\bar{e}'\in\mathcal{SA}\left(z^0|\bar{s},\bar{G}\right)}V^p\left(\bar{s},\bar{e}',\bar{G},z^0\right)$$

(iii) Customers internalize pooling externalities at z^0 if $\bar{G}$ solves

$$\max_{\bar{G}'\in\mathcal{SA}(z^0|\bar{s},\bar{e})}V^p\left(\bar{s},\bar{e},\bar{G}',z^0\right).$$

Next we state the equivalent of Theorem 2 in the current framework. Given $(\boldsymbol{s},\boldsymbol{e},\boldsymbol{G})$, we denote by $\eta_i^s(z)=d\ln m_i\left(\boldsymbol{s}_i(z),\boldsymbol{e}_i(z)\right)/d\ln s_i$ and $\eta_i^e(z)=d\ln m_i\left(\boldsymbol{s}_i(z),\boldsymbol{e}_i(z)\right)/d\ln e_i$. For simplicity, in order to avoid delving into corner conditions, in the statement below we assume that the equilibrium path originating from z^0 is such that we have $s_i^t, e_i^t > 0$ for every t, i.

Theorem 4. *Suppose that at a given state z^0, Problem (76) admits a unique optimal solution, and let $(\boldsymbol{s}, \boldsymbol{e}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{b})$ be an equilibrium allocation rule. Then the following statements hold:*[52]

(i) Carriers internalize thin/thick market externalities at z^0 if and only if, for every $t \geq 0$:

$$\forall i \in I: \sum_j \boldsymbol{G}_{ij}^t \left(z^t\right) \Delta_{ij}^s \left(z^t\right) = \eta_i^s \left(z^t\right) \sum_j \boldsymbol{G}_{ij} \left(z^t\right) \left(\Delta_{ij}^s \left(z^t\right) + \Delta_{ij}^e \left(z^t\right)\right).$$

(ii) Customers internalize thin/thick market externalities at z^0 if and only if, for every $t \geq 0$:

$$\forall i \in I: \sum_j \boldsymbol{G}_{ij} \left(z^t\right) \Delta_{ij}^e \left(z^t\right) = \eta_i^e \left(z^t\right) \sum_j \boldsymbol{G}_{ij} \left(z^t\right) \left(\Delta_{ij}^s \left(z^t\right) + \Delta_{ij}^e \left(z^t\right)\right)$$

(iii) Customers internalize pooling externalities at z^0 if and only if, for every $t \geq 0$, for each origin i,

$$\Delta_{ij}^s \left(z^t\right) = \max_{k \neq i} \Delta_{ik}^s \left(z^t\right)$$

for every ij such that $\boldsymbol{G}_{ij} \left(z^t\right) > 0$.

The following section provides the proof.

E.3 Proof of Theorem 4

The proof follows the same reasoning as under the steady state assumption, but has to overcome a number of technical difficulties due to the infinite dimensional form of the planner's optimization problem.

E.3.1 Preliminaries

Let X be a compact and convex subset of $\mathbb{R}^N$ for some $N \in \mathbb{N}$, and $\beta \in (0,1)$. For each pair of sequences $\bar{x}, \bar{y} \in X^{\mathbb{N}\cup\{0\}}$ we define the inner product $\langle \bar{x}, \bar{y} \rangle = \sum_{t=0}^{\infty} \beta^t x^t \cdot y^t$ for all $\bar{x}, \bar{y}$, where $\cdot$ denotes the standard

[52] Formally, the only if parts of statements (i) to (iii) hold for almost every sequence $\left(\bar{s}, \bar{e}, \bar{G}\right) \in \mathcal{SA}\left(z^0\right)$. That is, there exists a dense subset D of $\mathcal{SA}\left(z^0\right)$ such that the only if part of the statements hold whenever $\left(\bar{s}, \bar{e}, \bar{G}\right) \in D$. See Section E.3 for details.

inner product on $\mathbb{R}^N$. Define the norm $\|\bar{x}\| = \sqrt{\langle \bar{x}, \bar{x} \rangle}$ and $L^{2,\beta} = \left\{ \bar{x} \in X^{\mathbb{N} \cup \{0\}} : \|\bar{x}\| < \infty \right\}$. Then $\left(L^{2,\beta}, \|\cdot\| \right)$ is a Banach space.

Let $\mathcal{X} \subseteq L^{2,\beta}$ be a convex set and $f : \mathcal{X} \to \mathbb{R}$ be a continuous and concave function.

Definition 7. For every $\bar{x} \in \mathcal{X}$, the super gradient of f at $\bar{x}$, denoted $\partial f(\bar{x})$, is the set of all sequences $\bar{y} \in X^{\mathbb{N} \cup \{0\}}$ such that, for every $\bar{x}' \in \mathcal{X}$:

$$f(\bar{x}') - f(\bar{x}) \leq \langle \bar{x}' - \bar{x}, \bar{y} \rangle .$$

f is differentiable at $\bar{x}$ if its super gradient at $\bar{x}$ contains a unique element.

Lemma 3. *Let $D \subset \mathcal{X}$ be the set of sequences at which f is differentiable. Then D is a dense subset of $\mathcal{X}$.*

Proof. See (Asplund, 1968), Theorem 2. □

Lemma 4. *$\bar{x}$ maximizes f over $\mathcal{X}$ if and only if 0 belongs to $\partial f(\bar{x})$.*

Proof. Immediate from the definition of $\partial f(\bar{x})$. □

The following lemma will be useful in the derivation of each of the three statements internalizing the respective externalities in Theorem 4. As before, z denotes the state. The interpretation of the variables x, θ will change based on each externality considered; for instance, in the case of carrier thin/thick market externalities, x corresponds to s, while θ corresponds to e, G. The function f summarizes all constraints, H defines the state dynamics and u the welfare.

Lemma 5. *Let $L, M, N > 0$ and Z, X, Θ, be compact and convex subsets of $\mathbb{R}^L$, $\mathbb{R}^M$ and $\mathbb{R}^N$, respectively, $u : X \times \Theta \times Z \to \mathbb{R}$ be a concave and continuously differentiable function, $H : X \times \Theta \times Z \to Z$ be a linear function, and for each $k = 1, ..., K > 0$, let $f_k(x, \theta, z)$ be a continuously differentiable and concave* *function. Let* $z^0 \in Z$*, and* $\mathbf{\Theta} \subseteq \Theta^{\mathbb{N}\cup\{0\}}$ *be such that for every* $\bar{\theta} \in \mathbf{\Theta}$ *problem*

$$P\left(\bar{\theta}\right) : \max_{x \in X^{\mathbb{N}\cup\{0\}}} \sum_{t=0}^{\infty} \beta^t u\left(x^t, \theta^t, z^t\right)$$

$$\forall t, k : f_k\left(x^t, \theta^t, z^t\right) \geq 0$$

$$\forall t : z^{t+1} = H\left(x^t, \theta^t, z^t\right)$$

is feasible, and let $V\left(\bar{\theta}\right)$ *denote its value. Then* V *is concave. Moreover, suppose that* $\bar{\theta} \in \mathbf{\Theta}$ *is such that* $P\left(\bar{\theta}\right)$ *admits a unique optimal solution, and let* $\bar{x} \in X^{\mathbb{N}\cup\{0\}}$*,* $\bar{\lambda} \in \left(\mathbb{R}^K\right)^{\mathbb{N}\cup\{0\}}$ *and* $\bar{\phi} \in \left(\mathbb{R}^L\right)^{\mathbb{N}\cup\{0\}}$ *be such that, for every* t, k, l, m*:*

$$\lambda_k^t \geq 0 \text{ with equality if } f_k\left(x^t, \theta^t, z^t\right) > 0 \tag{77}$$

$$\frac{\partial u\left(x^t, \theta^t, z^t\right)}{\partial x_m} + \sum_k \lambda_k^t \frac{\partial f_k\left(x^t, \theta^t, z^t\right)}{\partial x_m} + \beta \sum_l \frac{\partial H_l\left(x^t, \theta^t, z^t\right)}{\partial x_m} \phi_l^{t+1} = 0 \tag{78}$$

where the sequence $\bar{\phi}$ *is defined recursively by:*

$$\phi_l^t = \frac{\partial u\left(x^t, \theta^t, z^t\right)}{\partial z_l} + \sum_k \lambda_k^t \frac{\partial f_k\left(x^t, \theta^t, z^t\right)}{\partial z_l} + \beta \sum_{l'} \frac{\partial H_{l'}\left(x^t, \theta^t, z^t\right)}{\partial z_l} \phi_{l'}^{t+1} \tag{79}$$

$$\lim_{t \to \infty} \beta^t \phi_l^t = 0$$

and the sequence $\bar{z}$ *is defined recursively by*

$$\forall t : z^{t+1} = H\left(x^t, \theta^t, z^t\right).$$

Then $\bar{\theta}$ *maximizes* V *over* $\mathbf{\Theta}$ *if*

$$\forall t, n : \frac{\partial u\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} + \sum_k \lambda_k^t \frac{\partial f_k\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} + \sum_l \frac{\partial H_l\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} \phi_l^{t+1} = 0 \tag{80}$$

and the above condition is also necessary whenever V *is differentiable at* $\bar{\theta}$*.*

Proof. Concavity of V can be proved using an argument analogous to Section A.3. For a generic sequence $\bar{a} = \left(a^t\right)_{t=0}^{\infty}$ and for every $T > 0$, we use the notation $\bar{a}^T \equiv \left(a^t\right)_{t=0}^{T}$ to denote the truncation of $\bar{a}$ at T.

When dealing with a sequence $\left(\bar{\theta}, \bar{x}\right)$ and an initial state z^0, unless stated otherwise, it is understood that $\bar{z}$ refers to the sequence of states induced by $\left(\bar{\theta}, \bar{x}\right)$ and the map H from z^0:

$$\forall t \geq 0 : \, z^{t+1} = H\left(x^t, \theta^t, z^t\right).$$

Let $\bar{\theta}, \bar{x}, \bar{\lambda}, \bar{\phi}$ be as in the statement. For every $T > 0$ consider the finite horizon problem,

$$P\left(T, \bar{\theta}\right) : \, V^T\left(\bar{\theta}^T\right) = \max_{\bar{x}^T \in X^{T+1}} \sum_{t=0}^{T} \beta^t u\left(x^t, \theta^t, z'^t\right) + \beta^{T+1} \sum_l z_l^{T+1} \phi_l^{T+1}$$
$$\text{s.t. } \forall t = 0, ..., T : \forall k : \, f_k\left(x^t, \theta^t, z^t\right) \geq 0.$$

By standard convex optimization theory, Conditions (77), (79) and (78) imply that $\left(\bar{x}^T, \bar{\lambda}^T\right)$ is an optimal dual pair for Problem $P\left(T, \bar{\theta}\right)$. Hence for every feasible sequence $\bar{x}'$ and for every $T > 0$ we have

$$\sum_{t=0}^{T} \beta^t u\left(x^t, \theta^t, z^t\right) + \beta^{T+1} \sum_l z_l^{T+1} \phi_l^{T+1} \geq \sum_{t=0}^{T} \beta^t u\left(x'^t, \theta^t, z'^t\right) + \beta^{T+1} \sum_l z_l'^{T+1} \phi_l^{T+1}$$

Sine Z and u are bounded[53], taking limits on both sides implies that $\bar{x}$ is optimal for $P\left(\bar{\theta}\right)$. Hence by our assumptions it must be the unique optimal solution for $P\left(\bar{\theta}\right)$. Define

$$\forall t, n : \, y_n^t = \frac{\partial u\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} + \sum_k \lambda_k^t \frac{\partial f_k\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} + \sum_l \frac{\partial H_l\left(x^t, \theta^t, z^t\right)}{\partial \theta_n} \phi_l^{t+1}.$$

We show that $\bar{y} \in \partial V\left(\bar{\theta}\right)$. From Marimon and Werner (2019) it follows that $\bar{y}^T \in \partial V\left(\bar{\theta}\right)$ for all $T > 0$:

$$\forall \bar{\theta}' \in \mathbf{\Theta} : \, V^T\left(\bar{\theta}'\right) - V^T\left(\bar{\theta}\right) \leq \sum_{t=0}^{T} \beta^t \sum_n y_n^t \left(\theta_n'^t - \theta_n^t\right)$$

Pick $\bar{\theta}' \in \Theta$ and let $\bar{x}'$ be an optimal solution for $P\left(\bar{\theta}'\right)$. For each T we have

$$\sum_{t=0}^{T} \beta^t u\left(x'^t, \theta^t, z'^t\right) + \beta^{T+1} \sum_l z_l'^{T+1} \phi_l^{T+1} \leq V^T\left(\bar{\theta}'\right)$$

[53] u is bounded, being a continuous function on a compact space.

and

$$V^T\left(\bar{\theta}\right)=\sum_{t=0}^{T}\beta^t u\left(x^t,\theta^t,z^t\right)+\beta^{T+1}\sum_l z_l^{T+1}\phi_l^{T+1}$$

hence

$$\sum_{t=0}^{T}\beta^t u\left(x'^t,\theta^t,z'^t\right)-\sum_{t=0}^{T}\beta^t u\left(x^t,\theta^t,z^t\right)+\beta^{T+1}\sum_l\left(z_l'^{T+1}-z_l^{T+1}\right)\phi_l^{T+1}\leq\sum_{t=0}^{T}\beta^t\sum_n y_n^t\left(\theta'^t-\theta_n^t\right)$$

Taking limits of both sides we get $V\left(\bar{\theta}'\right)-V\left(\bar{\theta}\right)\leq\sum_{t=0}^{\infty}\beta^t\sum_n y_n^t\left(\theta_n'^t-\theta_n^t\right)$. Since $\bar{\theta}'$ was arbitrary, this implies $y\in\partial V\left(\bar{\theta}\right)$. Hence, by Lemma 4, $\bar{\theta}$ maximizes V over $\mathbf{\Theta}$ if $\bar{y}=0$, and this condition is also necessary whenever V is differentiable at $\bar{\theta}$. This completes the proof. □

E.3.2 Proof of main result

This subsection is devoted to the proof of Theorem 4. We first establish two auxiliary lemmas.

Lemma 6. *The function f defined in equation (74) is continuously differentiable. Moreover, given a vector of choice probabilities $p\in\Delta I$, a vector $\phi\in\mathbb{R}^I$ and a scalar $\psi\in\mathbb{R}$, the following are equivalent:*

(i)

$$\psi=E_{\mathbb{P}}\max\left(\phi_j+\epsilon_j\right)\;\text{ and }\forall j:\;p_j=\mathbb{P}\left[\phi_j+\epsilon_j=\max_k\left(\phi_k+\epsilon_k\right)\right].$$

(ii)

$$\forall j:\;f\left(p\right)+\frac{\partial f\left(p\right)}{\partial p_j}-\sum_k p_k\frac{\partial f\left(p\right)}{\partial p_k}+\phi_j-\psi=0$$

Proof. It is well known (Galichon, 2018) that

$$\forall p\in\Delta I:\;-f\left(p\right)=\min_{\phi}\left[\sum_j p_j\phi_j-\mathrm{E}_{\mathbb{P}}\max_j\left(\phi_j+\epsilon_j\right)\right].$$

By our assumptions on $\mathbb{P}$, the objective function of the above problem is continuous and strictly convex, hence the set of its solutions is a singleton. By the envelope theorem, this implies that, $\partial f\left(p\right)=\{-\phi^*\left(p\right)\}$, where $\phi^*\left(p\right)$ is the unique optimal solution of the above problem. Hence f is differentiable. Moreover, continuity of ∂f follows by noting that ϕ^* is continuous by the Maximum Theorem.

For the second part of the statement, it is known (see Galichon, 2018) that (i) is equivalent to

$$\phi \in \partial \left(-f\left(p\right)\right) \text{ and } -f\left(p\right)+\psi=\sum_{j} p_j \phi_j .$$

When f is differentiable, the condition above is equivalent to (ii). This completes the proof. □

Lemma 7. *Let $\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}, \bar{\psi}, \bar{\mu}^e, \bar{\mu}^s, \bar{\Delta}, \bar{\phi}^s, \bar{\phi}^e$ be such that*

$$\lim_{t\to\infty} \beta^t \phi_{ij}^{s,t} = \lim_{t\to\infty} \beta^t \phi_{ij}^{e,t} = 0,$$

$\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right) \in \mathcal{A}\left(z^0\right)$ *and, for every t, i, j, $s_i^t, e_i^t > 0$ and the following conditions hold:*

$$\mu_i^{s,t} \geq 0 \textit{ with equality if } s_i^t < x_{ii}^t$$

$$\mu_{ij}^{e,t} \geq 0 \textit{ with equality if } e_i^t G_{ij}^t > y_{ij}^t$$

$$\Delta_{ij}^t \geq 0 \textit{ with equality if } q_{ij}^t < m_i\left(s_i^t, e_i^t\right) G_{ij}^t$$

$$\Delta_{ij}^t \geq w_{ij}\left(q^t\right) + \phi_{ij}^{s,t} - \beta \mu_{ij}^{e,t+1} - \psi_i^t, \textit{with equality if } q_{ij}^t > 0$$

$$\psi_i^t = E_{\mathbb{P}} \max_j \left(\phi_{ij}^{s,t} + \epsilon_{ij}\right)$$

$$P_{ij}^{b^t} = \mathbb{P}\left[\phi_{ij}^{s,t} = \max_k \left(\phi_{ik}^{s,t} + \epsilon_{ik}\right)\right]$$

$$\phi_{ii}^{s,t} = \mu_i^{s,t} + \beta \phi_{ii}^{s,t+1}$$

$$\phi_{ij}^{e,t} = \kappa_{ij} - \mu_{ij}^{e,t}$$

$$\phi_{ij}^{s,t} = -c_{ij}^s + \beta\left[d_{ij} \phi_{jj}^{s,t+1} + \left(1-d_{ij}\right) \phi_{ij}^{s,t+1}\right].$$

Then:

(i) $\bar{s}$ maximizes the function $\bar{s}' \mapsto V\left(\bar{s}', \bar{e}, \bar{G}, z^0\right)$ over $\mathcal{SA}\left(z^0|\bar{e}, \bar{G}\right)$ if, for every i, t:

$$- c_i^s + \frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial s_i} \sum_j G_{ij}^t \Delta_{ij}^t + \psi_i^t - \beta \phi_{ii}^{s,t+1} - \mu_i^{s,t} = 0. \tag{81}$$

This condition is also necessary whenever the function $\bar{s}' \mapsto V\left(\bar{s}', \bar{e}, \bar{G}, z^0\right)$ is differentiable at $\bar{s}$.

(ii) $\bar{e}$ maximizes the function $\bar{e}' \mapsto V\left(\bar{s}, \bar{e}', \bar{G}, z^0\right)$ over $\mathcal{SA}\left(z^0|\bar{s}, \bar{G}\right)$ if, for every i, t:

$$\frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial e_i} \sum_j G_{ij}^t \Delta_{ij}^t - \sum_j G_{ij}^t \left(c_{ij}^e + \kappa_{ij} - \mu_{ij}^{e,t} - \phi_{ij}^{e,t+1}\right) = 0. \tag{82}$$

This condition is also necessary whenever the function $\bar{e}' \mapsto V\left(\bar{s}, \bar{e}', \bar{G}, z^0\right)$ is differentiable at $\bar{e}$.

(iii) $\bar{G}$ maximizes the function $\bar{G}' \mapsto V\left(\bar{s}, \bar{e}, \bar{G}', z^0\right)$ over $\mathcal{SA}\left(z^0|\bar{s}, \bar{e}\right)$ if there exists a sequence $\bar{\omega}$ such that, for every i, t:

$$m_i\left(s_i^t, e_i^t\right) \Delta_{ij}^t - e_i^t \left(c_{ij}^e + \kappa_{ij} - \mu_{ij}^{e,t} - \phi_{ij}^{e,t+1}\right) \leq \omega_i^t \tag{83}$$
$$\text{with equality if } G_{ij}^t > 0.$$

This condition is also necessary whenever the function $\bar{G}' \mapsto V\left(\bar{s}, \bar{e}, \bar{G}', z^0\right)$ is differentiable at $\bar{G}$.

Proof. We apply Lemma 5 to Problem (76)

$$P\left(\bar{s}, \bar{e}, \bar{G}\right) : V^p\left(\bar{s}, \bar{e}, \bar{G}, z^0\right) = \max_{\bar{q}, \bar{b}} \sum_{t=0}^{\infty} \beta^t W^p\left(s^t, e^t, G^t, q^t, b^t; z^t\right)$$
$$\text{s.t. } \left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right) \in \mathcal{A}\left(z^0\right).$$

In doing so, notice that the assumptions of Lemma 5 are satisfied, since by Lemma 6 the function W^p is continuously differentiable, and we can take feasible allocations and states to live inside a compact set.[54]

We use the following notation for the Lagrangian multipliers:

[54] Indeed, let $M = \sum_{ij} x_{ij}^0$. Then for every $\left(\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}\right) \in \mathcal{A}\left(z^0\right)$ we must have

$$\forall t, i, j : 0 \leq s_i^t, q_{ij}^t, b_{ij}^t, x_{ij}^t \leq M.$$

multiplier	constraint
$\mu_{ij}^{e,t}$	$e_i^t G_{ij}^t \geq y_{ij}^t$
$\mu_i^{s,t}$	$x_{ii}^t \geq s_i^t$
ψ_i^t	$s_i^t = \sum_j \left(q_{ij}^t + b_{ij}^t\right)$
Δ_{ij}^t	$q_{ij}^t \leq m_i\left(s_i^t, e_i^t\right) G_{ij}^t$
$-\omega_i^t$	$\sum_j G_{ij}^t = 1$

Moreover, we denote $\bar{\phi} = \left(\bar{\phi}^s, \bar{\phi}^e\right)$, where $\bar{\phi}^s$ is the component of $\bar{\phi}$ associated with the supply component of the state and $\bar{\phi}^e$ is associated with the demand component. With this notation in hand, the set of Conditions (77) is given by

$$\begin{aligned}\forall t, i, j : \mu_i^{s,t} &\geq 0 \text{ with equality if } s_i^t < x_{ii}^t \\ \mu_{ij}^{e,t} &\geq 0 \text{ with equality if } e_i^t G_{ij}^t > y_{ij}^t \\ \Delta_{ij}^t &\geq 0 \text{ with equality if } q_{ij}^t < m_i\left(s_i^t, e_i^t\right) G_{ij}^t,\end{aligned}$$

the set of Conditions (79) is given by

$$\begin{aligned}\forall t, i, j : \phi_{ii}^{s,t} &= \mu_i^{s,t} + \beta \phi_{ii}^{s,t+1} \\ \phi_{ij}^{e,t} &= \kappa_{ij} - \mu_{ij}^{e,t} \\ \phi_{ij}^{s,t} &= -c_{ij}^s + \beta\left[d_{ij}\phi_{jj}^{s,t+1} + \left(1 - d_{ij}\right)\phi_{ij}^{s,t+1}\right]\end{aligned}$$

Moreover, letting e^*, q^*, b^* be a solution of

$$\begin{aligned}\max_{e,q,b\geq 0} W(q) - \sum_i e_i &\left(\min_j c_{ij}^e + \min_j \kappa_{ij}\right) \\ \text{s.t.} q_{ij} &\leq M \\ \sum_j q_{ij} &\leq m_i(M, e_i)\end{aligned}$$

we have that every sequence $\bar{e}$ such that $e_i^t > e_i^*$ for some t, i is clearly sub optimal, hence without loss of generality we can take

$$0 \leq e_i^t, y_{ij}^t \leq e_i^*.$$

and the set of Conditions (78) is given by

$$\forall t, i, j : \Delta_{ij}^t \geq w_{ij}\left(q^t\right) + \phi_{ij}^{s,t} - \beta\mu_{ij}^{e,t+1} - \psi_i^t \text{ with equality if } q_{ij}^t > 0$$

$$f\left(P_i^b\right) + \frac{\partial f\left(P_i^b\right)}{\partial P_{ij}} - \sum_k P_{ik}^b \frac{\partial f\left(P_i^b\right)}{\partial P_{ik}} + \phi_{ij}^{s,t} - \psi_i^t = 0.$$

By Lemma 6, the set of conditions in the second line above is equivalent to

$$\forall t, i, j:\ P_{ij}^{b^t} = \mathbb{P}\left[\phi_{ij}^{s,t} = \max_k\left(\phi_{ik}^{s,t} + \epsilon_{ik}\right)\right] \text{ and } \psi_i^t = \mathrm{E}_{\mathbb{P}} \max_j\left(\phi_{ij}^{s,t} + \epsilon_{ij}\right).$$

To prove Statement (i), we apply Lemma 5 to the function $\bar{s}' \mapsto V\left(\bar{s}', \bar{e}, \bar{G}, z^0\right)$. Given our assumption that $s_i^t > 0$ for every t, i, Condition (80) is given by

$$\forall i, t:\ -c_i^s + \frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial s_i} \sum_j G_{ij}^t \Delta_{ij}^t + \psi_i^t - \beta\phi_{ii}^{s,t+1} - \mu_i^{s,t} = 0.$$

To prove Statement (ii), we apply Lemma 5 to the function $\bar{e}' \mapsto V\left(\bar{s}, \bar{e}', \bar{G}, z^0\right)$. Given our assumption that $e_i^t > 0$ for every t, i, Condition (80) is given by

$$\forall i, t:\ \frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial e_i} \sum_j G_{ij}^t \Delta_{ij}^t - \sum_j G_{ij}^t\left(c_{ij}^e + \kappa_{ij} - \mu_{ij}^{e,t} - \beta\phi_{ij}^{e,t+1}\right) = 0.$$

To prove Statement (iii), we apply Lemma 5 to the function $\bar{G}' \mapsto V\left(\bar{s}, \bar{e}, \bar{G}', z^0\right)$. Condition (80) is given by

$$\forall i, j, t:\ m_i\left(s_i^t, e_i^t\right)\Delta_{ij}^t - e_i^t\left(c_{ij}^e + \kappa_{ij} - \mu_{ij}^{e,t} - \beta\phi_{ij}^{e,t+1}\right) - \omega_i^t + \mathrm{pos}_{ij}^t = 0$$

where pos_{ij}^t is the multiplier associated to the positivity constraint $G_{ij} \geq 0$, which must satisfy

$$\mathrm{pos}_{ij}^t \geq 0 \text{ with equality if } G_{ij}^t > 0.$$

This completes the proof. □

Proof of main result In order to prove the main result, let everything be as in the statement. Let $(V^{s,t}, U^{e,t}, \Delta^{s,t}, \Delta^{e,t})_{t=0}^{\infty}$ be the sequence of carriers and customers' value functions and meeting surpluses associated with the sequence $\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}$ evaluated at the state trajectory z^t, $t >= 0$. For every $t \geq 0$ define $\phi^{s,t} = V^{s,t}$, $\phi^{e,t} = U^{e,t}$, $\psi^t = \mathrm{E}_{\mathbb{P}} U^{s,t}(\epsilon)$, $\Delta^t = \Delta^{s,t} + \Delta^{e,t}$ and

$$\mu_i^{s,t} = \max\left\{-c_i^s + \boldsymbol{\lambda}_i^s\left(z^t\right)\sum_{j\neq i} G_{ij}^t \Delta_{ij}^{s,t} + U_i^{s,t} - \beta V_i^{s,t+1}, 0\right\}$$

$$\mu_{ij}^{e,t} = \kappa_{ij} - U_{ij}^{e,t}.$$

Then $\bar{s}, \bar{e}, \bar{G}, \bar{q}, \bar{b}, \bar{\psi}, \bar{\mu}^e, \bar{\mu}^s, \bar{\Delta}, \bar{\phi}^s, \bar{\phi}^e$ satisfies the conditions of Lemma 7. Moreover, notice that:

- Condition (81) can be written as

$$\forall i, t : \frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial s_i} \sum_j G_{ij}^t \left(\Delta_{ij}^{s,t} + \Delta_{ij}^{e,t}\right) - \boldsymbol{\lambda}_i^s\left(z^t\right) \sum_{j\neq i} G_{ij}^t \Delta_{ij}^{s,t} = 0.$$

Using $\boldsymbol{\lambda}_i^s\left(z^t\right) = m_i\left(s_i^t, e_i^t\right)/s_i^t$ and rearranging, this is equivalent to

$$\eta_i^s\left(z^t\right) \sum_j G_{ij}^t \left(\Delta_{ij}^{s,t} + \Delta_{ij}^{e,t}\right) = \sum_j G_{ij}^t \Delta_{ij}^{s,t}.$$

- Condition (82) can be written as

$$\forall i, t : \frac{\partial m_i\left(s_i^t, e_i^t\right)}{\partial e_i} \sum_j G_{ij}^t \left(\Delta_{ij}^{s,t} + \Delta_{ij}^{e,t}\right) - \sum_j G_{ij}^t \left(c_{ij}^e + U_{ij}^{e,t} - \beta U_{ij}^{e,t+1}\right) = 0.$$

Using $U_{ij}^{e,t} = -c_{ij}^e + \boldsymbol{\lambda}_i^e\left(z^t\right)\Delta_{ij}^{e,t} + \beta U_{ij}^{e,t+1}$, $\boldsymbol{\lambda}_i^e\left(z^t\right) = m_i\left(s_i^t, e_i^t\right)/e_i^t$ and rearranging, this is equivalent to

$$\eta_i^e\left(z^t\right) \sum_j G_{ij}^t \left(\Delta_{ij}^{s,t} + \Delta_{ij}^{e,t}\right) = \sum_j G_{ij}^t \Delta_{ij}^{e,t}.$$

- Condition (83) can be written as

$$\forall i, j, t : m_i\left(s_i^t, e_i^t\right)\Delta_{ij}^t - e_i^t\left(c_{ij}^e + U_{ij}^{e,t} - \beta U_{ij}^{e,t+1}\right) \leq \omega_i^t$$

with equality if $G_{ij}^t > 0$.

Using $U_{ij}^{e,t} = -c_{ij}^{e} + \boldsymbol{\lambda}_i^e\left(z^t\right)\Delta_{ij}^{e,t} + \beta U_{ij}^{e,t+1}$, $\boldsymbol{\lambda}_i^e\left(z^t\right) = \frac{m_i\left(s_i^t, e_i^t\right)}{e_i}$ and rearranging, this is equivalent to

$$\forall i, j, t : \Delta_{ij}^{s,t} \leq -\frac{\omega_i^t}{\boldsymbol{\lambda}_i^e\left(z^t\right)}.$$

with equality if $G_{ij}^t > 0$.

This completes the proof.