WORKING PAPERS

N° 1207

April 2023

"Market Structure, Investment and Technical Efficiencies in Mobile Telecommunications"

Jonathan Elliott, Georges V. Houngbonon Marc Ivaldi and Paul Scott

Toulouse School of Economics

Market Structure, Investment, and Technical Efficiencies in Mobile Telecommunications*

Jonathan T. Elliott,† Georges V. Houngbonon,‡
Marc Ivaldi,§ Paul T. Scott¶

April 25, 2023

Abstract

We develop a model of competition in prices and infrastructure among mobile network operators. Consolidation can increase market power, but economies of scale, which we derive from physical principles, lead to more efficient data transmission. Estimating our model with French consumer and infrastructure data, we find that while prices decrease with more firms, so do download speeds. Consumer surplus is maximized at six firms, but fewer firms would improve total surplus and welfare for high-income consumers. We also use our framework to assess the impact of another aspect of market structure: the allocation of spectrum.

Keywords: Market structure, scale efficiency, antitrust policy, infrastructure, endogenous quality, queuing, mobile telecommunications.

JEL Classification: D21, D22, L13, L40.

*This research has benefited from the support of Orange (through the TSE-Partnership research foundation) and the Agence Nationale de la Recherche grant ANR-17-EURE-0010. It has also benefited from the support of NYU IT High Performance Computing resources, services, and staff expertise. We thank Richard Chen, Stéphane Ciriani, Chris Conlon, Ying Fan, Fraida Fund, Mohamed Karray, Jakub Kastl, Michael Knox, Elena Krasnokutskaya, Paul LaFontaine, Marc Lebourges, Julienne Liang, Alessandro Lizzeri, Chitra Marti, Thomas Marzetta, Shivendra Panwar, Nancy Rose, Greg Rosston, Maher Said, David Salant, Jan Philip Schain, David Sibley, Alan Sorenson, Patrick Sun, Dan Waldinger, Aleks Yankelevich, and Ali Yurukoglu for insightful comments, discussions, and suggestions. The usual disclaimer applies. All estimation and simulation code (in Python) can be found at https://github.com/jonathantelliott/mobile-telecommunications.

†Johns Hopkins University, jonathan.elliott@jhu.edu

‡IFC-World Bank Group, ghoungbonon@ifc.org, gvivienh@gmail.com

§Toulouse School of Economics, marc.ivaldi@tse-fr.eu

¶New York University, ptscott@stern.nyu.edu

1 Introduction

In the mobile telecommunications industry, market structure is shaped by antitrust policy and the regulation of radio frequencies, or spectrum. Spectrum is necessary for the operation of a mobile network, and a firm's spectrum holdings (the set of frequencies it has the right to operate) impacts its quality of service (i.e., download speeds). Recently, mobile network operators in many countries have sought to merge and combine their spectrum holdings, with mixed responses from regulators concerned about increased concentration.[1] In recent discussions regarding both antitrust policy and spectrum allocation, quality of service has been a prominent concern.[2]

In this paper, we develop a structural model of the mobile telecommunications industry to capture the impact of changes in market structure (the number of network operators and the allocation of spectrum among them) on equilibrium outcomes such as prices, investment, download speeds, and welfare. The model allows us to assess the trade-off between market power and economies of scale, both in the traditional sense, where consolidation may result in higher or lower prices (Williamson, 1968), and in the sense that consolidation affects quality of service. As our notion of market structure includes not only the number of firms but also their spectrum holdings, our framework also allows us to consider the impact of changes in the allocation of spectrum to and within the industry.

Our structural model comprises firms, consumers, and data transmission. Firms (mobile network operators) choose the prices of their mobile service plans and their level of investment in infrastructure, which consumers rely on for data consumption. Consumers choose a mobile phone plan, as well as how much data to consume using that plan, given the download speeds associated with the plan. Our model of data transmission describes how download speeds emerge from firms' and consumers' decisions.

Download speeds, arguably the crucial measure of quality of service in this context, present two modeling challenges. First, due to congestion, download speeds depend on consumers' data

[1] Approved mergers include T-Mobile/Orange (UK, 2010), Hutchinson/VimpelCom (Italy, 2016), Sprint/T-Mobile (USA, 2020), and Teléfonica/Virgin (UK, 2020). Blocked mergers include AT&T/T-Mobile (USA, 2011), TeliaSonera/Telenor (Denmark, 2015), and Teléfonica/Hutchinson (UK, 2016). Anecdotally, network operators in some countries (e.g., France) have recently avoided proposing four-to-three mergers due to an expectation that they would be blocked by antitrust authorities.

[2] For instance, the Sprint/T-Mobile merger was allowed based on the finding "that quality benefits and dynamic competition serve as countervailing forces to the static analysis that substantially address its predicted harmful price effects" (Federal Communications Commission, 2019). Genakos, Valletti and Verboven (2018) study how concentration in mobile telecommunications is related to both prices and investment in infrastructure. Turning to spectrum allocation, the Federal Communications Commission's "National Broadband Plan" describes the potential consequences of insufficient spectrum allocation to mobile telecommunications: "higher prices, poor service quality, an inability for the U.S. to compete internationally, depressed demand and, ultimately, a drag on innovation" (Federal Communications Commission, 2010).

consumption decisions as well as firms' investments. Second, even ignoring congestion, there isn't a simple mapping from firm's investments to data transmission rates, as data transmission depends, among other things, on spectrum operated and the distance over which data is transmitted. We model download speeds based on engineering models of data transmission that capture how data is transmitted across space and how network load is handled (in particular, Błaszczyszyn, Jovanovicy and Karray, 2014).[3] These engineering relationships imply two types of economies of scale that have important economic implications, which we call *economies of density* and *economies of pooling*.

Economies of density result from path loss: as electromagnetic waves carrying data travel, they lose power. Therefore, a mobile network operator can serve a densely populated area more efficiently (meaning either a higher download speed at a given cost or the same download speed at a lower cost) than a sparsely populated area.[4] With symmetric firms, the population density served by each firm is inversely proportional to the number of firms. Consequently, for a given level of total investment in the industry, mobile data services are higher quality when the number of firms is small.[5]

Economies of pooling result from mobile network congestion. When many consumers request data at the same time, data requests enter a queue. Longer queues result in slower download speeds, and there are economies of scale in serving queues. For example, if two network operators were to combine both their customer bases and owned spectrum, the combined firm could more efficiently allocate network capacity among customers, thereby reducing congestion, resulting in higher average download speeds. More generally, the allocation of resources serving a stochastic demand process leads to economies of scale (Mulligan, 1983; De Vany, 1976; Carlton, 1978).

We estimate a model of demand for mobile plans and data consumption based on the French market in 2015. Our estimation relies on a unique data set from the French mobile market, with data on choices and consumption by nearly 15 million customers in October 2015 from a single mobile network operator, Orange Mobile.[6] We also incorporate measured download speeds from Ookla, detailed (publicly available) data on mobile network infrastructure from

[3] As we derive our production function and associated scale efficiencies transparently from physical principles, our study falls within the tradition of engineering production functions of Chenery (1949).

[4] For example, suppose that the number of base stations per person is held constant across different population densities, so that less population-dense areas have lower geographic base station density. Because signals in the sparsely populated areas will have to travel further on average, they will experience greater path loss, and sparsely populated areas will have inferior service despite receiving the same level of investment per capita.

[5] Of course, the equilibrium level of investment (per firm and in total) may change with the number of firms. Our model will allow for such changes endogenously, with firms choosing infrastructure investment strategically.

[6] In accordance with data protection and privacy concerns, we were provided with commune-level statistics about consumer choices and consumption rather than consumer-level data directly.

the radio frequency regulator (ANFR), and income distribution data from the French statistical office (INSEE). While we only observe consumers who subscribe to Orange Mobile, we observe the prices and characteristics for all contracts available in the market, and we prove that the estimation strategy of Berry, Levinsohn and Pakes (1995) can be employed in this setting.[7]

While our model of the supply side is mostly derived from engineering models, we recover a small number of cost parameters from firms' first-order conditions. Intuitively, once we have estimated demand, we can quantify marginal revenue. We can then use firms's first-order conditions and our understanding of marginal revenue to make inferences about firms' costs. Firms' pricing decisions provide information about their costs per user served. Furthermore, firms' infrastructure investment decisions (in particular the choice of how densely to build base stations) provide information about the costs of building base stations.

We use the estimated models of demand and supply to compute counterfactual equilibria under different numbers of firms. Consolidation presents a trade-off for consumers: faster downloads at the cost of higher prices. We find that consumer surplus is maximized at six firms, but low-income consumers prefer a market with more firms than do high income consumers, who have a higher willingness to pay for increased download speeds. Total surplus is maximized at three firms.

We also explore the marginal social value of allocating more spectrum to the mobile telecommunications industry and compare this value with an individual firm's willingness to pay for a marginal unit of spectrum. We find that the marginal social value is about five times greater than an individual firm's willingness to pay.[8] This result highlights limitations with using the results of spectrum auctions to guide high-level spectrum allocation decisions, such as how to allocate spectrum among different sectors. While spectrum auctions may reveal network operators' willingness to pay, willingness to pay may be a gross underestimate of spectrum's social value in mobile telecommunications. Thus, when deciding how much spectrum to allocate to mobile telecommunications, a structural model like ours may prove invaluable to regulators (although quantifying the opportunity cost of spectrum—its value when allocated to another sector—is beyond our scope).[9]

Our model is also well suited to addressing questions of within-industry spectrum allocation. Inspired by the entry of Free Mobile in 2012 in France, we consider two ways in which a

[7]Our model predicts shares for all products from all providers in the market, but we only require that the model rationalize product-level market shares for Orange. For other firms, we impose firm-level demand shocks and require the model to rationalize firm-level market shares. Chu (2010) uses a similar approach.

[8]Rosston (2003) found social value to be more than ten times firm willingness to pay.

[9]Our model takes spectrum allocation as given. Thus, while our framework allow us to quantify the impact of spectrum allocation on outcomes, we abstract away from concerns about the spectrum allocation mechanism, e.g. Milgrom and Segal (2020) and Doraszelski et al. (2019).

regulator might allocate more spectrum to mobile telecommunications: by giving it to a new entrant (inducing entry), or distributing it among incumbents. We find that the former is better for consumer surplus, but the latter is better for total surplus.

Related Literature Most theoretical studies on the relationship between competition and investment take total industry-wide investment as the outcome of interest (Arrow, 1962; Vives, 2008). However, the operation of mobile telecommunications networks features important sources of economies of scale, which introduces a potential wedge between industry-wide investment and industry performance. Even if total investment increases with the total number of firms, quality of service may decline as network resources are spread more thinly across firms. By augmenting a model of investment in infrastructure with an engineering-based model of data transmission, we can directly quantify these scale economies.

While our analysis assesses the impact of market structure on prices and quality of service, market structure in mobile telecommunications has broader potential impacts: on product proliferation and the types of contracts offered (Seim and Viard, 2011; Fan and Yang, 2020), on coordinated effects (Bourreau, Sun and Verboven, 2021), and on incentives to engage in vertical restrictions (Sinkinson, 2020).

A few papers also study investment in mobile telecommunications infrastructure. Granja (2022) studies investment decisions under universal service regulation in Brazil. Lin, Tang and Xiao (2022) analyze 4G technology investment under a hypothetical merger, finding that the merger would reduce investment in this technology. Grajek and Röller (2012) argue that the empirical evidence suggests that access regulation (forcing incumbents to share their infrastructure with entrants) reduces incentives to invest in telecommunications infrastructure. Björkegren (2022) also models endogenous investment in infrastructure, finding that adding a competitor increases investment in rural areas. Björkegren's setting is a less-developed country where geographic coverage is the key product characteristic affected by network operators' investments; ours is a developed country where we take full geographic coverage for granted, and quality of service is the key product characteristic.

There is a limited empirical literature studying imperfectly competitive markets in which firms optimally choose the quality of their products offered. In the seminal theory (Spence, 1975) and in well-studied empirical contexts such as newspapers (Fan, 2013) and cable television (Crawford and Shum, 2007; Chu, 2010; Crawford et al., 2018; Crawford, Shcherbakov and Shum, 2019), quality is a product characteristic that firms can directly control. However, in the context of mobile telecommunications, a challenge for accurately modeling quality of service is the simultaneous determination of download speeds and demand for data.

Consumer demand for a network operator's services depends on its quality of service, and its quality of service depends on consumer demand due to congestion externalities.[10] Most demand models for mobile services do not model the simultaneous determination of demand and quality of service (including Bourreau, Sun and Verboven (2021), Cullen, Schutz and Shcherbakov (2020), Fan and Yang (2020), Nevo, Turner and Williams (2016), Sinkinson (2020), Sun (2015)). Only El Azouzi, Altman and Wynter (2003) and Lhost, Pinto and Sibley (2015) model the simultaneous determination of service quality and choice of service provider using queuing theory like we do. Our study builds on these by incorporating path loss (and therefore economies of density) and by estimating a product-level demand model using detailed consumption and quality data (therefore allowing us to tackle questions of market power). Meanwhile, in the engineering literature, Hua, Liu and Panwar (2012) examine how integrating network resources benefits both from economies of density and pooling, but without an economic equilibrium framework that endogenizes consumers' choices and firms' investments.

Outline The remainder of this paper is organized as follows. Section 2 presents the data along with some descriptive statistics on usage and quality of mobile data. Section 3 presents the model of demand and infrastructural investment. Section 4 presents the estimation strategy, and Section 5 presents the results. Section 6 presents some counterfactual analyses.

2 Data and Background

2.1 Firms

We focus on the French telecommunications market in October 2015. During the period we study, the French mobile industry comprised four mobile network operators (MNOs): Orange (ORG), SFR-Numericable (SFR), Bouygues Telecom (BYT) and Free Mobile (FREE).

MNOs own and operate their network infrastructure (with some network sharing, which we will describe in Section 3.3). In contrast, mobile virtual network operators (MVNOs) sell plans to customers without owning their own network resources; instead, they rent access to MNOs' networks. Providing network access to MVNOs is mandatory and enforced by regulation, but the access charge is freely negotiated with the MNO. MVNOs accounted for 10.6% of the mobile contracts in late 2015 (ARCEP, 2016).

2.2 Products and Characteristics

We collect data on mobile phone plan terms (including monthly prices, data limits, voice limits) from online quarterly catalogs of offers proposed by the four MNOs and the largest

[10] Congestion externalities are negative network externalities. Related challenges arises in markets with positive network externalities; e.g., Lee (2013).

MVNO, EI Telecom. Here we describe how we interpret this catalog data at a high level; further details are available in Appendix C.1.

By 2015, wireless plans were largely differentiated based on data services, with more expensive plans coming with larger data allowances. Most plans featured unlimited voice allowances; only some low-end plans with zero or low data allowances had limited voice minutes. Furthermore, while data consumption was still growing rapidly through 2015, voice and text message consumption had stabilized.[11]

Table 1 describes our choice set, with monthly prices and data limits representing the main characteristics of interest. Monthly data limits are "soft," in the sense that customers can still use data services once the limit is exceeded, but download speeds will be throttled significantly.[12] We aggregate phone plans by data limit category (less than 500 MB, 500–2 999 MB, 3 000–6 999 MB, and more than 7 000 MB) and whether they include unlimited voice services. For each plan grouping and each firm, we choose a representative plan to include in our choice set.

Our representative plans do not feature bundled services like fixed broadband, fixed telephony, and television services. Moreover, plans have different contract lengths (no commitment, a 12 month commitment, or a 24 month commitment). Most consumers subscribe to plans with a contract length of 24 months, so our representative plans correspond to 24-month commitment plans. Within a group of a firm's plans, defined by data and voice limits, the representative plan we select is the one that is the least expensive (after adjusting the monthly price for a handset subsidy as described below) among the 24-month commitment plans. This plan is always one without home broadband and television services. Thus, our choice set of representative plans consists entirely of mobile-only plans.

The representative phone plans in our model's choice set have the characteristics of plans actually available in the market. The only characteristic that is adjusted is the monthly price. When a representative contract is associated with a handset subsidy, we use as the price the monthly price of the plan minus the value of that handset subsidy. See Appendix C.1.1 for details about how we calculate this subsidy.

In the customer database described below, we observe market shares for plans with only wireless services as well as plans with bundled services. Each actual plan in these data is then associated with a representative plan (by category defined by data limit and unlimited voice

[11] Source: Séries chronologiques annuelles (1998-2015) data released by ARCEP. Obtained from http://www.arcep.fr/fileadmin/reprise/observatoire/serie-chrono/series-chrono-annuelles-1998-2015p.xlsx September 23, 2022.

[12] The data allowances we measure are the baseline allowances associated with phone plans. We ignore add-on options.

dummy), and our estimation method takes the market shares of the representative plans to be the aggregate market share of all the actual products associated with them.[13] For instance, our empirical model features one high-data-limit plan for Orange. We treat the price of this plan as 38.74 €. This price corresponds to an observed price of 54.99 € for this plan and an adjustment of 16.25 € for the value of the associated handset subsidy. We measure the market share of this representative plan as the sum of market shares of eleven high-data-limit contracts offered by Orange.

Table 1: The Choice Set

Operator	**Price (€)**	**Data Limit (MB)**	**Unlimited Voice**	**Plans Represented**	**Min Price (€)**	**Max Price (€)**	**Min Limit (MB)**	**Max Limit (MB)**
Orange	12.07	50	No	11	4.99	30.99	0	50
Orange	14.99	1000	No	4	14.99	14.99	1 000	1 000
Orange	22.91	1000	Yes	2	22.91	24.99	1 000	1 000
Orange	30.91	4000	Yes	5	19.99	48.99	3 000	5 000
Orange	38.74	8000	Yes	11	38.74	165.99	8 000	20 000
Bouygues	8.07	0	No	6	3.99	11.32	0	20
Bouygues	14.99	1000	No	3	14.99	14.99	1 000	1 000
Bouygues	20.91	3000	Yes	4	19.99	29.99	3 000	5 000
Bouygues	33.74	10000	Yes	4	32.70	72.70	10 000	20 000
Free Mobile	2.00	50	No	1	2.00	2.00	50	50
Free Mobile	19.99	3000	Yes	1	19.99	19.99	3 000	3 000
SFR	12.07	100	No	5	5.99	14.99	100	200
SFR	14.99	1000	No	3	14.99	19.99	1 000	1 000
SFR	22.91	1000	Yes	3	22.91	29.99	1 000	1 000
SFR	31.91	5000	Yes	5	19.99	43.99	3 000	5 000
SFR	37.74	10000	Yes	9	36.70	149.99	10 000	20 000
MVNO	7.99	0	No	13	7.99	18.99	0	200
MVNO	17.99	1000	No	5	9.99	17.99	500	1 000
MVNO	19.99	500	Yes	10	19.99	35.99	500	2 000
MVNO	42.99	5000	Yes	13	12.99	61.99	3 000	5 000
MVNO	64.99	10000	Yes	4	64.99	76.99	10 000	10 000

Each row corresponds to an object in the choice set, i.e., a representative product. The minimum and maximum prices and data limits are over the set of plans represented by each representative product in the choice set. The minimum price can be greater than the price of the representative plan because the representative plan is the least expensive within the group of plans *that has a commitment of 24 months*. Because some plans have commitments of less than 24 months, it is possible for the minimum price to be lower than the representative plan's price.

We do not explicitly distinguish between pre- and postpaid phone plans. Most consumers subscribe to postpaid plans, which account for 83% of plans in late 2015 (ARCEP, 2016). While postpaid plans require consumers to pay for their consumption during a monthly billing period, prepaid customers require customers to pay as they go. Prepaid contracts generally

[13]The Orange customer database includes consumers on plans that are no longer available. These plans, like available plans, are all mapped to a representative plan and consumers subscribing to these plans will contribute to the market share of the associated representative plan.

involve low data limits and limited voice allowances.

For MVNOs, our choice set includes one representative plan for each category; that is, we effectively assume there is one representative MVNO firm. Representative MVNO contracts are selected from the contracts of the largest MVNO (EI Telecom) in the same way we select representative contracts for the MNOs.

2.3 Demand Data

Our main demand data source is a proprietary data set of 15 million residential mobile customers of one operator, Orange Mobile, in October 2015. This data set includes information on the phone plan subscribed to and the usage of mobile voice and data services. Note that we focus only on the residential market for mobile services, ignoring business customers. Residential customers represented 89% of the mobile market in 2015 (ARCEP, 2016).

The customer data set is complemented by data on the quality of mobile data services, as measured by download speeds. Due to congestion, delivered download speeds are not merely a function of infrastructure and geographic characteristics. Congestion arises because the available bandwidth is shared among users and, as a result, the greater the number of users, the lower the quality (as measured by download speed). At the same time, the number of users (and therefore the demand for data) on a network depends on quality. In our counterfactuals, we employ a model in which demand and quality of service are simultaneously determined, but for the purpose of estimation, we rely on a direct measure of download speeds as our measure of quality. Speedtest is a service offered by the firm Ookla that allows users to check their download and upload internet speeds. We use data from these speed tests that include measured download speed, the time of the test, the location of the user, and the mobile network operator. We use a proprietary data set provided by Ookla on over one million speed tests in France in the fourth quarter of 2015 to construct a measure of experienced download speeds for each mobile network operator in each municipality. Section C.3 in the data appendix explains the construction of this quality measure in detail.

Markets are defined as municipalities (French communes), and we limit our analysis to relatively populous markets, defined as those with a population greater than 10 000, for a total of 589 markets.[14] Municipality-level market size is defined as the population age 12 and older, obtained from the French Bureau of Statistics, INSEE. While France has about 36,000

[14]We limit ourselves to populous markets because active network sharing (where network operators share the transmitting components of their infrastructure) is relatively common in rural areas but not practiced in urban areas. Thus, for our sample, we are comfortable associating a firm's measured download speeds with that firm's own infrastructural investments. There are 592 municipalities with a population greater than 10 000, and we drop three of those municipalities due to insufficient download speed tests to construct quality measures. This yields a total of 589 markets in our sample.

municipalities, the 589 in our sample contain 43.5% of France's population.

For network operators other than Orange, we have only market shares at the national level from GSMA Intelligence. Table 2 presents the market shares for each firm in October 2015.

Table 2: Aggregate Market Shares of Alternatives

Market Size (millions)	ORG	SFR	BYT	FREE	MVNO	Non-users
56.5	29.4%	13.4%	17.2%	21.5%	10.6%	8.0%

Data reported by the regulator (ARCEP, 2016) provides the relative share of MVNOs and MNOs. Relative shares within MNOs obtained from GSMA Intelligence. Shares are adjusted to allow for 8% outside option share, consistent with CREDOC (2015).

We also construct a "Rest of France" municipality which aggregates the population and income distribution from all communes not included in our estimation sample. As we explain below, the Rest of France municipality plays a very limited role in the estimation; we include it primarily so that we can calculate aggregate market shares that can be compared to the national market shares in Table 2. Download speeds in the Rest of France municipality are computed as the average download speeds in all municipalities outside the 589 municipalities in our estimation sample. The Rest of France municipality is not involved in simulations or in estimating infrastructure costs, so we need not construct the infrastructure measurements described below for it. We also omit the Rest of France municipality from descriptive statistics presented below.

2.4 Infrastructure Data

Finally, we obtain detailed data on infrastructure from the national radio communications regulator (ANFR). These data describe the locations of all mobile telecom base stations, along with the number of antennas and frequencies operated by each network operator.[15]

Ultimately, we want to quantify the typical cell for each municipality, characterized by the area served by base stations and the bandwidth operated. For bandwidth, we simply compute the mean bandwidth operated across all base stations for each operator and each municipality. To measure the area of the typical cell, dividing municipality area by the number of base stations could be misleading. The concentration of base stations within uninhabited areas is typically low, but such areas have few users and low data demand. Thus, it could paint a misleading picture of the intensity of investment if we simply divided municipality area by the number of base stations, particularly in municipalities with large, uninhabited areas. We

[15] This database is publicly accessible at https://www.cartoradio.fr/.

instead consider a measure of the "adjusted area" of a commune. To this end, we compute the contraharmonic mean of population density across space (equivalently, the population density integrating across persons rather than space).[16] The adjusted area is defined as the municipality's population divided by the contraharmonic mean population density. We then measure the object of interest, the area served by a typical base station, as the adjusted area divided by number of base stations.[17] Note that if a municipality consists of a populated area (with a uniform population distribution) as well as an area with zero population density, the contraharmonic mean population density of the municipality will correspond to the population density in the populated area.

In addition to infrastructure data from ANFR, we use traffic data from OSIRIS, which is an internal database provided by Orange. OSIRIS provides the total volume of data traffic per network cell over time. We use these volumes to calculate data demand rates, which we then use to calibrate parameters of the data transmission model.

2.5 Descriptive statistics

Table 3 provides summary statistics for variables of interest.

Measured quality (download speeds) varies substantially both across and within markets. Across markets, the average standard deviation for an operator is 9.56 Mbps, and across operators, the average standard deviation for a market is 7.92 Mbps. Figure 1 displays histograms of measured quality across markets for each mobile network operator.[18] The raw Ookla data include few measurements for MVNO operators, so we set MVNO speeds equal to the average of Orange, Bouygues, and SFR within each market. Anecdotally, MVNOs contract with all three of these MNOs, but not with Free. Pooling the MVNO speed measurements that we do have across communes gives a roughly similar average download speed to this imputation: 22.7 Mbps in comparison to the mean of 24.7 Mbps in Table 3.

[16]The data we use for this is the Gridded Population of the World, v4, available from https://sedac.ciesin.columbia.edu/data/collection/gpw-v4.

[17]For example, Fontainbleau is a relatively populous commune consisting of a town surrounded by a forest. While the population density in the town is relatively high, the population density of the commune appears low if we divide by the commune's total area. Our measure of adjusted area for Fontainbleau is 69.6 square kilometers, while the raw municipality has an area of 172 square kilometers.

[18]There is a potential selection concern in these measures of download speeds. Because they come from voluntary speed tests, it may be the case that measurements tend to happen when consumers experience slow downloads. However, the levels of download speeds reported in Table 3 are consistent in the aggregate with the levels coming from other sources. We note that for Orange, Bouygues, and SFR, our average download speeds lie within the values reported by ARCEP for intermediate and urban density areas (the densities of areas in our sample). For Free, the 23 Mbps average download speed is actually higher than the 19 Mbps number reported by ARCEP. See https://www.arcep.fr/cartes-et-donnees/nos-publications-chiffrees/couverture-et-qualite-de-service-mobile-2g-3g-4g-5g/couverture-et-qualite-des-services-mobiles-juillet-2016.html (accessed November 7, 2022).

Table 3: Summary Statistics

	Mean	Std. Dev.	Min.	Max.
Customer data (Orange)				
Market average usage (MB)	1 044	194	555	1 702
Fraction users in market above data limit	0.18	0.03	0.10	0.28
Number of customers		4 425 831		
Quality and market data				
Quality Orange (Mbps)	32.82	11.11	3.97	84.98
Quality Bouygues (Mbps)	23.70	9.65	0.60	72.97
Quality Free (Mbps)	23.15	11.03	1.56	56.74
Quality SFR (Mbps)	17.57	8.58	0.39	52.30
Quality MVNO (Mbps)	24.70	7.04	5.13	48.87
Median income (Euros)	13 035	3 177	5 152	31 320
Number of markets		589		
Tariff data				
Price	23.47	14.22	2.00	64.99
Price (Orange)	23.92	9.90	12.07	38.74
Price (Others)	23.33	15.32	2.00	64.99
Data limit	3 081	3 484	0	10 000
Number of phone plans		22		
Infrastructure data				
Bandwidth per firm (MHz)	70.69	30.42	0.00	140.20
Number of base stations	7.47	21.47	0	511
Effective cell radius (km)	1.44	0.93	0.26	7.64

Customer, quality, market, and infrastructure data summary statistics are (unweighted) across markets. Tariff data summary statistics are across mobile phone plans.

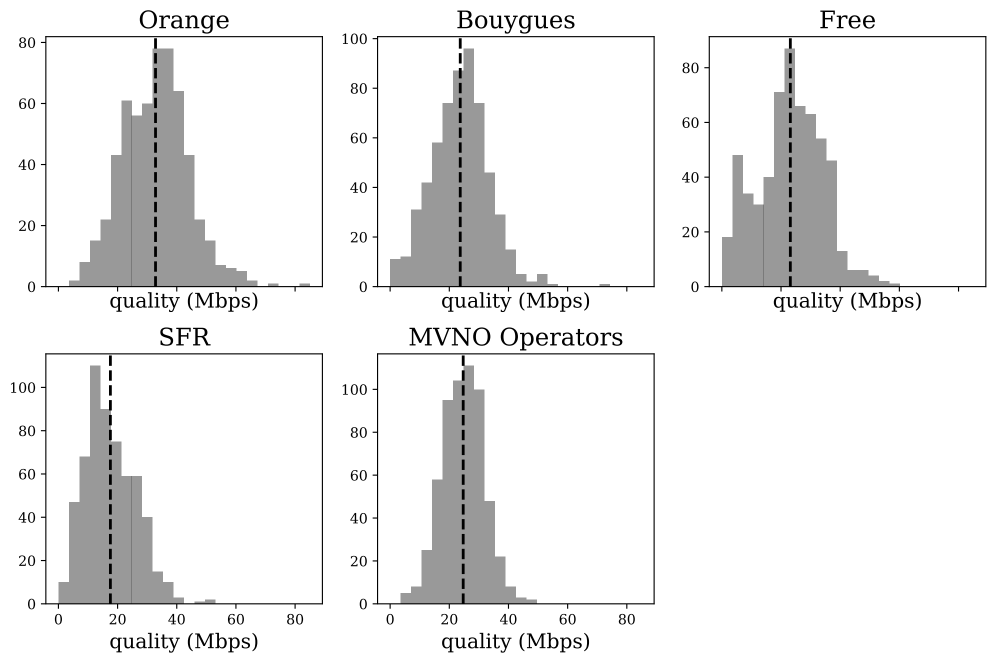
Data usage is positively correlated with measured quality. Figure 2 plots the relationship across markets between Orange download speeds and the observed average data usage for three different data limits.[19] Most consumers do not actually reach their data limit in a given month, and the average fraction of the data limit that is consumed is decreasing in the size of the data limit, as demonstrated in Figure 3, which plots the histograms of market-level average data consumption for three different data limits.[20]

Markets with higher median incomes tend to have higher market shares of expensive phone plans. Figure 4 plots the relationship between the median incomes in each market and the joint market share of the two most expensive Orange phone plans, with prices of 30.91 € and 38.74 €. Median incomes are positively correlated with the market share of these expensive plans, with a correlation coefficient of 0.495.

[19] The correlations for data limits 1 000 MB, 4 000 MB, and 8 000 MB are, respectively, 0.147, 0.270, 0.246.

[20] For the data limits 1 000 MB, 4 000 MB, and 8 000 MB, the fraction of the data limit that is consumed is, respectively, on average, 0.656, 0.578, and 0.534.

Figure 1: Histograms of qualities by operator



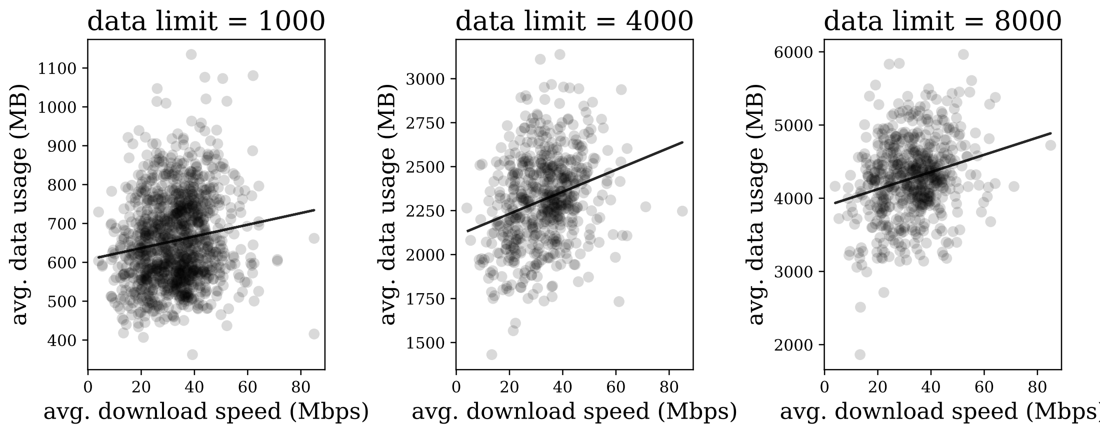
Note: Presented are the average download speeds at the market level for each operator. Dashed vertical lines represent the average of the market-level averages by operator, unweighted by population. The scale of the x-axis is the same across all subplots, allowing for comparisons in the distribution of average download speeds across operators.

3 Model

In this section we describe a formal model of consumer choice, how download speeds are determined, and firm competition. These components jointly provide a model of the mobile telecommunications industry that can capture how changes to market structure impact prices, quality of service, and welfare. We present each component in turn. The first component (Section 3.1), which captures how consumers choose mobile phone plans and how much data to consume, takes prices and download speeds as given. The second component (Section 3.2) maps consumer demand and infrastructure into download speeds, taking prices and infrastructural investments as given. The final component (Section 3.3) captures how firms choose the prices of mobile phone plans and the level of investment in infrastructure.

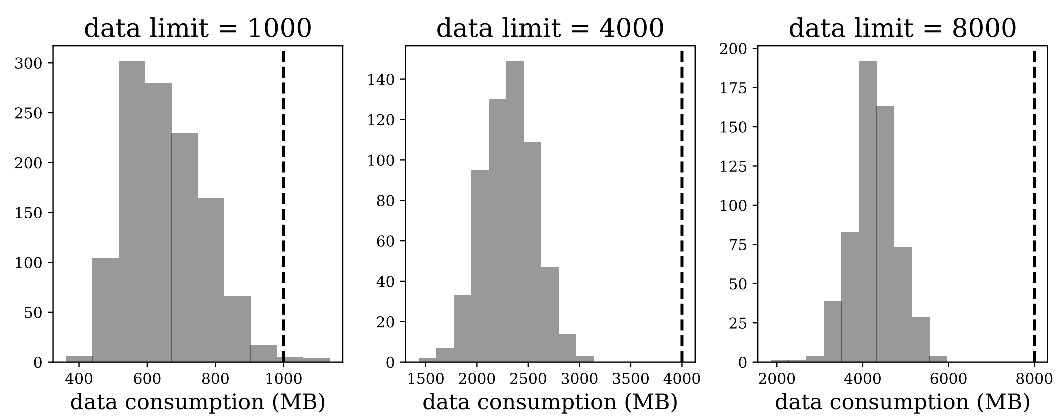
Before presenting each of these model components, we introduce some notation that is common to each of these components. There exist a set of mobile phone plans, $\mathcal{J}$, indexed by j. Each plan j belongs to a particular firm, $f(j)$, and the set of plans provided by a firm is given by $\mathcal{J}_f$. Consumers belong to different geographic markets, indexed by m, which vary by demographics and geography (the latter matters for the efficiency of data transmission). Table 15 in the

Figure 2: Average data usage vs. measured quality across markets



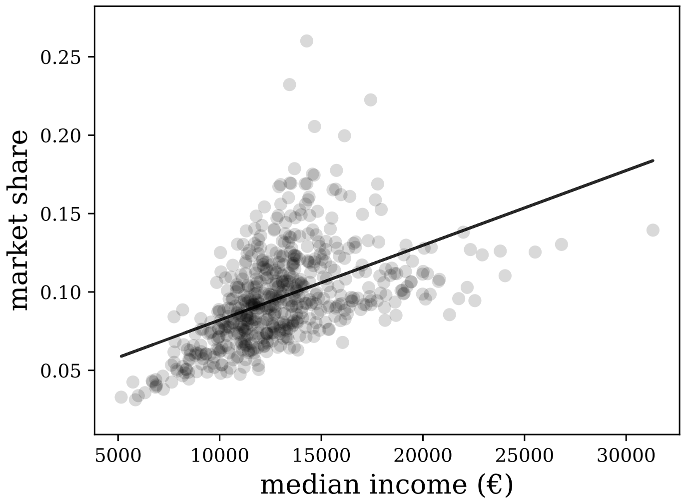
Note: Presented are scatterplots for three different Orange phone plans (corresponding to the three subplots) between the average download speed and the average amount of data consumed by each customer. Observations are at the market level. The line in each subplot is a line-of-best-fit for the observations.

Figure 3: Average data usage across markets



Note: Each subplot represents, for Orange phone plans with a particular data limit, a histogram of customers' average data consumption with that plan at the market level. Market-level average data consumption for each phone plan is obtained by, for each market, averaging data consumption across customers in that market subscribing to the phone plan. Dashed vertical lines represent the average of the market-level averages, unweighted by population.

Figure 4: Median income vs. expensive contract market share



Note: The figure presents a scatterplot of the median income in a market with the market share in that market of Orange's two most expensive phone plans (the 4 GB representative plan and the 8 GB representative plan). The line is a line-of-best-fit.

Appendix provides a list of all parameters used in the model and their definitions.

3.1 Demand Model

Consumers make decisions about to which mobile phone plan (if any) they subscribe and how much data to consume using that plan. Each mobile phone plan j in a market m is characterized by the download speed available in that market, $Q_{f(j),m}$;[21] the price of that phone plan, p_j; and a data consumption limit, $\bar{d}_j$. Note that download speeds are common across plans offered by the same firm, as firms do not discriminate across plans in the download speeds they offer. Note also that prices and data limits do not depend on the market. In France, mobile phone plan prices and characteristics (except download speeds) are set nationally.

A consumer's indirect utility for a plan j depends on the utility that they derive from consuming x megabytes of data and the product characteristics. This indirect utility is given by

$$u_{jm}\left(x, Q_{f(j),m}, P_j; \theta_i, \vartheta_i, \varepsilon_{ij}\right) = w_j\left(x, Q_{f(j),m}; \vartheta_i, \theta_i\right) + \theta_v v_j - \theta_{pi} P_j + \xi_{jm} + \varepsilon_{ij}, \quad (1)$$

where $w_j(\cdot)$ maps the plan j, data consumption x, and data quality $Q_{f(j),m}$ into the utility from consumption of mobile services. Other plan characteristics that enter the consumer's

[21]While consumers may be mobile, we assume that their choices depend on the network quality in their municipality of residence.

utility include the price, P_j; whether the plan has an unlimited voice allowance, captured by v_j (equal to 1 if plan j has an unlimited voice allowance, 0 otherwise); ξ_{jm}, the product-market-specific demand shock; and idiosyncratic tastes, ε_{ij}. The parameters θ and ϑ describe preferences. The preference parameter ϑ_i specifically captures how much consumer i values consuming data, described in detail in the following section.

3.1.1 Mobile Data Consumption

Subscribing to a particular plan j, a consumer chooses how much data to consume given the plan's data consumption limit, download speed, and the consumer's value of data consumption. They choose their level of consumption to maximize the utility from data consumption, $w_j(\cdot)$. To rationalize finite data consumption even when additional data consumption entails no monetary cost, our functional form of $w_j(\cdot)$ includes a term which corresponds to the disutility of download times. This disutility is proportional to the amount of data downloaded and inversely proportional to the download speed. It can be thought of as the opportunity cost of time spent downloading. Consumers will consume data until the marginal utility of extra data corresponds to the disutility of additional download time.

A consumer's utility of data consumption is given by the following functional form:

$$w_j(x, Q; \vartheta_i, \theta_i) = \vartheta_i \log(1 + x) - c_j(x, Q; \theta_i). \tag{2}$$

The first term captures the utility the consumer derives from consuming data. It exhibits decreasing marginal returns and depends on the parameter capturing how much the consumer values data consumed, ϑ_i. The second term, $c_j(\cdot)$, is the opportunity cost of the time spent downloading. It is given by the following formula:

$$c_j(x, Q; \theta_i) = \begin{cases} \theta_c \frac{x}{Q} & \text{if } x \leq \bar{d}_j \\ \theta_c \left(\frac{\bar{d}_j}{Q} + \frac{x - \bar{d}_j}{Q^L} \right) & \text{if } x > \bar{d}_j, \end{cases} \tag{3}$$

where θ_c is a preference parameter capturing how much the consumer dislikes waiting.

There is a discontinuity in download speeds when a consumer reaches their monthly data limit, $\bar{d}_j$, captured by the two cases in equation 3. Data consumed after reaching the data limit downloads at a throttled speed $Q^L \ll \mathbf{Q}$, where $\mathbf{Q}$ is download speeds stacked across firms and markets.[22]

This discontinuity in download speeds creates a discontinuity in the marginal cost of data

[22] MNOs in France typically use a throttled speed of 128 Kbps (see sectionC.1.2 in the appendix for more information about throttled download speeds). We use this value for throttled speeds in our estimation of demand and cost parameters as well as in our counterfactuals.

consumption. We let $x^*_{jm}(\cdot)$ denote the consumer's optimal data consumption:

$$x^*_{jm}\left(Q_{f(j),m};\vartheta_i,\theta_c\right) = \arg\max_{x\in\mathbb{R}_+}\left\{w_j\left(x, Q_{f(j),m};\vartheta_i,\theta_i\right)\right\}.$$

The first order condition and the structure of the marginal cost of data consumption yield four possible cases that determine the optimal data consumption:[23]

$$x^*_{jm}\left(Q_{f(j),m};\vartheta_i,\theta_i\right) = \begin{cases} 0 & \text{if } \vartheta_i \leq \frac{\theta_c}{Q_{f(j),m}} \\ \frac{\vartheta_i}{\theta_c/Q_{f(j),m}} - 1 & \text{if } \frac{\theta_c}{Q_{f(j),m}} \leq \vartheta_i < \left(\frac{\theta_c}{Q_{f(j)}}\right)\left(\bar{d}_j+1\right) \\ \bar{d}_j & \text{if } \frac{\theta_c}{Q_{f(j),m}}\left(\bar{d}_j+1\right) \leq \vartheta_i < \frac{\theta_c}{Q^L}\left(\bar{d}_j+1\right) \\ \frac{\vartheta_i}{\theta_c/Q^L} - 1 & \text{if } \vartheta_i \geq \frac{\theta_c}{Q^L}\left(\bar{d}_j+1\right). \end{cases} \tag{4}$$

The first case captures consumer types that would not consume any data.[24] The second case captures consumer types that consume less than $\bar{d}_j$ even without throttling. The third case captures consumer types that would consume greater than $\bar{d}_j$ if download speeds were not throttled, but under throttling, the marginal cost of an additional unit of data is greater than the marginal benefit, so they consume exactly the data limit. The final case captures consumer types that would consume greater than $\bar{d}_j$ even under throttled download speeds.[25]

3.1.2 Mobile Phone Plan Decision

A consumer i chooses the mobile phone plan that maximizes their expected utility. The expectation is with respect to the data consumption utility parameter ϑ_i, which is a random variable that we assume is distributed

$$\vartheta_i \sim \mathit{Exponential}\left(\theta_{di}\right).$$

That consumers do not know the realization of their ϑ_i prior to choosing a plan reflects that consumers may be unable to perfectly forecast their utility for data when choosing a phone plan. While consumers do not know their ϑ_i *ex ante*, they do know their θ_{di}.

Each market has an outside option, which is not subscribing to a phone plan. This option is represented by $j = 0$ and has indirect utility normalized to ε_{i0}.

Unlike with respect to the data consumption utility parameter ϑ_i, consumers do observe the realization of their vector of idiosyncratic taste shocks, $\boldsymbol{\varepsilon}_i$, prior to choosing a phone plan.

[23] We are using here the assumption that $Q^L \ll \mathbf{Q}$, which holds in our data.

[24] We interpret such consumers as those that do not need their mobile plan (e.g., they went out of the country for the month). In the data, we observe a point mass of consumers that consume zero data—even among those that adopt high data limit plans.

[25] Small data limit plans have hard data limits (i.e., there is no throttling). We therefore impose that all contracts with data limits less than 500 MB cannot consume greater than the associated data limit.

We assume a nested structure on the idiosyncratic shocks. Specifically,

$$\varepsilon_{ij} = \zeta_{ig(j)} + (1-\sigma)\,\eta_{ij},$$

where η_{ij} is i.i.d. extreme value and ζ_{ig} has the distribution such that ε_{ij} is extreme value. The value $\sigma \in [0,1)$ is the nesting parameter.[26] All phone plans (but not the outside option) belong to a single nest. The addition of a nest for all plans except the outside option allows for more flexible substitution patterns to the outside option.

Observing the utility parameter θ_i and idiosyncratic taste shocks $\boldsymbol{\varepsilon}_i$, consumer i chooses the phone plan that maximizes their utility, taking an expectation over their data consumption (i.e., their realization of ϑ_i). Their choice of phone plan j^*_{im} is therefore given by:

$$j^*_{im}\left(\mathbf{Q}_m, \mathbf{P}; \theta_i, \boldsymbol{\varepsilon}_i\right) = \arg\max_{j\in\mathcal{J}\cup\{0\}} \left\{\mathbb{E}\left[u_{jm}\left(x^*_j\left(Q_{f(j),m}; \vartheta_i, \theta_i\right), Q_{f(j),m}, P_j; \theta_i, \vartheta_i, \varepsilon_{ij}\right)\right]\right\}, \quad (5)$$

where the expectation is over ϑ_i.[27]

Integrating over idiosyncratic taste shocks, we obtain market shares for each mobile phone plan conditional on consumer type θ_i, given by

$$s_{ijm}\left(\mathbf{Q}_m, \mathbf{P}; \theta_i\right) = \int \mathbb{1}\left\{j = j^*_{im}\left(\mathbf{Q}_m, \mathbf{P}; \theta_i, \boldsymbol{\varepsilon}_i\right)\right\} dF\left(\boldsymbol{\varepsilon}_i\right), \quad (6)$$

and integrating over consumer types we get market shares:

$$s_{jm}\left(\mathbf{Q}_m, \mathbf{P}\right) = \int s_{ijm}\left(Q_{f(j),m}, P_j; \theta_i\right) dF_m\left(\theta_i\right). \quad (7)$$

These market shares, along with data consumption, given by equation 4, yields the average data consumed in a market m by consumers subscribed to a phone plan j:[28]

$$\bar{x}_{jm}\left(\mathbf{Q}_m, \mathbf{P}\right) = \int\int \frac{s_{ijm}\left(\mathbf{Q}_m, \mathbf{P}; \theta_i\right)}{s_{jm}\left(\mathbf{Q}_m, \mathbf{P}\right)} x^*_{jm}\left(Q_{f(j),m}; \vartheta_i, \theta_i\right) dF\left(\vartheta_i|\theta_i\right) dF_m\left(\theta_i\right). \quad (8)$$

3.2 Data Transmission Model

In this section, we lay out a formal model of how download speeds are jointly determined by bandwidth allocations, infrastructure investment decisions, and the load imposed on a network

[26] Note that if $\sigma = 0$, the model is equivalent to a random coefficients model without nesting.

[27] The expected value of utility from data consumption, $\mathbb{E}\left[w_j\left(x^*\left(Q;\theta_i\right), Q\right)\right]$, is feasible to derive analytically but extremely cumbersome to write. It is therefore omitted from the draft but may be found in our code. See function `E_u` in https://github.com/jonathantelliott/mobile-telecommunications/blob/main/code/demand/dataexpressions.py.

[28] An analytic expression for $\bar{x}_{jm}\left(\mathbf{Q}_m, \mathbf{P}\right)$ exists and may be found in our code. See function `E_x` in https://github.com/jonathantelliott/mobile-telecommunications/blob/main/code/demand/dataexpressions.py.

by consumers. We rely on standard telecommunications engineering models to determine how these components map to experienced download speeds and are particularly indebted to Błaszczyszyn, Jovanovicy and Karray (2014).

In this model, firms own and operate their own networks with no sharing of infrastructure. While *passive network sharing* (the sharing of the physical structure of base stations and the cost of electric power) is common, our cost function specification is in a sense robust to it, as we discuss below. During 2015, *active network sharing* (which occurs when equipment that transmits data is shared) occurred primarily in areas with low population density. Because we want to associate each firm's quality of service with its own investment decisions, we ultimately focus on the higher-density areas of France in our analysis. See Appendix C.4 for further discussion.

3.2.1 Base Station Infrastructure and Data Transmission

For each mobile network operator f and municipality m, we assume that each municipality has homogeneous population density and that the full land area is divided into equally-sized hexagonal cells, so each cell is identical for a given operator and municipality.[29] We assume that each cell is served by a single base station transmitting an omni-directional signal. A crucial aspect of infrastructure is owned spectrum or bandwidth, B_{fm}, which measures the range of frequencies a firm has the right to operate (in MHz). Bandwidth is not a choice variable in our model, but it is an aspect of market structure we vary in our counterfactual analysis.

The size of network operator f's cells in market m is characterized by R_{fm}, which is the cell radius (more precisely, a hexagonal cell's maximal radius, which is equal to its side length). In this section we take cell size as given and consider how firms choose the sizes of their cells in the next section. We could also think of this choice variable of the firms as being the number of base stations in a given municipality, N_{fm}. We assume the area served by each cell is $A_m/N_{fm} = \frac{3\sqrt{3}}{2}R_{fm}^2$, where A_m is municipality m's effective land area, and $3\sqrt{3}R_{fm}^2/2$ is the area of a regular hexagon with maximum radius R_{fm}.[30] Note that we take for granted that firms will serve the full municipality area. This is standard practice in recent engineering-based studies of mobile service provision in developed countries, reflecting the idea that quality, not coverage, is the relevant non-price characteristic that network operators now compete on in developed countries. We assume that the municipality's area can be divided into equally-

[29] We use the terminology "network operator" in this subsection since mobile virtual network operators do not own their own network resources, as explained in Section 2.1. All network operators are a firm (and therefore denoted by f), but not all firms are mobile network operators.

[30] When implementing the model empirically, we use an adjusted measure of land area because the raw land area may overstate the area that operators need to cover (at least with high quality) when large unpopulated areas are present. See Section 2.4 above for details.

sized hexagons, effectively ignoring municipality geometry and other spatially explicit details. Heterogeneity in municipality topography and other features that affect radio transmission can be captured in a municipality-level spectral efficiency parameter, introduced later in this section.

The point of the remainder of this section is to derive channel capacity, $\bar{Q}_{fm}\left(R_{fm}, B_{fm}\right)$, from basic principles. Channel capacity captures the average data transmission rate that a base station can achieve ignoring congestion. It is decreasing with respect to cell radius R_{fm} because signals must travel further (on average) as the area of a cell increases. Readers uninterested in the details of this derivation may skip to Section 3.2.2.

If and when a base station is serving a consumer at location ℓ, they will receive download speed $B_{fm}q_{m\ell}$ in megabits per second (data transmission rates generally scale linearly with bandwidth used to transmit data). Note that we assume that, when a consumer is served by a base station, that base station devotes its full bandwidth to that consumer. We will soon introduce a precise function to describe how $q_{m\ell}$ depends on the consumer's location within the cell, but generally, $q_{m\ell}$ will be lower for consumers located further from the base station due to path loss (the phenomenon of signals losing power as they travel across space).

Suppose that data demand has a uniform distribution over space. Normalizing data demanded per unit area to unity (this is harmless as the demand rate per unit area would cancel out of equation 10 below), the total data demanded within a cell will be equal to its area. We will denote a cell's area by $A\left(R_{fm}\right) = \frac{3\sqrt{3}}{2}R_{fm}^2$.

To determine the amount of time consumers spend downloading data, we need to integrate over $q_{m\ell}$. Specifically, for one unit of data downloaded, a consumer at location ℓ will spend $\left(B_{fm}q_{m\ell}\right)^{-1}$ seconds downloading. Integrating over consumers (or, equivalently, the cell's area, since consumers are uniformly distributed), the total time spent downloading is given by

$$\int_{\ell \in H\left(R_{fm}\right)} \frac{1}{B_{fm}q_{m\ell}} d\ell, \tag{9}$$

where $H\left(R_{fm}\right)$ is the set of locations composing a hexagon with radius R_{fm}.

To determine average download speeds, we divide total data downloaded by total time spent downloading. That is, we divide by $A\left(R_{fm}\right)$ by equation 9. This yields a harmonic mean of $q_{m\ell}$, multiplied by bandwidth:

$$\bar{Q}_{fm}\left(R_{fm}, B_{fm}\right) = \frac{B_{fm}A\left(R_{fm}\right)}{\int_{\ell \in H\left(R_{fm}\right)} q_{m\ell}^{-1} d\ell}. \tag{10}$$

The above equation expresses *channel capacity*, describing how feasible download speeds are influenced by the firm's choice of cell radius R_{fm} and its bandwidth B_{fm}. Importantly, channel capacity is not the same as *delivered* download speed. Below, we will model how delivered download speeds also depend on consumption (i.e., congestion) using queuing theory.

Next, we consider the individual download speed function $q_{m\ell}$, which gives download speed measured in bits per second (per Hertz of bandwidth):

$$q_{m\ell} = \gamma_m \log_2 \left(1 + SINR_\ell\right), \tag{11}$$

where $SINR_\ell$ is the signal-to-noise-and-interference ratio, which we will explain below, and γ_m is a spectral efficiency parameter.

When the spectral efficiency parameter is set equal to unity ($\gamma_m = 1$), equation 11 represents the Shannon-Hartley Theorem (Shannon, 1948), which provides the theoretical upper bound on data transmission rates as a function of SINR. The Shannon-Hartley Theorem's bound is much higher than the data transmission rates typically achieved in practice. Actual rates of data transmission are affected by the encoding technology, topography, weather, and the presence of buildings and other physical barriers. This means that the efficiency of data transmission may vary by market. We therefore employ a market-specific spectral efficiency parameter. We calculate these spectral efficiency parameters to match our model's predicted delivered download speeds with observed download speeds. See Section 4.3 for further discussion.

This spectral efficiency parameter can absorb many aspects of the data transmission technology, and in particular, anything that affects the level of download speeds without affecting how they decline with distance. For instance, one might be concerned that our measure of spectrum includes all the frequencies owned by an operator, and therefore the frequencies used for *both* downloads as well as uploads by mobile customers, but we're using this measure of bandwidth to model *only* download speeds. Operators could manage outgoing transmissions (downloads) and incoming transmissions (uploads) by using half of their spectrum for each (in practice, they have more sophisticated strategies). In this case, the relevant measure of spectrum for determining download speeds would be half of the spectrum owned by each operator; therefore, we would want to rescale our measure of bandwidth by a factor of 0.5. By calibrating our spectral efficiency parameter to observed download speeds, we implicitly achieve such a rescaling. Avoiding interference between incoming and outgoing signals is just one of many factors that tends to make $\gamma_m < 1$. Ultimately, we find a mean value of γ_m equal to 0.165 across municipalities.

The signal-to-noise-and-interference ratio (SINR) is given by the ratio of signal power to the sum of noise and interference power:

$$SINR_{\ell}\left(R_{fm}\right)=\frac{S_{\ell}}{N+I_{\ell}\left(R_{fm}\right)}, \tag{12}$$

where S_ℓ is signal power density (signal power per unit of bandwidth), N is noise power density, and $I_\ell\left(R_{fm}\right)$ is interference density. Note that signal power S_ℓ depends on location due to path loss. Interference power $I_\ell\left(R_{fm}\right)$ depends on both location and the cell's radius because cell size determines how far neighboring base stations are.

We assume that base stations transmit signals at the maximum power permitted by regulation. As the signal travels away from the base station, its power diminishes (path loss). We take this into account by using the Hata model of path loss (Hata, 1980), in which the signal power received by a consumer depends on their distance from the base station. In a vacuum, the signal power would be proportional to the squared inverse of the distance traveled. In telecommunications jargon, this is a *path loss exponent* of two. In the Hata model we use, the path loss exponent is 3.522, reflecting the fact that signal strength drops off more quickly as it travels along the Earth's surface than it would in a vacuum. See Appendix A.1.1 for the precise functional form of S_ℓ.

Noise power N is constant, and set equal to Johnson-Nyquist noise. Interference power I_ℓ is set equal to 30% of the signal power from the six adjacent cells. See Appendix A.1.2 for details regarding the units and formulas for the noise and interference variables.

3.2.2 Queuing

Consumers' download requests do not arrive uniformly over time. This means that channel capacity $\overline{Q}_{fm}$ derived above will not represent the actual delivered download speed in practice.

To derive a relationship between channel capacity and average delivered download speed, we follow Błaszczyszyn, Jovanovicy and Karray (2014) and Lhost, Pinto and Sibley (2015) and assume that download requests arrive according to a Poisson process and that download requests are served through a M/M/1 queue (a queuing system in which a single server serves jobs on a first-come, first-served basis). Then, the average download speed, Q_{fm}, will be

$$Q_{fm}=\overline{Q}_{fm}-Q_{fm}^{D}, \tag{13}$$

where Q_{fm}^D is the arrival rate of download requests among consumers served by the base station.[31] It comes from the demand model and is provided explicitly in equation 15 below. Each of the terms in equation 13 should be understood as rates, e.g., as values measured in

[31]For a derivation of this formula, see Taylor, Karlin and Taylor (1998), pp. 548-549.

megabits per second.

3.2.3 Download Speeds

Our demand system described how consumption depended on download speeds. The queuing theory model above describes how download speeds depend on consumption. We now consider how the engineering relationships described above come together with demand to simultaneously determine delivered download speeds in equilibrium. To be clear, in this section we consider equilibrium in terms of download speeds and consumer demand, taking prices and infrastructure as given. Formally, the equilibrium we now consider is conditional on a vector of prices of mobile phone plans $\mathbf{P}$ and infrastructure variables $(\mathbf{R}_m, \mathbf{B}_m)$, where $\mathbf{R}_m$ and $\mathbf{B}_m$ are the stacked cell radii and bandwidths of the network operators.

The total demand for downloads on network operator f's network over a month can be broken down into the product of three terms, which come from the demand component of our model:

$$X_{fm}\left(Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right) = pop_m \times s_{fm}\left(Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right) \times \bar{x}_{fm}\left(Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right), \tag{14}$$

where pop_m is the number of potential consumers in the market, and the market share and data consumption functions, $s_{fm}(\cdot)$ and $\bar{x}_{fm}(\cdot)$, come from the phone plan-level analogues in Section 3.1, summed across the phone plans offered by network operator f.[32]

The demand rate for downloads on network operator f's network is the total downloads serviced by operator f over a month, $X_{fm}(\cdot)$, distributed across time and across base stations. This rate is given by:

$$Q^D_{fm}\left(R_{fm}, Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right) = \frac{X_{fm}\left(Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right)}{H \times N_{fm}\left(R_{fm}\right)}, \tag{15}$$

where H is the number of seconds in a month and $N_{fm}(\cdot)$ is the number of base stations network operator f has in market m.[33]

Combining equations 10, 13, and 15, we have

$$\forall f = 1, \ldots, F: \qquad Q_{fm} = \bar{Q}_{fm}\left(R_{fm}, B_{fm}\right) - Q^D_{fm}\left(R_{fm}, Q_{fm}, \mathbf{P}_f, \mathbf{Q}_{-f,m}, \mathbf{P}_{-f}\right). \tag{16}$$

[32] MVNOs use MNOs' infrastructure for their own plans. Therefore, in our empirical analysis, we incorporate the load that results from the plans offered by the MVNOs on the MNOs' networks. ORG, BYG, and SFR all allow MVNOs to use their infrastructure, and (lacking data on these relationships) we assume MVNO load is distributed equally among these three MNOs.

[33] In our empirical application and counterfactuals, we use $H = 31 \times 8 \times 3600$. That is, we try to capture download speeds during peak hours when most of the downloads occur, and we assume that days effectively consist of eight peak hours.

Given prices and infrastructure variables, the vector of equilibrium download speeds $\mathbf{Q}_m^*$ is defined as the vector of values of Q_{fm} that solves equation 16. The download speed function $\mathbf{Q}_m^*(\mathbf{P}, \mathbf{R}_m, \mathbf{B}_m)$ describes equilibrium download speeds given prices and infrastructure variables.

3.2.4 Economies of Scale

Our model allows for two sources of scale efficiencies: *economies of pooling* and *economies of density*.

Economies of Pooling It has long been recognized in the economics literature that "there are economies of scale in servicing a stochastic market" (Carlton, 1978).[34] In operations management, the same phenomenon has been referred to as the "Pooling Principle" (Cattani and Schmidt, 2005). Thus, we use "economies of pooling" to describe economies of scale coming from consolidating bandwidth.

It is easy to see how economies of scale result from our queuing theory model. Equation 13 holds that the average delivered download speed corresponds to the difference between channel capacity and the download demand rate. Crucially, channel capacity is linear in bandwidth. Thus, if two identical firms combine their bandwidth and their customer bases (holding the download demand rate per customer fixed), then both terms on the right-hand side of equation 13 would double. Consequently, download speeds (the left-hand side) would also double.

Economies of Density Due to path loss, captured by the $q_{m\ell}$ function, the closer users are to a base station, the more efficiently that station can serve them. Thus, if we increase the density of users served by a firm while keeping constant the number of users per base station, users will be closer to base stations serving them on average, improving download speeds. Consequently, urban areas tend to be less costly to serve than rural areas (at a given quality level). Mapping this to our analysis, having fewer firms implies that individual firms face higher population densities of customers. Consequently, at a given level of investment in terms of base stations per customer, a market with fewer firms will have less path loss and higher download speeds (per unit of bandwidth).

We can quantify these economies of density by comparing the channel capacities that result from the case of two network operators to that of one network operator with the same number of base stations as the two combined, but arranged into a hexagonal grid with smaller cell

[34] Robinson (1948) was perhaps the first to describe the phenomenon, under the heading of "the economy of the large machine." De Vany (1976) was an early application using queuing theory to derive economies of scale. Mulligan (1983) shows formally how economies of scale result from queuing theory.

sizes. If each of the two operators have cell radii of R, then the combined operator would have a radius of $R/\sqrt{2}$. The combined firm will enjoy higher channel capacity per cell due to decreased path loss, but the degree of improvement is very sensitive to the baseline cell radius. If the two-operator case has a radius $R = 1$ km, then the single operator with the same number of base stations has a channel capacity that is just 0.1% larger (per unit of bandwidth operated). If $R = 5$ km, however, the combined operator would have a more substantial improvement of channel capacity, 19.4%. In our infrastructure data, the effective cell radii cover a range of values that includes both 1 km and 5 km (see Table 3), but they tend to be much closer to 1 km. This foreshadows one message from our counterfactual results: while economies of density can matter in principle, they have little impact for the typical cell sizes in our data. We revisit this discussion in Section 6.5, where we simulate equilibria for different population densities.

3.3 Firm Competition

In this section, we present how firms choose prices and infrastructure to maximize profits. We can understand the network equilibrium model in the previous section as holding at the market level m with potentially different infrastructural variables in each market, $(\boldsymbol{R}_m, \boldsymbol{B}_m)$. However, prices are set nationally, so we will not introduce subscripts on the price vectors. From now on, when the infrastructure variables appear without market subscripts, they refer to the stacked vector of infrastructure variables for all markets.

Firms set prices and infrastructure simultaneously in all markets in a static game. We consider the first-order conditions with respect to each competitive variable in turn.

3.3.1 Price Competition

Variable profits are given by

$$\left(\mathbf{P}_f - \mathbf{c}_f^u\right) \cdot \sum_m pop_m \mathbf{S}^*_{fm}\left(\mathbf{P}, \mathbf{R}_m, \mathbf{B}_m\right), \tag{17}$$

where $\mathbf{c}^u$ is the variable cost per customer, pop_m is the size of market m, and $\mathbf{S}^*_{mf}\left(\cdot\right)$ denotes the vector of product-level shares for phone plans offered by firm f in market m. This market share function is derived from the demand system and our download speed model (equation 16) as follows:

$$\mathbf{S}^*_{fm}\left(\mathbf{P}, \mathbf{R}_m, \mathbf{B}_m\right) = \mathbf{s}_{fm}\left(Q^*_{mf}\left(\mathbf{P}, \mathbf{R}_m, \mathbf{B}_m\right), \mathbf{Q}^*_{m,-f}\left(\mathbf{P}, \mathbf{R}_m, \mathbf{B}_m\right), \mathbf{P}_f, \mathbf{P}_{-f}\right),$$

where the $\mathbf{s}_{fm}\left(\cdot\right)$ corresponds to the stacked vector of firm f's phone plan-level market shares given by equation 7.

We assume that firms choose prices to maximize the variable profits expressed in 17. Note that equilibrium download speeds depend on price, so the first-order condition for optimal price-setting must not only take into account the direct effect of lowering price on consumer demand, but also the indirect effect of endogenous download speeds. The indirect effect lowers price elasticities because as demand for firm f falls, its download speeds increase due to reduced network load, which has a positive effect on demand, thereby dampening the demand reduction. We discuss demand elasticities further in Section 6.

3.3.2 Costs and Infrastructure Competition

Firms also decide on their infrastructural investments in each market, measured by R_{fm}. Infrastructure costs in market m are given by the following function:

$$C_{fm}\left(R_{fm}, B_{fm}\right) = c^{s}_{fm} \frac{A_m}{A\left(R_{fm}\right)} B_{fm}, \tag{18}$$

where A_m is the land area of market m, and c^{s}_{fm} captures costs per base station and unit of bandwidth (which may vary by network operator and by market), and $A\left(R\right) = 3\sqrt{3}R^2/2$ is the area of a hexagonal cell with radius R.

This cost function reflects the idea that the main costs associated with a base station are the electricity costs, the cost of installing antennas, and other costs that are proportional to the bandwidth being operated. An advantage of this cost function is that, if we suppose that all firms operate at the same base station locations, then redistributing bandwidth among firms and/or changing the number of firms does not change the total costs incurred within the industry. Thus, this cost function shuts down a potential source of economies of scale associated with the duplication of fixed costs.[35]

This cost function also rules out any gains from passive network sharing. Because costs are proportional to bandwidth, firms would not change their total costs by combining their network resources at a given location. While our analysis does not explicitly incorporate passive network sharing, this does *not* lead us to overstate the case for consolidation. That is, one might worry that some of the predicted counterfactual efficiency gains from consolidation will be overstated because those efficiency gains can be realized among firms without consolidating. Because this source of cost savings does not exist in our baseline model, this is not a concern when interpreting our main counterfactuals.

That said, it is natural to think that there are some fixed costs associated with operating a base station, such as rents or setup costs, that don't scale with the bandwidth being operated. We

[35]See Peha (2017) for an analysis of economies of scale in mobile services coming from fixed costs per base station (without the economies of density and economies of pooling we consider).

conduct robustness exercises with an alternative cost function that treats all infrastructure costs as fixed costs per base station (that is, dropping the B_{fm} term from equation 18). Appendix D includes results for this alternative cost function.

We can define market-level profits as follows:

$$\Pi_{fm}\left(\boldsymbol{P},\boldsymbol{R}_m,\boldsymbol{B}_m\right)=\left(\boldsymbol{P}_f-\mathbf{c}_f^u\right)\cdot\sum_m pop_m\boldsymbol{S}_{fm}^*\left(\boldsymbol{P},\boldsymbol{R}_m,\boldsymbol{B}_m\right)-C_{fm}\left(R_{fm},B_{fm}\right). \tag{19}$$

Finally, we can define the national profit function for each firm f:

$$\Pi_f\left(\boldsymbol{P},\boldsymbol{R},\boldsymbol{B}\right)=\sum_m\Pi_{mf}\left(\boldsymbol{P},\boldsymbol{R}_m,\boldsymbol{B}_m\right). \tag{20}$$

Equation 20 defines the profit function for each firm, summing across all 589 markets, and we assume that each firm unilaterally and simultaneously chooses a (national) price vector $\boldsymbol{P}_f$ and a vector of cell radii (a cell radius for each municipality) $\boldsymbol{R}_f$ to maximize their profits, taking other firms' price and infrastructure choices as given, yielding equilibrium prices $\boldsymbol{P}^*$ and radii $\boldsymbol{R}^*$.

4 Estimation

In this section we describe our method of estimating the parameters of model described in Section 3. We first describe how we estimate the demand model using a modified version of Berry, Levinsohn and Pakes (1995), described below. After estimating demand, we infer firm's costs based on the assumption that firms set prices and invest in quality optimally. Finally, we describe how we use the data transmission model to calibrate spectral efficiency parameters.

4.1 Demand Estimation

We seek to estimate the distribution of consumer parameters θ_i. Specifically, we have the following parameters

$$\theta_i=\left[\theta_{pi},\theta_c,\theta_{di},\theta_v\right]'.$$

Note that we have two heterogeneous parameters that we allow to vary by income. Specifically, we assume

$$\begin{pmatrix}\log\left(\theta_{pi}\right)\\ \log\left(\theta_{di}\right)\end{pmatrix}=\begin{pmatrix}\theta_{p0}\\ \theta_{d0}\end{pmatrix}+\begin{pmatrix}\theta_{pz}\\ \theta_{dz}\end{pmatrix}z_i, \tag{21}$$

where z_i is the consumer's income.

4.1.1 Unobserved Demand Component

As is standard in the demand estimation literature, we use market shares to back out the unobserved demand components $\boldsymbol{\xi}$. We observe the set of products (in our setting, phone plans) offered by all firms, but we only observe detailed market share data at the plan-market-level for Orange. For plans offered by other firms, we observe market shares at an aggregate firm-level. The standard BLP contraction mapping used to solve for $\boldsymbol{\xi}$ cannot recover the unobserved demand components with market shares at different levels of aggregation. We therefore use a modified technique (similar to Chu (2010)) that is able to handle market shares at different levels of aggregation.

Our modified estimation technique rationalizes plan-level market shares for Orange plans and only the firm-level aggregate market shares for the other firms. Formally, we assume

$$\forall j \in \mathcal{J}_{-O}, \forall m: \quad \xi_{jm} = \xi_{f(j)},$$

where $\mathcal{J}_{-O}$ is the set of non-Orange plans, and $f(j)$ is the firm that corresponds to plan j.[36] Appendix B.1 describes a modified version of the BLP contraction mapping that is capable of solving for the unique vector $\boldsymbol{\xi}$ under the above assumption.

Our estimates indicate that the variation in ξ's across firms is more substantial than the variation within firm across municipalities and products. In our preferred specification, the standard deviation of estimated ξ_{jm} for Orange is 0.122 while the standard deviation of firm-level $\xi_{f(j)}$ terms is 0.418 (including the mean value for Orange).

4.1.2 Elasticity and Nesting Parameter Imputations

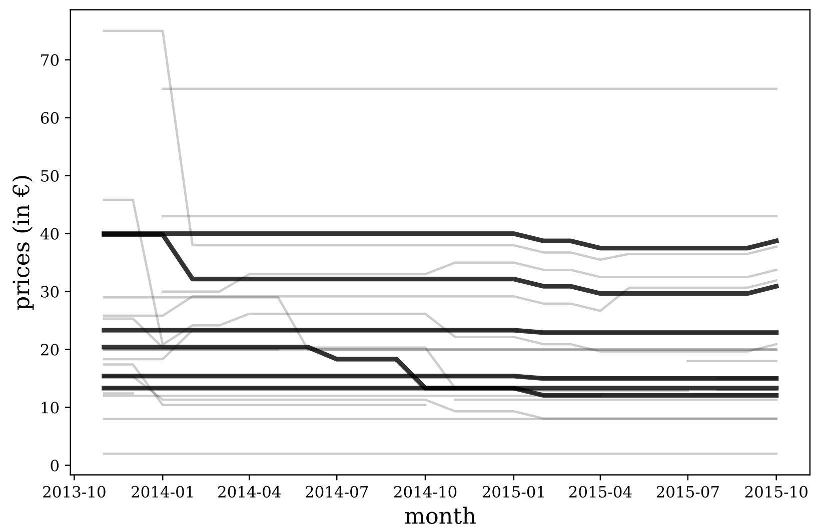
Prices are set nation-wide and do not vary by market. Moreover, prices varied very little over time around our sample period.[37] See Figure 5 for prices over the two years prior to our sample period. Prices of Orange phone plans are in black, and the prices of other operator plans are in light gray. Given the lack of price variation, it is difficult to identify price elasticities from the data.

We therefore take an approach where we impute price elasticities over a wide range of possible

[36] This is where we use the Rest of France municipality. When we find a value of the national shock $\xi_{f(j)}$ to rationalize the national market shares for a firm other than Orange, the data we use are the national market shares described in Table 2. Therefore, when computing national market shares predicted by the model, we want to sum over all of France, rather than summing over the 589 urban and suburban municipalities that we focus on for the purposes of understanding the infrastructural investment game.

[37] Note that Bourreau, Sun and Verboven (2021) consider a time period that includes the entry of Free Mobile in 2012. Following this entry, there were substantial price changes as the incumbent MNOs reacted to the new low-cost competitor. In contrast, during the two years leading up to our sample period, price variation was quite limited.

Figure 5: Prices of Orange phone plans over two years



elasticities. For each elasticity considered, we impose that the price elasticity of Orange products corresponds to the imposed elasticity. Formally, we calculate the implied Orange products price elasticity in market m, defined as follows:

$$e_m^O(\theta) = \frac{\mathbf{s}_{O,m}(1.01\boldsymbol{P}_O, \boldsymbol{P}_{-O}, \mathbf{Q}_m; \theta) - \mathbf{s}_{O,m}(\boldsymbol{P}_O, \boldsymbol{P}_{-O}, \mathbf{Q}_m; \theta)}{0.01\mathbf{s}_{O,m}(\boldsymbol{P}_O, \boldsymbol{P}_{-O}, \mathbf{Q}_m; \theta)},$$

where $\mathbf{s}_{O,m}(\cdot)$ is the vector of market shares of phone plans offered by Orange, as defined by equation 7, and $\mathbf{P}_O$ and $\mathbf{P}_{-O}$ represent the prices of plans offered, respectively, by Orange and by the other firms.

For a range of price elasticities $E \in \mathcal{E}$, we require that

$$\mathbb{E}\left[e_m^O(\theta) - E\right] = 0$$

as a moment in our estimation procedure, described below.

Our reference point for these imputations is Bourreau, Sun and Verboven (2021), who study the French market around the entry of Free Mobile, a few years earlier than our sample period. Free's entry was disruptive, resulting in considerable price and choice set variation, but prices settled down before the period covered by our data. Bourreau, Sun and Verboven (2021) estimate $\mathbb{E}\left[e_m^O(\theta)\right]$ to be approximately -2.5; we will treat this value as our baseline imputation, and we will also consider imputations of -1.8 and -3.2 as robustness checks.

We also impute values for the nesting parameter, σ. Lack of variation in the set of phone plans available prevents us from being able to feasibly estimate this parameter. For a given own-price elasticity imputation, the nesting parameter effectively controls how much substitution goes to the outside option. As the imputed σ approaches one, there is effectively no outside option. At the opposite extreme, $\sigma = 0$ yields a mixed logit model with no nesting.

While these imputations represent strong assumptions, note that there are still important aspects of consumer demand to be estimated, particularly how consumers trade off prices and download speeds (and how consumers differ in such preferences). In this paper, we present estimates as well as counterfactual results for a range of elasticity and nesting parameter imputations. These results are located in Appendix D.

4.1.3 Identification

Data utility parameters θ_{d0}, θ_{dz}, and θ_c are identified, in part, by matching predicted data consumption with observed data consumption. Formally, from the data we construct $\bar{x}_{jm}$, which is the average data consumption across consumers using phone plan j in market m. Given θ, we can construct the predicted mean data consumption across consumers in market m that choose phone plan j using equation 8.

Matching observed and predicted data consumption effectively identifies the average θ_{di} conditional on θ_c. The covariance across markets between the median income in the market and the average data consumption helps to identify how θ_{di} varies by income. We therefore use a moment interacting the difference between predicted and observed data consumption and median market income.

Simply matching data consumption does not separately identify consumption behavior from the level of the utility derived from consuming data, however. We need additional moments to be able to jointly identify θ_{d0}, θ_{dz}, and θ_c. Data consumption limits and download speeds change the costs of consuming data, so variation in these two phone plan characteristics creates variation in the utility coming from data consumption.

Download speeds may be endogenous to demand shocks due to congestion and if firms' investment decisions depend on local demand conditions. Therefore, we instrument download speeds with (log) population densities, which influence experienced download speeds by changing the level of path loss. Another reason for using an instrument is attenuation bias. Our measures of download speeds are based on limited sample sizes (see Appendix C.3 for details), and therefore there is a degree of measurement error in the variable we use. Our use of an instrument that is based on unrelated measurements alleviates concerns about attenuation bias.

Parameters associated with other plan characteristics are identified in a straightforward way. The imputed elasticity moment effectively identifies the average θ_{pi}. Variation in median incomes across markets helps to identify how this parameter varies by income. We assume that the demand shocks $\boldsymbol{\xi}$ are uncorrelated with median incomes. Voice allowances are also assumed to be uncorrelated with the demand shocks.

In summary, we have the following moments that we use to identify the distribution of preference parameters θ. Note that the moments are only evaluated for Orange plans since we only observe data consumption and plan-market shares for Orange.

Moments
$\mathbb{E}\left[e_m^O(\theta) - E\right] = 0$
$\mathbb{E}\left[\xi_{jm}(\theta) inc_m^{med}\right] = 0$
$\mathbb{E}\left[\bar{x}_{jm}(\theta) - \bar{x}_{jm}\right] = 0$
$\mathbb{E}\left[\left(\bar{x}_{jm}(\theta) - \bar{x}_{jm}\right) inc_m^{med}\right] = 0$
$\mathbb{E}\left[\xi_{jm}(\theta) \log\left(pop\_density_m\right)\right] = 0$
$\mathbb{E}\left[\xi_{jm}(\theta) \bar{d}_j\right] = 0$
$\mathbb{E}\left[\xi_{jm}(\theta) v_j\right] = 0$

Our data come from consumers subscribing to plans that don't always have the same data limit as the representative plan used in the model. Recall that we aggregate plans by data category, so, for example, a consumer might subscribe to an Orange plan with a 3GB-per-month limit, but in our econometric model, we treat them as having subscribed to Orange's representative plan with a 4GB limit. To avoid concerns about matching predicted average data consumption $\bar{x}_{jm}(\theta)$ with observed average data consumption $\bar{x}_{jm}$, we adjust the consumption data so that the proportion of the representative plan's data limit consumed corresponds to the proportion of the true data limit consumed in the raw data. See Appendix C.2 for details.

We use two-stage efficient GMM to estimate θ, searching for θ in an outer loop and solving for $\boldsymbol{\xi}(\theta)$ in an inner loop using the modified contraction mapping described in Appendix B.1. Further details regarding our estimation procedure can be found in Appendix B.2.

4.2 Cost Estimation

There are two types of cost parameters to be estimated: c_j^u, the cost per user of phone plan j, and c_{fm}^s, the cost per base station in market m for network operator f.

4.2.1 Costs per User

From equation 17, the first-order condition from the price setting game is

$$\sum_m pop_m \boldsymbol{S}_{mf}^*\left(\boldsymbol{P}, \boldsymbol{R}_m, \boldsymbol{B}_m\right) + \left(\sum_m N_m J_f \boldsymbol{S}_{mf}^*\left(\boldsymbol{P}, \boldsymbol{R}_m, \boldsymbol{B}_m\right)\right)\left(\boldsymbol{P}_f - \mathbf{c}_f^u\right) = 0, \qquad (22)$$

where J_f represents the Jacobian operator with respect to $\boldsymbol{P}_f$.

Therefore, an estimate of per-user marginal costs is given by

$$\hat{\mathbf{c}}_f^u = \boldsymbol{P}_f + \left(\sum_m pop_m J_f \boldsymbol{S}^*_{mf}\left(\boldsymbol{P}, \boldsymbol{R}_m, \boldsymbol{B}_m\right)\right)^{-1} \sum_m N_m \boldsymbol{S}^*_{mf}\left(\boldsymbol{P}, \boldsymbol{R}_m, \boldsymbol{B}_m\right). \tag{23}$$

4.2.2 Infrastructure Costs

Given the demand estimates and the model of how the infrastructure variables $(\mathbf{R}, \mathbf{B})$ map into delivered quality, we can simulate how equilibrium revenues change as the infrastructure is changed. Intuitively, we can measure the marginal revenue of infrastructure, and this allows us to infer the marginal cost of infrastructure.

Formally, we approximate the marginal operating income with respect to cell radius using numerical differentiation from each market based on a 0.01 km change in cell radius:

$$MR_{fm}^R\left(\boldsymbol{R}_m, \boldsymbol{B}_m\right) = \frac{\Pi_{fm}\left(\boldsymbol{P}, \left(R_{fm} + 0.01, \boldsymbol{R}_{-f,m}\right), \boldsymbol{B}_m\right) - \Pi_{fm}\left(\boldsymbol{P}, \left(R_{fm} - 0.01, \boldsymbol{R}_{-f,m}\right), \boldsymbol{B}_m\right)}{0.02}. \tag{24}$$

Note that these profit functions are defined in terms of the equilibrium download speeds that result from the infrastructural investment and prices. Thus, the above expressions for marginal operating income should be understood as implicitly taking into account how quality changes as infrastructural investment is changed. Furthermore, note that profits Π_{fm} include per-user costs; hence our use of "operating income" rather than "revenue."

Next, assuming that infrastructure investments are chosen to maximize profits, we can use the marginal operating income above to recover the remaining cost function parameters. Specifically, the marginal cost of increasing R_{fm} is obtained by differentiating the cost function in equation 18. For each firm and municipality, our estimated cost parameter c_{fm}^s sets this marginal cost equal to the marginal operating income in equation 24.

4.3 Spectral Efficiency Calibration

We calibrate the spectral efficiency parameter γ_m using delivered download speed data for each municipality. This is done by solving for the value of γ_m that makes equation 16 hold for Orange (we do not have usage data for other operators). In this calibration, the average experienced download speed Q_{fm} is the average download speed in Mbps in the delivered download speed data obtained from Ookla. Q^D comes from the OSIRIS infrastructure usage data. For each market, we determine Q^D by calculating the amount of data requested of Orange per second between noon and 1 pm and dividing by the number of Orange base stations in that market. Solving for the γ that makes equation 16 hold yields a market-specific spectral efficiency, $\hat{\gamma}_m$, for each market.[38] Across municipalities, the mean value of $\hat{\gamma}_m$ is 0.165, and its standard deviation is 0.048.

We recover γ_m focusing on Orange (for which we have data on infrastructure usage to construct Q^D) and assume that this commune-level spectral efficiency parameter applies to every firm. This is reasonable given that the firms have access to the same technologies. One potential reason for differences across firms would be heterogeneous spectrum holdings. In the Hata model of path loss, higher frequency spectrum is associated with greater loss (see equation 29). Thus, a firm with spectrum holdings primarily in high-frequency bands would experience greater loss than a firm with considerable holdings of low-frequency spectrum. Reassuringly, the spectrum holdings of Orange, SFR, and Bouygues were very similar in 2015.

Free owned little low-frequency spectrum in 2015. This is a potential cause for concern when estimating Free's costs using our structural model with the value of γ_m recovered using Orange. Free could have a different spectral efficiency parameter. Furthermore, another reason to worry about using equation 24 to recover Free's infrastructure costs is that Free benefited from active network sharing with Orange, meaning its delivered quality did not only depend on its own infrastructure investments, but also on Orange's (see Section C.4). Despite these concerns, our estimates of Free's infrastructure costs are not drastically different from those of the other firms (see Table 14). Thus, dropping Free's cost estimates before running the counterfactuals described below (which involve symmetric firms with costs parameters that are averages across our estimates) makes little difference to the results.

5 Results

5.1 Demand Estimates

Demand parameter estimates are listed in Table 9 in Appendix D.1 for a range of imputed price elasticities and imputed nesting parameters. The price elasticity implied by Bourreau, Sun and Verboven (2021) is approximately -2.5, the middle imputed price elasticity, which we regard as our preferred specification. For all imputations, price sensitivity is decreasing in income. The data utility parameter is increasing in income, which implies an inverse relationship between income and the value of data consumption, suggesting a higher opportunity cost of time spent

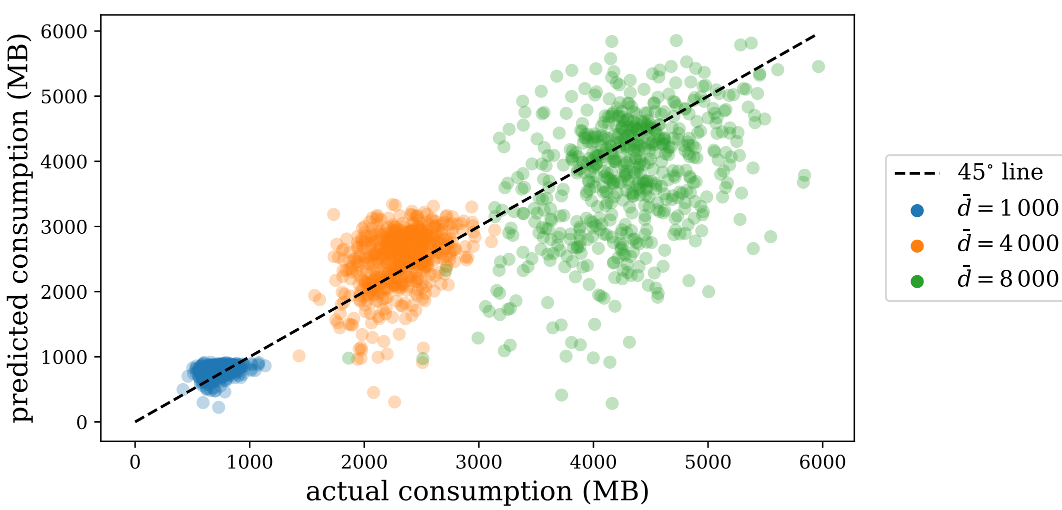
[38] Some frequencies are used for 3G technology, while others are used for 4G. We account for these technology differences by calculating the channel capacity for each technology separately (i.e., $\bar{Q}(R, B^{3G})$ and $\bar{Q}(R, B^{4G})$). We then adjust the 3G channel capacity to its 4G-equivalent by using the ratio of 3G-to-4G maximum link spectral efficiencies (respectively, 2.5 and 4.08 (Kim, 2015)). Therefore, in determining $\hat{\gamma}_m$ in each market, we use for the channel capacity

$$\bar{Q}_{fm}\left(R_{fm}, B^{3G}_{fm}, B^{4G}_{fm}\right) = \frac{2.5}{4.08}\bar{Q}_{fm}\left(R_{fm}, B^{3G}_{fm}\right) + \bar{Q}_{fm}\left(R_{fm}, B^{4G}_{fm}\right).$$

downloading for higher income individuals. The variance parameter is increasing in income. While signs are consistent across elasticities, the parameter estimates appear to be sensitive to the price elasticity chosen, especially price, voice allowance, and Orange dummy coefficients.

To help interpret the results above, Tables 10–12 in Appendix D.1 convert the parameter estimates into willingness to pay for certain contract characteristics across income percentiles. Figure 6 considers how well our model predicts actual data consumption by plotting predicted and actual average data consumption across markets for three Orange contracts with different data limits.[39] The diagonal line is a 45-degree line. Markets in which predicted average consumption equals observed average consumption will lie upon the line. Our estimated model correctly predicts the average level, even though this level is not a constant fraction of the data limit. While it predicts across-market heterogeneity less well, it does weakly predict high data consumption for markets with high observed data consumption and low data consumption for markets with low observed data consumption. The correlation coefficients between actual and predicted consumption for the three contracts across markets are, respectively, 0.304, 0.384, and 0.404.

Figure 6: Predicted vs. actual average data consumption



5.2 Cost Estimates

Estimated costs for our elasticity and nesting parameter imputations are given in Table 13 (per user costs) and Table 14 (infrastructure costs) in Appendix D.2.

In our preferred specification mean estimates cost of a 75MHz base station is approximately 120 000 €. Our estimates fit squarely within the range of estimates of the costs of constructing

[39]The predicted average data consumption is based on parameter estimates for the imputed elasticity -2.5 and a nesting parameter of 0.75.

large base stations, which tend to range from 50 000 to 250 000 € (Nikolikj and Janevski, 2014; Analysys Mason, 2015; Smail and Weijia, 2017).

6 Counterfactual Simulations

Our framework can address questions of market structure, both in terms of traditional antitrust questions as well as questions related to the management of the electromagnetic spectrum. In Section 6.1, we consider the optimal number of firms and the trade-off between market power and scale economies. Then, in Section 6.2, we consider the marginal value of spectrum allocated to mobile telecommunications and find that the marginal contribution to consumer surplus far exceeds firms' willingness to pay. In Section 6.3 we consider two different ways of allocating new spectrum in the industry: sponsoring the entry of a new firm, or allocating it among existing firms. In Section 6.4 we take a short-run focus, considering a change in the number of firms while holding infrastructure fixed. Finally, in Section 6.5 we consider how differences in population density impact equilibrium prices, quality of service, and the optimal number of firms.

In all our counterfactual simulations, we take market structure (described by the number of firms and spectrum allocation among them) as exogenous. This may seem at odds with the fact that spectrum allocations are endogenously determined through auctions. However, the recent literature has expressed increasing concerns about whether spectrum auctions lead to efficient post-auction outcomes, especially given that achieving efficiency in a spectrum auction is not as simple as achieving an efficient allocation of resources among auction participants; the regulator cares about outcomes for consumers, not just the bidding firms.[40]

Economists have considered how to design auctions to take into account post-auction market structure (Cramton et al., 2011; Rey and Salant, 2017), and it is well beyond our purview to argue that telecommunications regulators should simply regulate market structure more directly. That said, it is possible in principle for a regulator to directly control spectrum allocations, and our view is that it is worthwhile to consider the problem of a strong regulator that controls market structure at the level we define it (i.e., a vector of spectrum allocations across firms). Furthermore, our framework can complement the auction design paradigm; auction outcomes correspond to market structure, and our framework provides an understanding of how market structure maps to broader equilibrium outcomes.

[40]For instance, Jehiel et al. (2003) argue that features of multi-unit auction design that lead to more efficient allocation among auction participants can exacerbate post-auction market structure concerns. Eső, Nocke and White (2010) point out that increasing capacity, when capacity is efficiently allocated from the perspective of firms, can actually lead to a reduction in consumer welfare. Ershov and Salant (2022) present empirical evidence that some spectrum auctions have adverse impacts on market structure.

For each of the counterfactuals that we consider, we solve for the equilibrium for a representative commune in which each firm offers two mobile phone plans: one with a moderately low data limit of 1 GB and one with a very high data limit (in 2015) of 10 GB, which is the largest of the representative contracts. The representative commune that we construct has an income distribution matching the overall income distribution in our sample, population density equal to the contraharmonic mean density in France (2 792 people / km^2), available bandwidth equal to the population-weighted mean of the sum of frequencies operated in each market, a spectral efficiency parameter equal to the population-weighted mean in the calibration described in Section 4.3, and cost parameters equal to the mean estimated with equation 18, weighted by population. Both phone plans have an unlimited voice allowance, demand shocks equal to the average of those estimated for the Orange phone plans,[41] and per-user costs equal to the average of the estimated per-user costs for similar phone plans (those with $\bar{d}_j < 5$ GB for the low data limit plan and those with $\bar{d}_j \geq 5$ GB for the high data limit one).

One might worry about whether the focus on a representative commune yields results that hold for France when considered as a whole. In particular, does the representative commune, with its moderate population density, yield the same optimal number of firms that we would find for France, which comprises a mixture of high and low population-density areas? In Section 6.5, we find that the optimal number of firms is basically invariant to population density. Since the optimal number of firms for the representative commune is also optimal for high- and low-density areas, the optimum for the representative commune will also correspond to France in the aggregate.

While we compute the counterfactual equilibria for a wide range of imputed elasticities and nesting parameters (see Section 4.1.2 for the details regarding these imputations), we present results in this section for a single choice of these values: an overall price elasticity of -2.5 for Orange and a nesting parameter of 0.75. Results for other possible elasticities and nesting parameters are located in Appendix D.3. This price elasticity for Orange is approximately the same value as the price elasticity for Orange implied by Bourreau, Sun and Verboven (2021).[42] Our reason for preferring a high value of the nesting parameter is that substitution to the outside option appears to be relatively unimportant in the industry. Almost all adults own a mobile phone (in France in 2015, 92% of residents age 12 and up had a mobile phone), and anecdotally, few people even consider not having a phone or using a second. Table 4 presents the rate at which customers would substitute to the outside option in response to a

[41] Specifically, we set $\xi_{j,m} = \theta_O$, where θ_O is described in Appendix B.2.

[42] Bourreau, Sun, and Verboven report an own-price elasticity of -2.9 for Orange's postpaid contracts. While postpaid contracts represent the majority of Orange's mobile contract sales, we are interested in the elasticity of overall demand for Orange's products with respect to a price change for all their products. Using the market shares, diversion ratios, and elasticities reported by Bourreau, Sun, and Verboven, we compute Orange's overall price elasticity to be -2.4.

10% increase in all mobile plan prices. Note that a nesting parameter of $\sigma = 0$ (equivalent to multinomial logit with no nesting) features 3.29% of consumers switching to the outside option, and for a nesting parameter of $\sigma = 0.5$, we still get more than half of this rate of outside option substitution. Nesting parameters of $\sigma = 0.75$ or $\sigma = 0.85$ entail considerably less outside option substitution. We present in this section the nesting parameter $\sigma = 0.75$, which yields 0.95% of consumers switching to the outside option after a 10% increase in prices.

Table 4: Proportion of consumers who switch to outside option after a 10% overall price increase

Elasticity	$\sigma = 0.0$	$\sigma = 0.5$	$\boldsymbol{\sigma = 0.75}$	$\sigma = 0.85$
-3.2	3.90%	2.32%	1.21%	0.73%
$\mathbf{-2.5}$	3.29%	1.86%	**0.95%**	0.68%
-1.8	2.58%	1.39%	0.69%	0.45%

Displayed are proportions of consumers with a phone plan who would switch from a phone plan to the outside option after a 10% increase in the prices of *all* plans. We hold download speeds fixed at the values observed in the data. The row in bold corresponds to the imputed elasticity and nesting parameter we present in this section.

6.1 Market Power and Scale Efficiencies

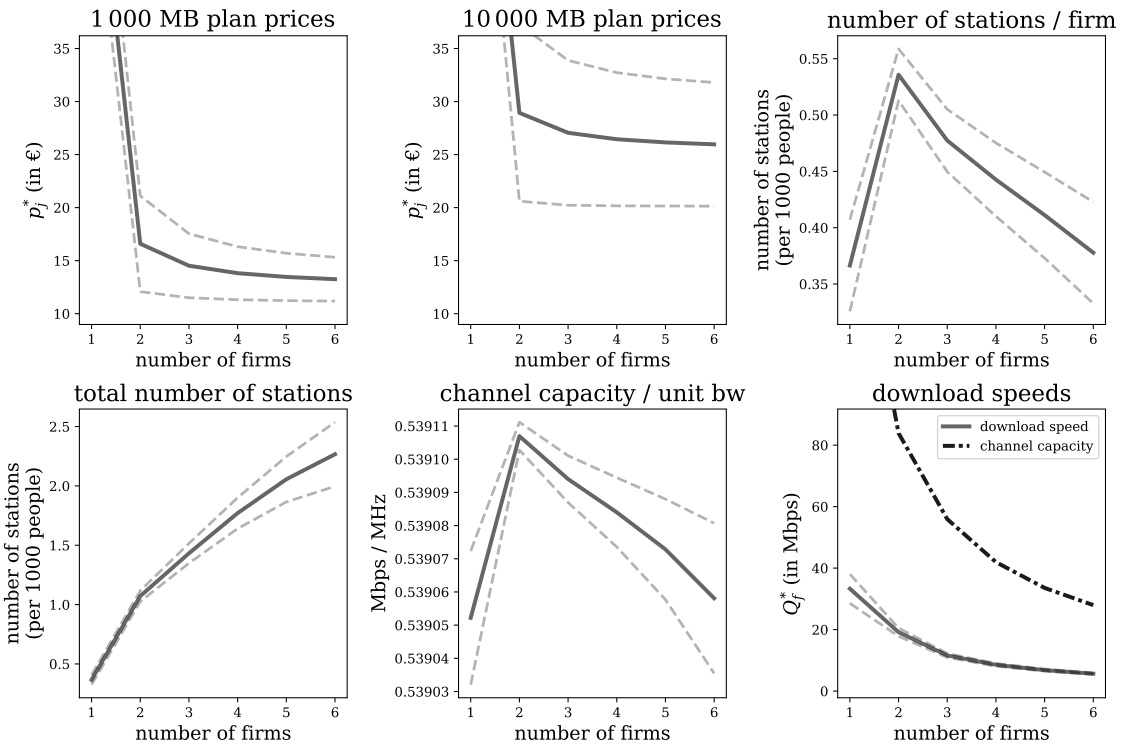
In this section, we explore the trade-off between market power and economies of scale by considering the optimal number of firms in a static equilibrium. Fewer firms gives each firm more market power but results in a higher density of consumers (lowering average path loss) and more pooling of consumers (reducing congestion at a base station). Given the gradual nature of network deployment in the industry, this exercise cannot hope to capture the short-run impacts of a potential merger; instead, we aim to capture the long-run trade-offs associated with consolidation (given the supply and demand models estimated using 2015 data).

The optimal number of firms depends on how equilibrium prices, investment, and download speeds vary based on the number of firms. Figure 7 displays these endogenous variables for symmetric equilibria with between one and six firms. Total bandwidth available to the industry is divided equally among the firms, which optimally set prices and investment levels. That is, each firm owns and operates spectrum $B_{fm} = B_0/n$, where B_0 is the total bandwidth available to the industry, and n is the number of firms.

We focus on symmetric equilibria with firms with identical spectrum endowments. When considering the optimal number of firms, a reason for our focus on symmetric spectrum holdings is that asymmetric spectrum allocations can be inefficient (Peha, 2017).

Equilibrium prices are declining in the number of firms but remain well above per-user marginal costs, which are 7.93 € and 18.18 € for the low and high data limit plans, respectively. Prices determine to which plan consumers subscribe and therefore the amount of data consumed. As a firm lowers its price, it attracts more customers, causing the load on its

Figure 7: Counterfactual prices and qualities



Note: Channel capacity is per base station. Download speeds are the average speed of transmission received by a user, including wait times. Dashed lines represent 95% confidence intervals.

network to increase, lowering download speeds. Lower download speeds dampen the appeal of the lowered price. The relevant elasticity for the purposes of setting optimal prices, therefore, involves a full derivative that takes into account this indirect effect of changing prices on download speeds. Figure 8 displays how this indirect effect from download speeds influences optimal price setting behavior by displaying two elasticities: *partial price elasticities* and *full price elasticities*. Partial price elasticities are the price elasticities holding the quality of service fixed, evaluated at equilibrium prices. Full price elasticities allow quality of service to adjust with the price. They decline less with the number of firms than the partial elasticities. The reason for the divergence between the full and partial price elasticities is the worsening of the indirect quality effect as the number of firms grows. When there are many firms, a firm's own capacity is small relative to the number of consumers that they can potentially attract from other firms, making quality of service degrade more for a given price increase.

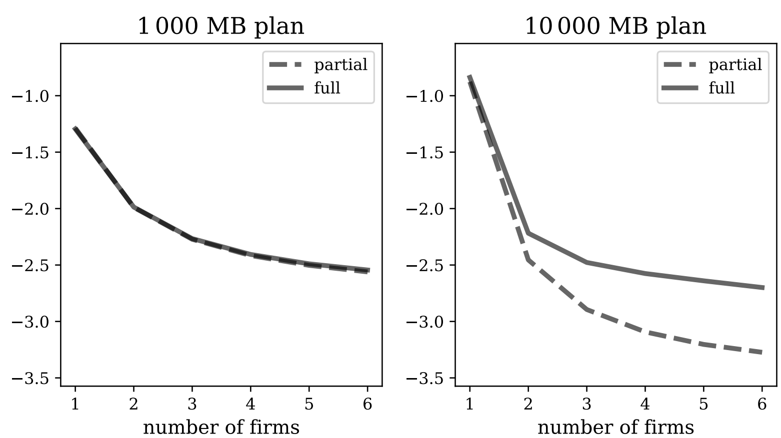
Investment patterns display a non-monotonic relationship in the number of firms. For a small number of firms, the number of base stations each firm builds is increasing in the number of the firms (alternatively, the cell radius characterizing each base station is decreasing). Increasing the number of firms beyond 2, however, decreases investment at the firm level: for each increase in the number of firms, each firm builds fewer base stations (increases the cell radius).

Despite this non-monotonicity in investment, download speeds are always decreasing in the number of firms. Comparing the monopoly case to the duopoly one, despite fewer base stations for the monopolist, we observe higher download speeds, reflecting economies of scale.

Closer inspection reveals that these economies of scale are driven largely by economies of pooling, rather than economies of density. Path loss will reduce channel capacity per unit of bandwidth, and when firms invest in more base stations, those base stations will serve closer customers, reducing path loss. As expected, channel capacity per unit of bandwidth follows the same shape as the number of base stations per firm, but note the scale of the graph for channel capacity per unit of bandwidth; the differences are trivial. In other words, firms are not seeing significant gains in data transmission by avoiding path loss here.

In contrast, economies of pooling have a large impact. We see that channel capacity is roughly inversely proportional to the number of firms, which is driven by channel capacity's proportionality to bandwidth operated (see equation 10 and the fact that total available bandwidth is being spread across the firms, i.e., $B_{fm} = B_0/n$).

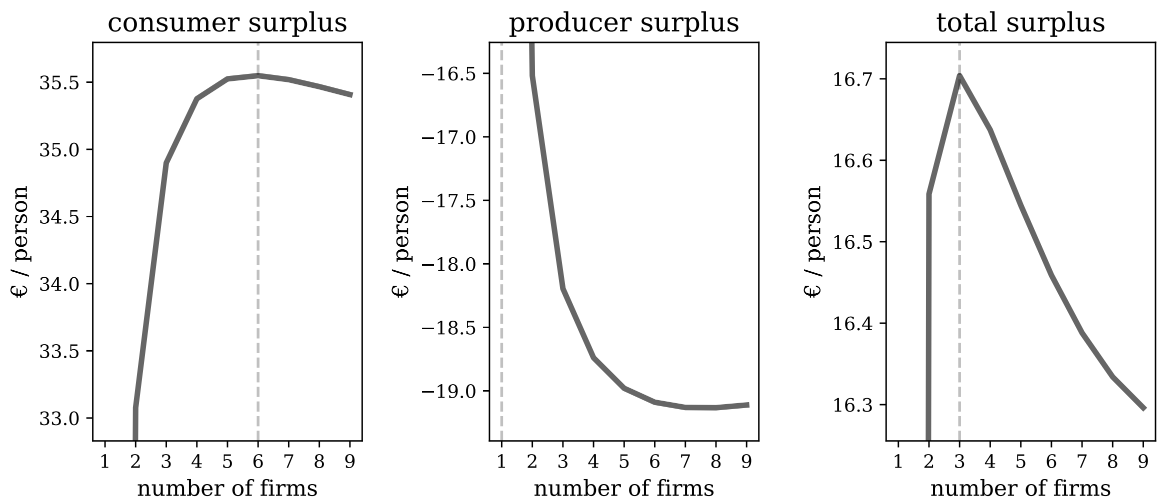
Figure 8: Full and partial price elasticities



Note: Partial elasticities are derivatives in which download speeds are held fixed. Full elasticities take into account how download speeds change endogenously as prices are changed. Price elasticities are evaluated at the equilibrium prices and quantities.

With both prices and quality declining in the number of firms, the optimal number depends on the trade-off between price and quality. Figure 9 considers welfare compared to the monopoly case as the number of firms is varied. For our preferred demand specification (elasticity of -2.5 and nesting parameter of 0.75), the optimal number of firms is three in terms of total surplus, and six in terms of consumer surplus. We present here consumer surplus calculated without including the logit error terms (ε_{ij}). This means that our consumer surplus results reflect differences in prices and download speeds rather than being mechanically driven by the number of firms.

Figure 9: Counterfactual welfare



Note: Welfare is measured in euros per capita relative to monopoly.

As Figure 10 illustrates, however, consumers do not agree on the optimal number of firms. We plot welfare for various income deciles against the number of firms for our preferred specification. While consumer surplus is increasing in the number of firms for most consumers (up to six or seven firms), the optimal number of firms for high-income consumers is four. In all our simulations, we have observed that the optimal number of firms is (weakly) decreasing with income.

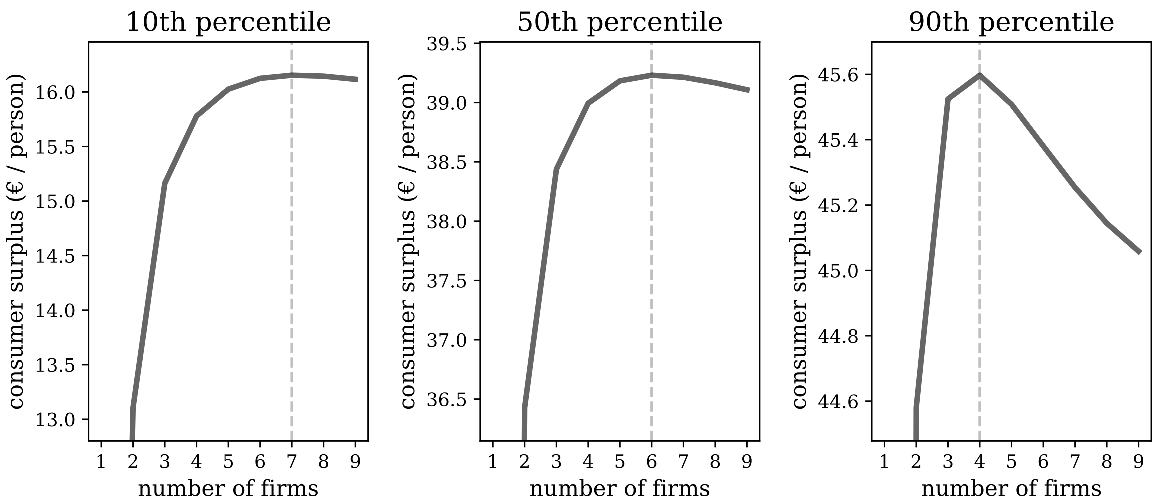
We note that Appendix D.3 indicates that the optimal number of firms is sensitive to the elasticity imputation (and considerably less sensitive to the nesting parameter). This points to the importance of careful demand analysis in future work, particularly in contexts where the estimates of Bourreau, Sun and Verboven (2021) may not apply.

6.2 Allocating Spectrum to the Industry

Regulators such as the FCC in the US and ARCEP and ANFR in France are tasked with bandwidth allocation, determining which industries (and firms) are allowed to operate which frequencies of electromagnetic spectrum and for what purposes. It is therefore crucial for such agencies to understand how allocating bandwidth to mobile telecommunications will affect social welfare.[43]

[43] The FCC's mandate is explicitly in "the public interest." To allocate spectrum optimally among different industries—or to allocate the optimal amount of spectrum to mobile telecommunications—one would need to

Figure 10: Counterfactual welfare by income level



Note: Welfare is measured in euros per capita relative to monopoly.

In this section, we quantify how allocating more bandwidth to the telecommunications industry affects firm profits, consumer welfare, and total surplus. When a firm receives a larger bandwidth allocation—holding prices and infrastructure fixed—its download speeds increase and it gains market share. The derivative

$$\frac{d\Pi_f\left(\mathbf{R}^*\left(B_f, \mathbf{B}_{-f}\right), \left(B_f, \mathbf{B}_{-f}\right)\right)}{dB_f} \tag{25}$$

captures how a firm's profit changes when just that firm receives a larger bandwidth allocation, taking into account how equilibrium investment and prices respond. This value captures an individual firm's willingness to pay for more bandwidth at the margin. The derivative

$$\frac{d\Pi_f\left(\mathbf{R}^*\left(B_f, B_{f'}, \mathbf{B}_{-f,f'}\right), \left(B_f, B_{f'}, \mathbf{B}_{-f,f'}\right)\right)}{dB_{f'}} \tag{26}$$

captures the impact of this increase in allocated bandwidth on *other* firms' profits.[44]

In a simple spectrum auction, the firms' bids will be related to the difference between these two expressions. A firm's bid reflects its own gain in profits from the increased bandwidth should it win the auction relative to losing the auction and the spectrum being allocated to another firm.

A regulator's decision of whether to allocate spectrum to mobile telecommunications should

quantify the social opportunity cost of spectrum, which is beyond our scope.

[44]We are assuming, as in Section 6.1, that firms are symmetric (prior to changing bandwidth allocations), so the identity of f' does not matter so long as $f' \neq f$.

be based not on the firms' bids, however, but on the marginal social value of allocating the bandwidth to the industry (compared to the marginal social value of allocating it to other industries and purposes). This marginal social value is captured by the following two derivatives. The first derivative,

$$\frac{d\Pi_f\left(\mathbf{R}^*\left(B\mathbf{1}\right), B\mathbf{1}\right)}{dB}, \tag{27}$$

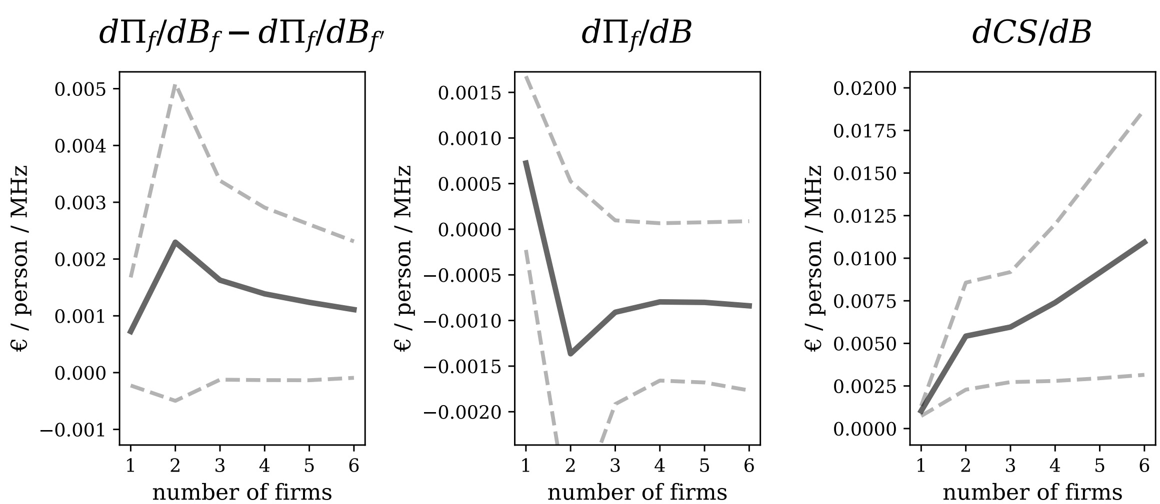
captures how the equilibrium profit of an individual firm changes when *all* firms are allocated more bandwidth. The second derivative,

$$\frac{dCS\left(\mathbf{R}^*\left(B\mathbf{1}\right), B\mathbf{1}\right)}{dB}, \tag{28}$$

captures how consumer surplus changes as all firms are allocated more bandwidth.

Our framework and estimates allow us to calculate each of these values. This allows us to not only calculate the marginal social value of allocating more bandwidth to the industry, but also to compare that value to that of the difference between expressions 25 and 26. Since spectrum auctions provide a signal of the marginal willingness to pay for spectrum (rather than it being allocated to a rival), this comparison is informative about how similar that willingness to pay is to the social value relevant to a regulator allocating spectrum across industries and uses.

Figure 11: Bandwidth derivatives



Note: Derivatives are evaluated at the symmetric equilibrium values. The derivative of own profits with respect to another firm's bandwidth ($d\Pi_f/dB_{f'}$) is undefined in the monopoly case. In the first subplot, therefore, what is reported in the case of only one firm is simply the derivative of own profits with respect to own bandwidth ($d\Pi_f/dB_f$). Dashed lines represent 95% confidence intervals.

As Figure 11 shows, with four firms, the firm's willingness to pay for additional bandwidth (the left panel) is about five times less than a unit of bandwidth allocated to the industry would add to consumer surplus (the right panel). This reflects the importance of using a structural model such as ours to quantify the social value of bandwidth. While auctions may allow us to observe signals of operators' willingness to pay for spectrum, such measures may be far lower than the social value of spectrum.[45]

In late 2015, the same time as our analysis, France auctioned off 60 MHz of spectrum (20-year licenses) in the 700 MHz band that was previously used by television broadcasters. The auction raised 2.8 billion euros, which, dividing by France's population at the time (66.55 million), results in 0.70 €/person/MHz. Our estimate of a firm's marginal willingness to pay for spectrum (rather than have it be allocated to a rival) at four symmetric firms is 0.28 €/person/MHz if we assume a monthly discount rate of 0.5% and therefore multiply the value in Figure 11 by $\frac{1}{1-0.995} = 200$. While this number is below the value in the auction, using our alternative base station cost function in which costs do not scale with bandwidth results in a willingness to pay of 1.19 €/person/MHz, higher than the value implied by the auction.[46]

6.3 Allocating Spectrum within the Industry

Spectrum allocation questions go well beyond the question of how much spectrum to allocate to each industry. In particular, how should spectrum be allocated among firms? In this section, we consider two ways of allocating new spectrum to the mobile telecommunications industry. First, the regulator could distribute the new spectrum among existing operators. Alternatively, it could sponsor the entry of a new operator, as happened in France with Free Mobile, which received regulatory approval to become France's fourth MNO in 2009 and launched in 2012.

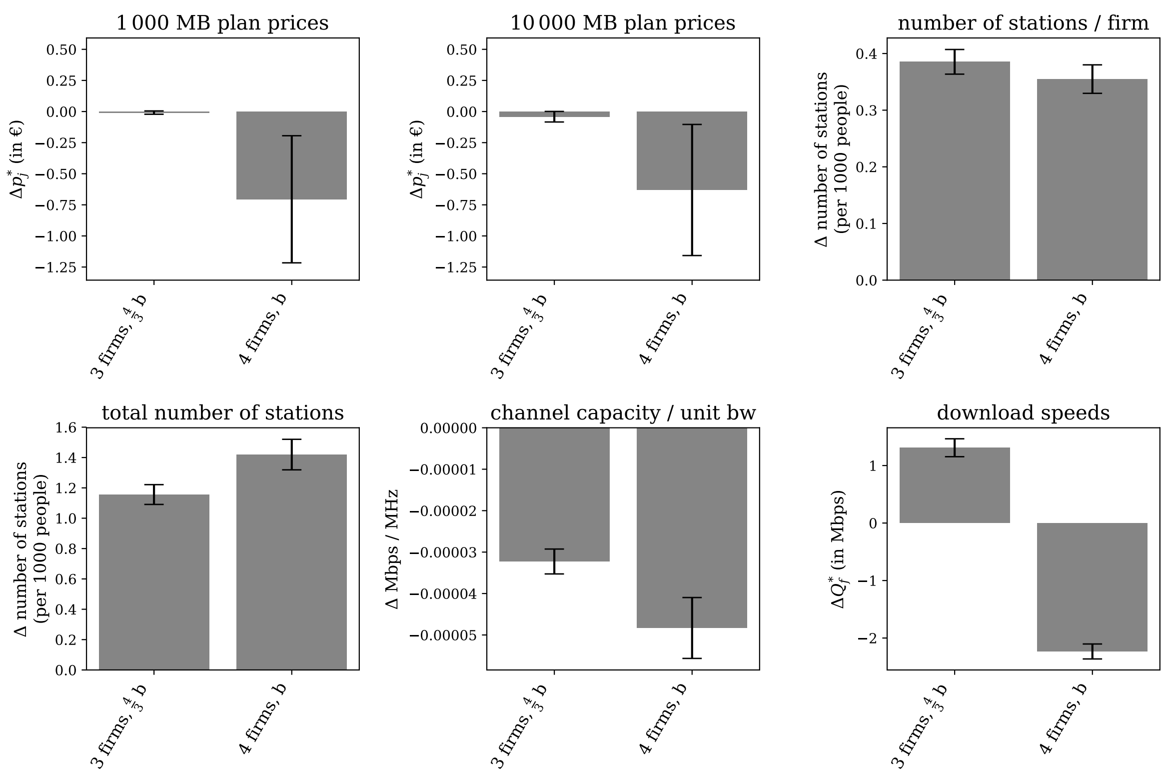
To study how spectrum should be allocated among firms, we consider two different ways of increasing the total amount of spectrum in the industry by 33.3%. Our baseline equilibrium is the symmetric equilibrium with three firms from Section 6.1. We compare this baseline to the equilibria resulting from two alternative ways of distributing extra bandwidth. The first equilibrium we consider increases each existing firm's bandwidth holdings by 33.3%. The second we consider is adding another firm, so that there are four firms total, with the new entrant having the same amount of bandwidth as the three individually do in the baseline (thus increasing total industry bandwidth by the same amount as in the first equilibrium).

[45]Of course, a regulator seeking to maximize total surplus would also need to consider the middle panel, but these values are small relative to the right one since firms compete away the surplus from additional bandwidth, so the point that the value of additional bandwidth is many times larger than that captured by spectrum auctions still stands.

[46]This under- and over-estimation depending on the cost specification points to a possible way of estimating a more flexible cost function. Rather than assuming base station costs are either fully proportional or fully invariant to bandwidth, one could estimate the proportion of the costs that are fixed and variable with bandwidth so that the model's implied willingness to pay for spectrum matches the value implied by the auction.

Figure 12 illustrates how various endogenous variables change with the additional bandwidth compared to the baseline equilibrium. Unsurprisingly, introducing a new firm leads to lower prices than increasing bandwidth per firm. However, download speeds benefit considerably more when bandwidth per firm is increased and actually decrease when a fourth firm is added.

Figure 12: Counterfactual prices and qualities

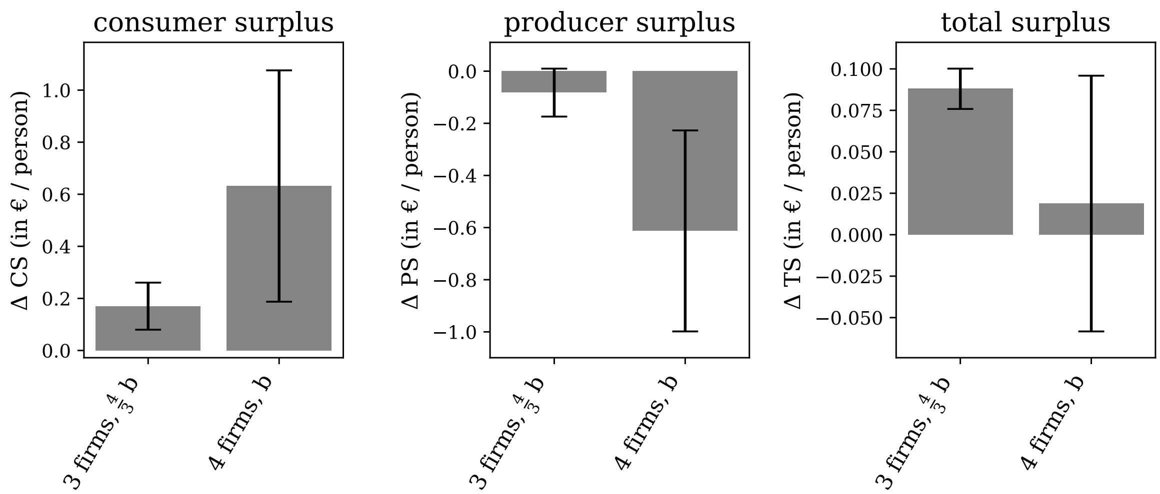


Note: Error bars represent 95% confidence intervals.

Figure 13 considers the overall effects on welfare and presents an interesting tension. Increasing the number of firms is better for consumer surplus (and consumers of all income deciles prefer that allocation to the one with more bandwidth per firm). Increasing bandwidth per firm is better for total surplus, however.

We note that heterogeneity in spectrum holdings plays an important role in some questions of spectrum allocation – e.g., heterogeneous spectrum holdings was an important source of claimed efficiency gains n the Sprint-T-Mobile merger (Asker and Katz, 2022). Our model could be extended to address such questions given that the model of path loss can capture how the loss of power varies for different spectral bands (see Section A.1.1 and in particular footnote 48).

Figure 13: Counterfactual welfare



Note: Error bars represent 95% confidence intervals.

6.4 Short-Run Analysis

The comparative statics exercise with respect to the number of firms in Section 6.1 should be interpreted with caution when extrapolating to merger analysis. Because those counterfactuals involve static equilibria, they certainly cannot capture the short-run impacts of mergers, for infrastructure cannot be rearranged instantaneously and costlessly in response to a change in the number of firms.

In this section, we consider the impact of consolidation in the short-run. That is, we change the number of firms and recompute an equilibrium without allowing infrastructure to adjust.

Tables 5 and 6 describe how outcomes change when we move from the four-firm equilibrium of Section 6.1 to an equilibrium with three firms with the base station radius fixed at the equilibrium radius from the original four-firm equilibrium. That is, we are crudely approximating a symmetric four-to-three merger in the short run in which infrastructure is fixed but prices can freely adjust.

To be clear, bandwidth is redistributed so that each firm in the three-firm equilibrium has 33.3% more bandwidth than each firm in the four-firm equilibrium. Furthermore, we impose that each firm's base station radius (or cell size) is the same in the three-firm and four-firm equilibrium. What we imagine is that all firms are sharing passive infrastructure, meaning they all have base stations located at the same places, and they each operate their own antennas on shared physical structures. When we consolidate to three firms, the antennas (and bandwidth) are simply consolidated, and each of the three operators now owns one third of the network infrastructure at each base station site rather than one quarter.

Table 5 shows that phone plan prices increase relative to the four-firm equilibrium in both the short-run and the long-run equilibrium (which corresponds to the equilibrium presented in Section 6.1). Prices of both the low- and high-end phone plans increase more in the short-run equilibrium.

In Section 6.1, there was a higher number of base stations per firm with three firms than with four firms. In this short-run equilibrium, we have fixed the number of base stations per firm, so the four-to-three comparison involves a more modest gain in download speeds here, where the gains in quality of service are driven entirely by consolidation of bandwidth and higher equilibrium prices (which increase download speeds by reducing congestion).

Table 5: Three firms with four-firm base station density: endogenous variables

	Δ 1 000 MB plan prices (in €)	Δ 10 000 MB plan prices (in €)	Δ download speeds (in Mbps)
short-run	0.710 (0.264)	0.642 (0.270)	2.425 (0.074)
long-run	0.702 (0.266)	0.607 (0.282)	3.044 (0.120)
difference	0.008 (0.002)	0.035 (0.013)	−0.619 (0.091)

Table 6 shows that relative to the four-firm equilibrium, consumer surplus declines and producer and total surplus improve. The decline in consumer surplus is larger in the short-run equilibrium, reflecting the higher prices and slower download speeds. Producers gain more in the short-run, as they do not compete with each other to increase the density of cells. The increase in total surplus is smaller in the short-run than the long-run, though the difference is quite small.

Table 6: Three firms with four-firm base station density: welfare

	Δ CS	Δ PS	Δ TS
short-run	−0.571 (0.179)	0.631 (0.213)	0.060 (0.028)
long-run	−0.478 (0.181)	0.544 (0.159)	0.067 (0.034)
difference	−0.093 (0.036)	0.086 (0.055)	−0.007 (0.019)

Welfare measured in euros per capita per month.

6.5 Impact of Population Density

Thus far, our counterfactuals have focused on a market with the contraharmonic mean population density in France, i.e., the mean population density when the mean is taken over people, rather than space. This density of 2 792 persons / km^2 roughly corresponds to a high-density suburb.

A natural question is whether the population density affects the trade-off between market power and scale efficiencies, perhaps changing the optimal number of firms. We first note that, with no path loss, the equilibrium comparative statics with respect to population density would be very straightforward.

No Path Loss Without path loss, channel capacity is fixed by the bandwidth owned and operated by the firm. The cell radius will not affect channel capacity. The decision of cell radius amounts to a decision of how many customers to serve with each base station, with the firm effectively choosing the optimal level of congestion. The population density will not affect this choice when we think about it in terms of the optimal number of consumers per base station (or the optimal level of congestion). As population density increases, the optimal number of consumers per station remains constant, implying base station area will be inversely proportional to population density. Equilibrium outcomes like prices and delivered download speeds remain the same. See Section A.2 for a more formal account.

In addition to France's contraharmonic mean population density (2 792 people/km^2), we consider three alternative population densities: the raw population densities of the continental USA (43.1) and France (123.9)—note that these are both quite low densities as both countries involve large unpopulated areas—and the population density of Paris (20 588).

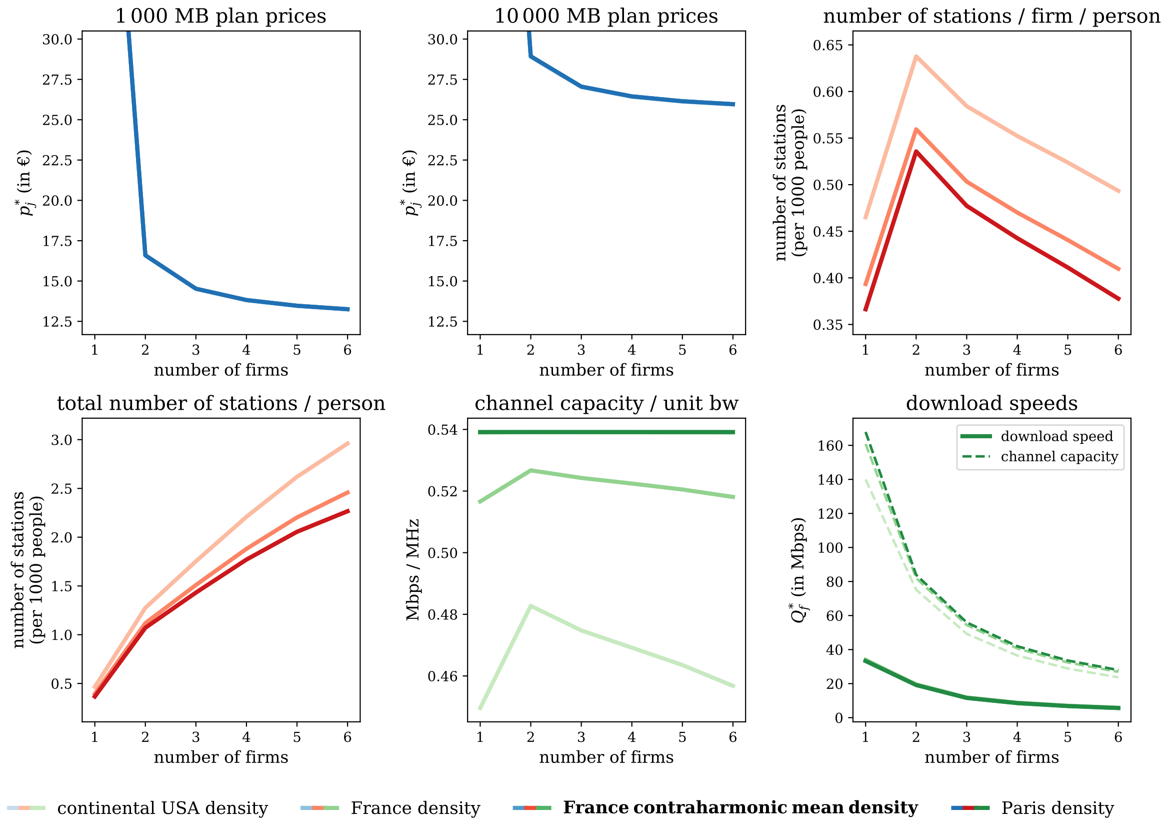
Figures 14 illustrates how equilibrium outcomes for these different population densities. Certain outcomes are indeed affected by population density. Naturally, path loss is more severe when serving a less dense market, demonstrated by lower channel capacities per unit of bandwidth in Figure 14 (despite higher levels of investment per person).[47]

Otherwise, the comparative statics with respect to population density are very similar to what we would expect without path loss. In other words, we do not see substantial economies of density. Figure 16 depicts channel capacity as a function of the cell's radius. For radii corresponding to the equilibrium radii in our counterfactuals, this function is quite flat, which is consistent with economies of density not being substantial across these population densities, although they may be at extremely low densities.

The optimal number of firms (for consumer or total surplus), depicted in Figure 15, is quite robust to the population density. Equilibrium outcomes like prices and delivered download speeds are extremely similar for different population densities. A takeaway is that, given the equilibrium cell sizes we observe, economies of density only appear to be a significant concern in very sparely populated areas.

[47] For each of these densities, we use the Hata model of path loss presented in Appendix A.1.1. This Hata model is for small cities. We have also simulated these counterfactual densities with rural and suburban Hata models of path loss for the associated densities, which exhibit less path loss as a function of distance. Results look similar but correspond more closely to the case of no path loss (in which the density does not matter).

Figure 14: Counterfactual prices and qualities by density



Note: Channel capacity is per base station. Download speeds are the average speed of transmission received by a user, including wait times.

7 Conclusion

The regulation of the mobile telecommunications industry, including antitrust policy and spectrum allocation, calls for an understanding of scale efficiencies as well as market power. Our approach has effectively been an interdisciplinary one, drawing from tools in empirical industrial organization to understand market power, and from wireless engineering to understand scale efficiencies. Our simulations show how our framework can shed light on many issues related to industry structure, including the optimal number of firms, across-industry spectrum allocation, and within-industry spectrum allocation.

Figure 15: Counterfactual welfare by density

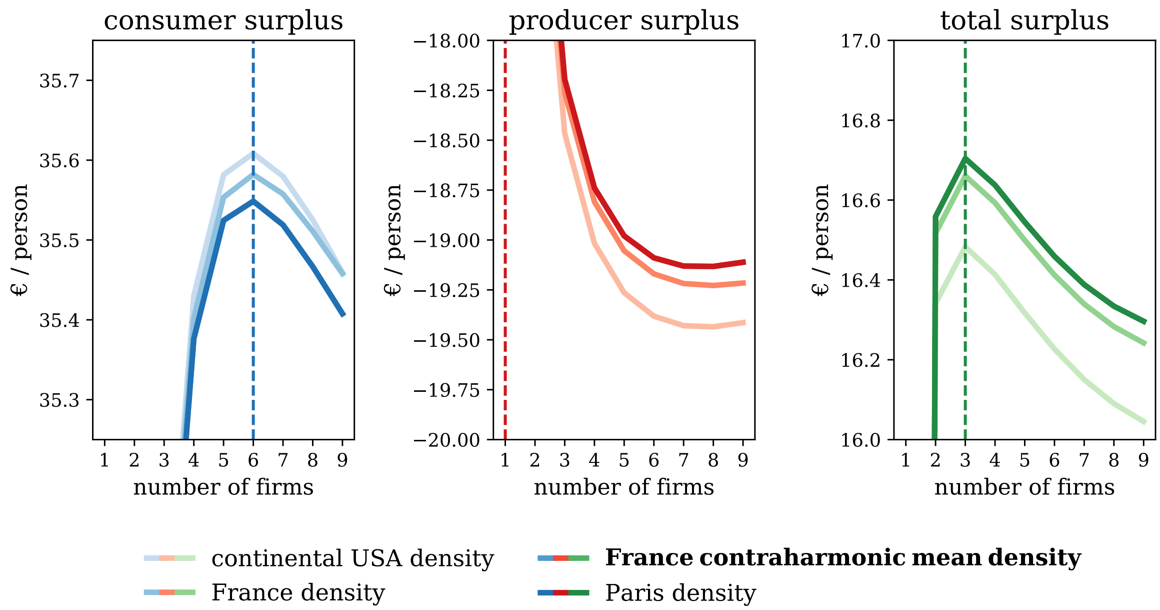
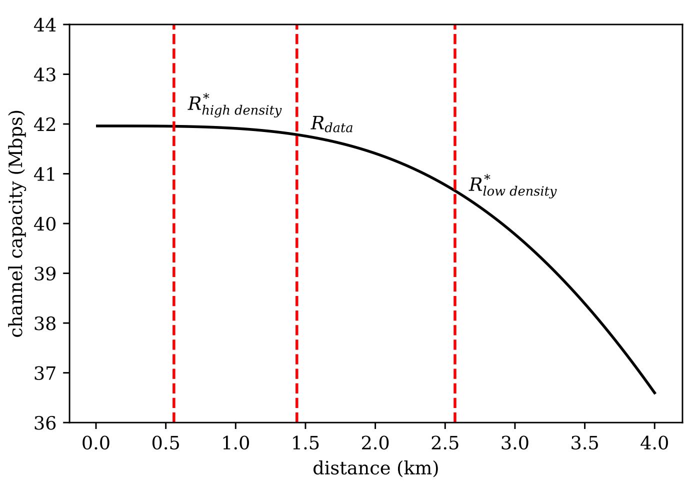


Figure 16: Channel capacity as function of radius



Note: R_{data} corresponds to the average radius of a cell in our data. $R^*_{\text{low density}}$ and $R^*_{\text{high density}}$ correspond to the equilibrium radius chosen in the four-firm equilibrium when the market has, respectively, a density of France and a density equal to the contraharmonic mean density of France. Bandwidth is set equal to the same total bandwidth as in the rest of our counterfactuals divided by four (for four firms), and spectral efficiency is also set to the same value as in the rest of our counterfactuals.

References

Analysys Mason. 2015. "MCT review 2015-2018: Mobile network cost modelling."

Arrow, Kenneth Joseph. 1962. "Economic Welfare and the Allocation of Resources for Inventions." *The Rate of Direction of Inventive Activity: Economic and Social Factors.*

Asker, John, and Michael L Katz. 2022. "The Sprint/T-Mobile Merger." In *Antitrust Economics at a Time of Upheaval.* , ed. John Kwoka, Tommaso Valletti and Lawrence White. Competition Policy International.

Autorité de Régulation des Communications Électroniques (ARCEP). 2011*a*. "Décision no 2011-0597 de l'Autorité de régulation des communications électroniques et des postes."

Autorité de Régulation des Communications Électroniques (ARCEP). 2011*b*. "Décision no 2011-0599 de l'Autorité de régulation des communications électroniques et des postes."

Autorité de Régulation des Communications Électroniques (ARCEP). 2016. "Observatoire des marchés des communications électroniques 4 février 2016."

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–90.

Björkegren, Daniel. 2022. "Competition in network industries: Evidence from the Rwandan mobile phone network." *RAND Journal of Economics*, 53(1): 200–225.

Błaszczyszyn, Bartlomiej, Miodrag Jovanovicy, and Mohamed Kadhem Karray. 2014. "How user throughput depends on the traffic demand in large cellular networks." 611–619, IEEE.

Bourreau, Marc, Yutec Sun, and Frank Verboven. 2021. "Market entry, fighting brands, and tacit collusion: Evidence from the French mobile telecommunications market." *American Economic Review*, 111(11): 3459–99.

Carlton, Dennis W. 1978. "Market behavior with demand uncertainty and price inflexibility." *The American Economic Review*, 68(4): 571–587.

Cattani, Kyle, and Glen M Schmidt. 2005. "The pooling principle." *INFORMS Transactions on Education*, 5(2): 17–24.

Centre De Recherche Pour L'étude Et L'observation Des Conditions De Vie (CREDOC). 2015. "Baromeètre du numérique (Édition 2015)."

Chenery, Hollis B. 1949. "Engineering Production Functions." *The Quarterly Journal of Economics*, 63(4): 507–531.

Chu, Chenghuan Sean. 2010. "The effect of satellite entry on cable television prices and product quality." *The RAND Journal of Economics*, 41(4): 730–764.

Conlon, Christopher, and Jeff Gortmaker. 2020. "Best practices for differentiated products demand estimation with pyblp." *The RAND Journal of Economics*, 51(4): 1108–1161.

Cramton, Peter, Evan Kwerel, Gregory Rosston, and Andrzej Skrzypacz. 2011. "Using spectrum auctions to enhance competition in wireless services." *The Journal of Law and Economics*, 54(S4): S167–S188.

Crawford, Gregory S, and Matthew Shum. 2007. "Monopoly quality degradation and regulation in cable television." *The Journal of Law and Economics*, 50(1): 181–219.

Crawford, Gregory S, Oleksandr Shcherbakov, and Matthew Shum. 2019. "Quality overprovision in cable television markets." *American Economic Review*, 109(3): 956–95.

Crawford, Gregory S, Robin S Lee, Michael D Whinston, and Ali Yurukoglu. 2018. "The welfare effects of vertical integration in multichannel television markets." *Econometrica*, 86(3): 891–954.

Cullen, Joseph, Nicolas Schutz, and Oleksandr Shcherbakov. 2020. "The Welfare Effects of Early Termination Fees in the US Wireless Industry."

De Vany, Arthur. 1976. "Uncertainty, waiting time, and capacity utilization: A stochastic theory of product quality." *Journal of Political Economy*, 84(3): 523–541.

Doraszelski, Ulrich, Katja Seim, Michael Sinkinson, and Peichun Wang. 2019. "Ownership concentration and strategic supply reduction." National Bureau of Economic Research.

El Azouzi, Rachid, Eitan Altman, and Laura Wynter. 2003. "Telecommunications network equilibrium with price and quality-of-service characteristics." In *Teletraffic science and engineering.* Vol. 5, 369–378. Elsevier.

Ershov, Daniel, and David J Salant. 2022. "Auctions and Mobile Market Competition: Evidence from European 4G Auctions." *Available at SSRN 4168166.*

Eső, Péter, Volker Nocke, and Lucy White. 2010. "Competition for scarce resources." *The RAND Journal of Economics*, 41(3): 524–548.

Fan, Ying. 2013. "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market." *American Economic Review*, 103(5): 1598–1628.

Fan, Ying, and Chenyu Yang. 2020. "Competition, product proliferation, and welfare: A study of the US smartphone market." *American Economic Journal: Microeconomics*, 12(2): 99–134.

Federal Communications Commission. 2010. "Connecting America: The National Broadband Plan."

Federal Communications Commission. 2019. "Memorandum opinion and order, declaratory ruling, and order of proposed modification." *WT Docket 18-197.*

Genakos, Christos, Tommaso Valletti, and Frank Verboven. 2018. "Evaluating market consolidation in mobile communications." *Economic Policy*, 33(93): 45–100.

Grajek, Michał, and Lars-Hendrik Röller. 2012. "Regulation and investment in network industries: Evidence from European telecoms." *The Journal of Law and Economics*, 55(1): 189–216.

Granja, João. 2022. "Regulation and Service Provision in Dynamic Oligopoly: Evidence from Mobile Telecommunications."

Grigolon, Laura, and Frank Verboven. 2014. "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation." *Review of Economics and Statistics*, 96(5): 916–935.

Hata, Masaharu. 1980. “Empirical Formula for Propagation Loss in Land Mobile Radio Services.” *IEEE Transactions on Vehicular Technology*, 29(3): 317–325.

Hua, Sha, Pei Liu, and Shivendra S Panwar. 2012. “The urge to merge: When cellular service providers pool capacity.” 5020–5025, IEEE.

Jehiel, Philippe, Benny Moldovanu, Marco Ottaviani, and Carol Propper. 2003. “An economic perspective on auctions.” *Economic Policy*, 18(36): 269–308.

Kim, Haesik. 2015. “Coding and Modulation Techniques for High Spectral Efficiency Transmission in 5G and Satcom.” 2746–2750, IEEE.

Lee, Robin S. 2013. “Vertical integration and exclusivity in platform and two-sided markets.” *American Economic Review*, 103(7): 2960–3000.

Lhost, Jonathan, Brijesh Pinto, and David Sibley. 2015. “Effects of spectrum holdings on equilibrium in the wireless industry.” *Review of Network Economics*, 14(2): 111–155.

Lin, Zhongjian, Xun Tang, and Mo Xiao. 2022. “Endogeneity in Games with Incomplete Information: U.S. Cellphone Service Deployment.”

Milgrom, Paul, and Ilya Segal. 2020. “Clock auctions and radio spectrum reallocation.” *Journal of Political Economy*, 128(1): 1–31.

Mulligan, James G. 1983. “The economies of massed reserves.” 73(4): 725–734.

Nevo, Aviv, John L. Turner, and Jonathan W. Williams. 2016. “Usage-Based Pricing and Demand for Residential Broadband.” *Econometrica*, 84: 411–443.

Nikolikj, Vladimir, and Toni Janevski. 2014. “A cost modeling of high-capacity LTE-advanced and IEEE 802.11ac based heterogeneous networks, deployed in the 700 MHz, 2.6 GHz and 5 GHz Bands.” *Procedia Computer Science*, 40: 49–56.

Peha, Jon M. 2017. “Cellular economies of scale and why disparities in spectrum holdings are detrimental.” *Telecommunications Policy*, 41(9): 792–801.

Rey, Patrick, and David Salant. 2017. “Allocating essential inputs.”

Robinson, Edward Austin Gossage. 1948. “Structure of competitive industry.”

Rosston, Gregory L. 2003. “The long and winding road: the FCC paves the path with good intentions.” *Telecommunications Policy*, 27(7): 501–515.

Seim, Katja, and V Brian Viard. 2011. “The effect of market structure on cellular technology adoption and pricing.” *American economic journal: Microeconomics*, 3(2): 221–51.

Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *Bell Systems Technical Journal*, 27: 379–423,623–656.

Sinkinson, Michael. 2020. “Pricing and entry incentives with exclusive contracts: Evidence from smartphones.”

Smail, Ghoul, and Jia Weijia. 2017. "Techno-economic analysis and prediction for the deployment of 5G mobile network." 9–16, IEEE.

Spence, A. Michael. 1975. "Monopoly, Quality, and Regulation." *Bell Journal of Economics*, 6(2): 417–429.

Sun, Patrick. 2015. "Quality competition in mobile telecommunications: Evidence from Connecticut."

Taylor, H.M., S. Karlin, and H.E. Taylor. 1998. *An Introduction to Stochastic Modeling.* Elsevier Science.

Varadhan, Ravi, and Christophe Roland. 2008. "Simple and globally convergent methods for accelerating the convergence of any EM algorithm." *Scandinavian Journal of Statistics*, 35(2): 335–353.

Vives, Xavier. 2008. "Innovation and competitive pressure." *The Journal of Industrial Economics*, 56(3): 419–469.

Williamson, Oliver E. 1968. "Economies as an Antitrust Defense: The Welfare Tradeoffs." *The American Economic Review*, 58(1): 18–36.

A Technical Appendix (for online publication)

A.1 Data Transmission Details

A.1.1 Signal Power

Equation 30 in Section 3.2 provides the formula we use for signal power. It is based on the Hata model of path loss (Hata, 1980). We use the Hata model for urban environments since we focus our analysis on urbanized areas. This model provides us with the following formula for path loss:

$$L(r) = 68.75 + 27.72 \log_{10}(f) - 13.82 \log_{10}(h) + (44.9 - 6.55 \log_{10}(h)) \log_{10}(r), \tag{29}$$

where $L(r)$ is in decibels, r is the distance from the antenna (in km), h is the height of the base station antenna (in m), and f is the frequency (in MHz).[48]

The specific values in our path loss equation can be derived as follows. We assume a base station height of 30 m and a signal frequency of 1900 MHz, which is approximately the median operated frequency in France in 2015. These values yield

$$L(r) = 139.2232 + 35.2249 \log_{10}(r).$$

The signal power in dBm at a distance r from the antenna is

$$A - L(r),$$

where A is the transmitted power. We assume a signal power of 61 dBm (or 1259 W) per 5 MHz of bandwidth at the base station, which corresponds to the regulated limit on effective isotropic radiated power for the 2600 band (ARCEP, 2011a); similar limits apply for lower frequencies (ARCEP, 2011b).

Converting the units to watts, this yields the following formula for signal power, in W per 5 MHz of bandwidth:

$$S_\ell = S(r_\ell) = \exp(-24.92)\, r_\ell^{-3.522}, \tag{30}$$

[48] In the Sprint/T-Mobile merger, heterogeneity in the merging parties' spectrum holdings played an important role in the claimed efficiency gains. T-Mobile had substantial holdings of low-frequency spectrum, and Sprint owned only high-frequency spectrum (Asker and Katz, 2022). Notice that frequency f enters positively into equation 29, meaning that the signal power of high-frequency spectrum will be lower, and its signal power level will approach the levels of noise and interference power at shorter distances. It would be a straightforward extension to our model to capture this heterogeneity in spectrum holdings using equation 29 and integrating over the appropriate set of frequencies for each firm. Such a model could capture how a firm holding only low-frequency spectrum would experience higher costs of service, especially in areas of low population density.

where r_ℓ is location ℓ's distance from the base station. These values yield a path loss exponent of 3.522. Most engineering studies use a path loss exponent between 3.5 and 4.[49] In contrast, signal strength in a vacuum would have a path loss exponent of 2, but signals decay more quickly on the Earth's surface.

A.1.2 Noise and Interference

Noise power N is set equal to Johnson-Nyquist noise, -107.01 dBm per 5 MHz of bandwidth. Note that our expression for signal strength in equation 30 yields approximately $\exp(-24.92) = 1.5\text{e}{-11}\text{W}$ signal power density at a distance of 1 km from the base station. The noise power density is $10^{(-107.01/10)}/1000 \approx 2\text{e}{-14}\text{W}$. Thus, signal power is orders of magnitude larger than noise power at such distances from a base station.

Interference power is set equal to 30% of the signal power from the six adjacent cells. The 30% number follows Błaszczyszyn, Jovanovicy and Karray (2014) and reflects the fact that adjacent cells won't always be in use, and modern systems use directional signals to limit interference. To illustrate its magnitude, note that the edge of a 1 km cell, at the midpoint between the serving base station and an adjacent identical cell's base station, we would have $1.5\text{e}{-11}\text{W}$ of signal power (per 5 MHz) from the cell being used, and $0.3 \cdot 1.5\text{e}{-11}\text{W}$ of interference power from the immediately adjacent cell's base station. Ignoring interference from other neighboring cells, this would lead to a SINR ratio of approximately 0.3^{-1}, the ratio of signal to interference. At these signal levels, interference dominates the denominator, and noise power plays little role.

Ultimately, the way we calculate interference power (at each point within a cell) is to sum interference from the neighboring six cells, pictured in Figure 17. For a given point in the center cell, we compute the distances between that point and the centroids of the adjacent cells, which is the location of the antennas corresponding to each cell.

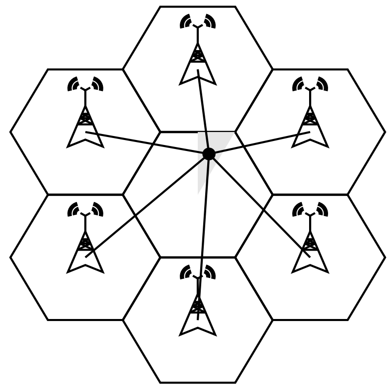
Let $\mathcal{L}$ be the locations of the centroids of the six adjacent hexagons to a hexagon centered at the origin of a Euclidean plane, when all seven hexagons are regular with a (maximum) radius of unity. That is,

$$\mathcal{L} \equiv \left\{ \left(0, \sqrt{3}\right), \left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right), \left(\frac{3}{2}, -\frac{\sqrt{3}}{2}\right), \left(0, -\sqrt{3}\right), \left(-\frac{3}{2}, -\frac{\sqrt{3}}{2}\right), \left(-\frac{3}{2}, \frac{\sqrt{3}}{2}\right) \right\}.$$

These points correspond to the locations of the adjacent base stations pictured in Figure 17. Let $d(\ell, \ell')$ represent the Euclidean distance between two points ℓ and ℓ'. Ultimately,

[49]For instance, Błaszczyszyn, Jovanovicy and Karray (2014) assume a path loss exponent of 3.8.

Figure 17: A hexagonal cell and its six adjacent cells



Note: The figure depicts the distance between an individual at a random location in the center cell and the base stations that correspond to the six adjacent cells. In determining the channel capacity of the cell, we integrate over the entire area of the center cell, taking into account this interference at each point.

interference power is calculated as follows:

$$I_{\ell}\left(R_{fm}\right) = 0.3 \sum_{\ell' \in \mathcal{L}} S\left(R_{fm} d\left(\ell, \ell'\right)\right), \tag{31}$$

where the signal power function $S(\cdot)$ is defined in equation 30 above. In other words, interference power is 30% of the summed signal powers from the adjacent six hexagons.

To calculate channel capacity (equation 10), we need to integrate signal power S_{ℓ} and interference power $I_{\ell}\left(R_{fm}\right)$ over the locations ℓ within a hexagonal cell. To perform this integration, it suffices to focus on one of the twelve right triangles that compose the hexagon and then multiply by twelve (each of the twelve triangles has the same distribution of interference). Specifically, we integrate over the shaded triangle in Figure 17:

$$\bar{Q}_{fm}\left(R_{fm}, B_{fm}\right) = \gamma_m B_{fm} A\left(R_{fm}\right) \left[12 \int_0^{\frac{\sqrt{3}}{2} R_{fm}} \int_0^{\frac{y}{\sqrt{3}}} \frac{1}{\log_2\left(1 + \frac{S_{(x,y)}}{N + I_{(x,y)}\left(R_{fm}\right)}\right)} \mathrm{d}x \mathrm{d}y \right]^{-1}. \tag{32}$$

A.2 Equilibrium without Path Loss

Here we show that in symmetric equilibria the optimal number of base stations per consumer is constant when there is no path loss or interference.

Let N_{fm} represent the number of base stations operated by operator f in municipality m. The number of consumers within each cell is given by $\frac{D_m A_m}{N_{mf}}$, where D_m is the population density and A_m is the municipality's area. We now rewrite equation 15 as

$$Q_{fm} = \overline{Q}_{fm} - \frac{D_m A_m}{N_{mf}} q^D \left(\boldsymbol{P}_{fm}, \boldsymbol{Q}_{fm}, \boldsymbol{P}_{-fm}, \boldsymbol{Q}_{-fm} \right), \tag{33}$$

where $q^D \left(\boldsymbol{P}_{fm}, \boldsymbol{Q}_{fm}, \boldsymbol{P}_{-fm}, \boldsymbol{Q}_{-fm} \right)$ represents equilibrium data consumption per capita. Note that channel capacity per base station $\overline{Q}_{fm}$ is exogenous without path loss and interference. Bandwidth is endowed, so there are no choice variables to influence channel capacity. The firm's only infrastructure choice here is effectively how many consumers they want to serve with each base station.

Consider firm f's variable profit function, equation 17, now written in per-consumer terms and as a function of quality:

$$\Pi^V_{fm} \left(\boldsymbol{P}_f, \boldsymbol{Q}_{fm} \right) \equiv \left(\boldsymbol{P}_f - \boldsymbol{c}^u_f \right) \cdot \boldsymbol{s}_f \left(\boldsymbol{P}_{fm}, \boldsymbol{Q}_{fm}, \boldsymbol{P}_{-fm}, \boldsymbol{Q}_{-fm} \right).$$

Let $\lambda_{fm} = \frac{D_m}{N_{fm}}$, and note that λ_{fm} can represent the firm's infrastructure choice variable. Rewrite variable profits as

$$\Pi^V_{fm} \left(\boldsymbol{P}_f, \lambda_{fm} \right) \equiv \left(\boldsymbol{P}_f - \boldsymbol{c}^u_f \right) \cdot \boldsymbol{s}_f \left(\boldsymbol{P}_{fm}, \lambda_{fm}, \boldsymbol{P}_{-fm}, \boldsymbol{\lambda}_{-fm} \right),$$

noting that the share function can be expressed as a function of λ_{fm} since delivered download speeds are determined by the congestion equation 33, and here $\lambda_{fm} = \frac{D_m}{N_{fm}}$ defines the congestion equation above.

Given the cost function expressed in equation 18, infrastructure costs are $c^s_{fm} B_{fm} N_{fm}$, and costs per capita can be expressed as

$$c^s_{fm} B_{fm} \frac{N_{fm}}{D_m A_m} = c^s_{fm} B_{fm} \lambda^{-1}_{fm} A^{-1}_m.$$

Both variable profits and infrastructure costs depend on population density D_m and the number of base stations N_{fm} only through their ratio $\lambda_{fm} = \frac{D_m}{N_{fm}}$. Therefore, the firm's optimum and the equilibrium level of investment entail a value for λ, or a number of base stations per consumer. Therefore, when we do comparative statics with respect to population density, the equilibrium number of base stations will be proportional to population density.

B Demand Estimation Details (for online publication)

B.1 Contraction Mapping

Here we consider an alternative version of the Berry, Levinsohn and Pakes (1995) (BLP) contraction mapping in which we observe market shares at the product-market level for Orange products but only aggregate firm-level market shares for the other products. We first show in Section B.1.2 that if we observe market shares at the firm-market level, the problem can be rewritten in such a way that the BLP contraction mapping proof holds. In Section B.1.3 we extend this result to the nested logit setting. Finally, in Section B.1.4 we show that if we observe some firm market shares only at the aggregate level (as is our case), the problem can still be rewritten to fit into the BLP contraction mapping proof setup.

B.1.1 Standard BLP Contraction Mapping Setup

We will start with the standard BLP setting in order to introduce notation. In this setting, there are products $j \in \mathcal{J} = \{1, \ldots, J\}$, and we observe market shares ς_{jm} for each product. We can express an individual's utility for a product as $u_{ijm} = \delta_{jm} + \mu_{ijm} + \varepsilon_{ijm}$, which yields the type-specific market shares

$$s_{ijm} = \frac{\exp\left(\delta_{jm} + \mu_{ijm}\right)}{\sum_{j'} \exp\left(\delta_{j'm} + \mu_{ij'm}\right)}.$$

Aggregate market shares are given by

$$s_{jm}\left(\delta\right) = \int \frac{\exp\left(\delta_{jm} + \mu_{ijm}\right)}{\sum_{j'} \exp\left(\delta_{j'm} + \mu_{ij'm}\right)} dF\left(\mu_m\right).$$

The existence of the contraction mapping implies that there is a unique vector δ such that $s_m\left(\delta\right) = \varsigma_m$ for any observed vector of shares ς_m.

B.1.2 Grouped Products Extension

Our setting is one in which market shares are observed only for certain groupings of products. That is, let $\mathcal{J}$ be partitioned into subsets $\mathcal{J}_f$ with $f \in \mathcal{F} = \{1, 2, \ldots F\}$. For each f, we observe only the market share ς_{ft} for all the products within $\mathcal{J}_f$. The subsets $\mathcal{J}_f$ may include individual products (i.e., in our application each Orange product would have its own $\mathcal{J}_f$ set) or several products (i.e., each non-Orange firm has one $\mathcal{J}_f$ group that includes all that firm's products).

Providing a parametric form, let $\delta_{jm} = \theta_1 x_{jm} + \xi_{jm}$, where θ_1 would capture what is often referred to as "linear parameters," i.e., parameters that can typically be estimated outside of the contraction mapping because they only shift the mean utility component δ_{jm} that the contraction mapping aims to recover. In this extension, the θ_1 parameters must be included in the contraction mapping.

We cannot recover δ_{jm} (or ξ_{jm}) separately for different $j \in \mathcal{J}_f$. We assume $\xi_{jm} = \xi_{fm}$ for all $j \in \mathcal{J}_f$ for each f.

Let $\bar{x}_{fm}$ be the mean value of x_{fm} for those products within $\mathcal{J}_f$. Then, we have $\delta_{jm} = \theta_1 \bar{x}_{fm} + \theta_1 x^d_{jm} + \xi_{fm}$, where $x^d_{jm} := x_{jm} - \bar{x}_{fm}$. We define $\widetilde{\delta}_{fm} = \theta_1 \bar{x}_{fm} + \xi_{fm}$, and $\widetilde{\mu}_{ijm} = \theta_1 x^d_{jm} + \mu_{ijm}$. This very nearly allows us to re-define the model in terms where we could apply the original BLP proof strategy to establish the contraction mapping. The only problem is that $\widetilde{\mu}_{ijm}$ is defined over j, where we would need it to be defined over f in order to apply the same proof strategy. Let's consider the aggregation over j to f:

$$s_{ifm}\left(\widetilde{\delta}\right) = \sum_{j \in \mathcal{J}_f} \frac{\exp\left(\widetilde{\delta}_{fm} + \widetilde{\mu}_{ijm}\right)}{\sum_{j' \in \mathcal{J}} \exp\left(\widetilde{\delta}_{f(j')m} + \widetilde{\mu}_{ij'm}\right)},$$

where $f(j')$ refers to the f associated with product j'.

Defining $\widetilde{\mu}_{ifm} = \log\left(\sum_{j \in \mathcal{J}_f} \exp\left(\widetilde{\mu}_{ijm}\right)\right)$, it follows that

$$\sum_{j \in \mathcal{J}_f} \exp\left(\widetilde{\delta}_{fm} + \widetilde{\mu}_{ijm}\right) = \exp\left(\widetilde{\delta}_{fm} + \widetilde{\mu}_{ifm}\right),$$

and therefore

$$s_{ifm}\left(\widetilde{\delta}\right) = \sum_{j \in \mathcal{J}_f} \frac{\exp\left(\widetilde{\delta}_{fm} + \widetilde{\mu}_{ifm}\right)}{\sum_{f'} \exp\left(\widetilde{\delta}_{f'm} + \widetilde{\mu}_{if'm}\right)}.$$

We can then aggregate up to market-level shares s_{fm} by integrating over the $\widetilde{\mu}_{ifm}$, and we have rewritten our extended setting in a way that allows us to apply the BLP proof strategy.

B.1.3 Grouped Products Extension with Nested Logit

In the more general random coefficients nested logit (RCNL) model introduced by Grigolon and Verboven (2014) (henceforth, GV), we can construct analogous formulas that will allow us to recover group-specific mean demands $\widetilde{\delta}$.

In the RCNL model, type-specific market shares are as follows:

$$s_{ijm} = \frac{\exp\left(\frac{\delta_{jm} + \mu_{ijm}}{1-\sigma}\right)}{\exp\left(\frac{I_{ig(j)}}{1-\sigma}\right)} \frac{\exp\left(I_{ig(j)}\right)}{\exp\left(I_i\right)},$$

where $\sigma \in [0,1)$ is the nesting parameter, $g(j)$ return the nest to which j belongs,[50] and

$$\begin{aligned} I_{ig} &= (1-\sigma)\log\left(\sum_{j\in\mathcal{J}_g}\exp\left(\tfrac{\delta_{jm}+\mu_{ijm}}{1-\sigma}\right)\right), \\ I_i &= \log\left(1+\sum_{g\in\mathcal{G}}\exp\left(I_{ig}\right)\right). \end{aligned}$$

In this extension, we redefine $\widetilde{\delta}_{fm}$ and $\widetilde{\mu}_{ifm}$ to incorporate σ. Let $\widetilde{\delta}_{fm} = \frac{\theta_1\bar{x}_{fm}+\xi_{fm}}{1-\sigma}$, $\widetilde{\mu}_{ijm} = \frac{\theta_1 x^d_{jm}+\mu_{ijm}}{1-\sigma}$, and $\widetilde{\mu}_{ifm} = \log\left(\sum_{j\in\mathcal{J}_f}\exp\left(\widetilde{\mu}_{ijm}\right)\right)$. Then

$$s_{ifm} = \frac{\exp\left(\widetilde{\delta}_{fm}+\widetilde{\mu}_{ifm}\right)}{\exp\left(\frac{I_{ig(f)}}{1-\sigma}\right)}\frac{\exp\left(I_{ig(f)}\right)}{\exp\left(I_i\right)}$$

where $I_{ig} = (1-\sigma)\log\left(\sum_{f\in\mathcal{F}_g}\exp\left(\widetilde{\delta}_{fm}+\widetilde{\mu}_{ifm}\right)\right)$ and $\mathcal{F}_g = \{f\in\mathcal{F} : g(f) = g\}$.

GV note that, substituting in our notation,

$$f\left(\widetilde{\delta}\right) = \widetilde{\delta} + \log(\varsigma) - \log\left(s\left(\widetilde{\delta}\right)\right)$$

is a contraction mapping if

$$1 - \frac{1}{s_f}\frac{\partial s_f}{\partial\widetilde{\delta}_f} \geq 0.$$

Unlike in GV, this holds in our case. Explicitly,

$$\frac{\partial s_f}{\partial\widetilde{\delta}_f} = \left(1 - \frac{\sigma}{1-\sigma}s_{f|g} - s_f\right)s_f,$$

and so

$$1 - \frac{1}{s_f}\frac{\partial s_f}{\partial\widetilde{\delta}_f} = \frac{\sigma}{1-\sigma}s_{f|g} + s_f \geq 0 \quad \Leftrightarrow \quad \sigma s_{f|g} + (1-\sigma)s_f \geq 0.$$

This condition holds for all $\sigma \in [0,1)$.

B.1.4 Market Aggregation Extension

In our setting we observe market shares only at the aggregate level for some firms. We assume in this extension $\xi_{jm} = \xi_{f(j)}$ for all j, m and recover ξ_f for each f. We will proceed in this section using the non-nested setting introduced in Section B.1.2, but the results hold using the analogues to the RCNL expressions introduced in Section B.1.3.

Analogous to the previous setup, let $\bar{x}_f$ be the mean value of x_{jm} across products $j \in \mathcal{J}_f$ and

[50] We will assume that products produced by the same firm belong to the same group. Formally, for each f, $g(j) = g_f$ for all $j \in \mathcal{J}_f$.

markets m, $\bar{x}_f = \frac{1}{MJ_f}\sum_m \sum_{j\in\mathcal{J}_f} x_{jm}$. Then, $\delta_{jm} = \theta_1 \bar{x}_{f(j)} + \theta_1 x^d_{jm} + \xi_{f(j)}$. where we now define $x^d_{jm} := x_{jm} - \bar{x}_{f(j)}$. Analogously defining $\widetilde{\delta}_f = \theta_1\bar{x}_f + \xi_f$, $\widetilde{\mu}_{ijm} = \theta_1 x^d_{jm} + \mu_{ijm}$, and $\widetilde{\mu}_{ifm} := \log\left(\sum_{j\in\mathcal{J}_f}\exp\left(\widetilde{\mu}_{ijm}\right)\right)$, then

$$\bar{s}_{if}(\widetilde{\delta}) = \sum_m w(m)\frac{\exp\left(\widetilde{\delta}_f + \widetilde{\mu}_{ifm}\right)}{\sum_{f'}\exp\left(\widetilde{\delta}'_f + \widetilde{\mu}_{if'm}\right)}.$$

We can aggregate up to aggregate firm shares $\bar{s}_f$ by integrating over $\tilde{\mu}_{ifm}$:

$$\bar{s}_f = \int\sum_m w(m)\frac{\exp\left(\widetilde{\delta}_f + \widetilde{\mu}_{ifm}\right)}{\sum_{f'}\exp\left(\widetilde{\delta}_{f'} + \widetilde{\mu}_{if'm}\right)}dF(\widetilde{\mu}_{ifm}) = \int\frac{\exp\left(\widetilde{\delta}_f + \widetilde{\mu}_{ifm}\right)}{\sum_{f'}\exp\left(\widetilde{\delta}_{f'} + \widetilde{\mu}_{if'm}\right)}dG(\widetilde{\mu}_{ifm}).$$

The final expression makes clear that the BLP contraction mapping proof strategy still holds in this aggregate setting.

When coding the contraction mapping, we follow Conlon and Gortmaker (2020) in implementing the SQUAREM algorithm (Varadhan and Roland, 2008).

B.2 Implementation Details

The setup outlined in Section B.1.4 is more restrictive than is necessary given our data. We observe product-level market shares for every market for Orange products. We therefore allow ξ_{jm} to differ by product and market for all $j\in\mathcal{J}_O$, where O denotes Orange.

The moments used in our GMM estimation procedure, listed in Section 4.1.3, are imposed only for Orange products. To center Orange demand shocks, we add an Orange dummy variable O_j defined as follows

$$O_j = \begin{cases} 1 & \text{if } f(j) = \text{Orange} \\ 0 & \text{otherwise,}\end{cases}$$

and O_j enters utility additively so that Equation 1 becomes

$$v\left(j,x,m;\theta_i,\vartheta_i,\varepsilon_i\right) \equiv u_j\left(x, Q_{m,f(j)};\vartheta_i,\theta_i\right) + \theta_v v_j - \theta_{pi}p_j + \theta_O O_j + \xi_{jm} + \varepsilon_{ij}.$$

The inclusion of the term $\theta_O O_j$ allows Orange products to differ in a systematic way from the products offered by other firms, restoring the validity of moments of the form presented in Section 4.1.3. To identify the parameter θ_O, we impose the following additional moment

$$\mathbb{E}\left[\xi_{jm}\left(\theta\right)O_j\right] = 0.$$

To ensure the correct sign for θ_c (which must be positive) while searching over the space of demand parameters, we search for $\log(\theta_c)$ rather than θ_c directly.[51]

Incomes are in units of 10 000 €. Data limits are in GB and quality measures are in GBps.[52]

C Data Appendix (for online publication)

This appendix provides additional description of our main datasets and variables. Section C.1 presents the characteristics of mobile tariffs and the tariff dataset. Section C.2 describes the Orange customer dataset and socioeconomic characteristics. Section C.3 describes the measurement of the quality of mobile data.

C.1 Product Data

C.1.1 Product Characteristics

We collect data on mobile phone plans released between November 2013 and October 2015, along with their characteristics, from operators' quarterly catalogs. It includes postpaid plans from the four MNOs and the largest MVNO (EI Telecom) as well as their prepaid plans.[53] Promotional plans, typically released during summer and Christmas, are not included in the dataset.

Plan characteristics include tariff, voice and data limits, international voice or data roaming, handset subsidy, length of commitment, and whether or not plans were bundled with fixed services. As described in Section 2.2, we choose representative mobile-only plans for each firm and adjust monthly prices based on contract duration and handset subsidies.

We take over 100 contracts from catalogs, and from them we construct 21 representative products in our model's choice set. We define categories of plans according to their level of data limits: less than 500 MB, 500–3 000 MB, 3 000–7 000 MB and more than 7 000 MB. These thresholds are chosen following discussions with industry experts and the statistical distribution of chosen plans. The second data limit category—that is, contracts with 500–3 000 MB—we have further split according to their voice allowances: unlimited or not, making a total of five categories of phone plans. Low data limit plans typically do not have unlimited voice, and high data limit contracts typically come with unlimited voice allowance, so we do

[51] The value reported in the demand estimates, Table 9 in Appendix D.1 is therefore the estimate of $\log(\theta_c)$.

[52] Note that quality measures are in Giga*bytes* per second (GBps), not Giga*bits* per second (Gbps). This conversion is needed so that the second term in Equation 3 has the interpretation of seconds spent downloading data.

[53] ORG's contracts include not only those that are sold through its main brand, but also others sold under alternative brands such as SOSH, BNP Paribas Mobile, FNAC Mobile, Click Mobile, Carrefour Mobile, etc.

not split these categories by the voice limit. We exclude plans bundled with fixed broadband or television.

We choose the least expensive plan in each category as the category's representative plan. Some customers keep old plans that are no longer available, so we fill these missing data by using the most similar representative plan. While some plans with handset subsidies have corresponding standalone versions, some do not. We adjust the prices of these latter plans using data on the price of handsets and the upfront payment required by Orange. We collect these data for both iPhone and Samsung, the two most popular handsets. We then distribute the handset cost over 24 months and update the monthly plan price by subtracting off the monthly cost of the handset. In addition, we assume that Orange's handset subsidies apply to other operators' subsidized contracts because we do not observed their upfront costs.

C.1.2 Soft Data Limits

For plans with data limits, the download speed is reduced for usage above allowance if no add-on is purchased. The maximal download speed under throttling is typically 128 Kbps. With this download speed, it would take over half-an-hour to download a 30 MB file, compared to 2 minutes under a theoretical non-throttled speed of 2 Mbps in a 3G network, and 24 seconds given a moderate 4G download speed of 10 Mbps. Basically, only emails and light web pages can be opened under throttling. As presented in Table 7 below, this download speed is not always specified by operators in their contracts. When it is, it may depend on the location of the usage (local or abroad). The actual download speed experienced by customers is a function of the number of simultaneous users, its location and handset. In our demand model, however, we assume that any data consumption over the data limit yields a speed of exactly 128 Kbps.

Table 7: Maximal download speed under throttling (Kbps)

Operator	National	Roaming
ORG	128*	ns
SFR	ns	ns
BYT	128	32
FREE	ns	ns

*:except video streaming.

ns ≡ not specified.

Source: operators' contracts

C.2 Consumer Data

C.2.1 Mean Data Consumption

We use the Orange customer data to construct market-level measures of mean data consumption for each Orange phone plan. Note that because we only observe data consumption for consumers of Orange plans, we cannot construct these measures for plans of other firms. Plans are aggregated based on the associated data limit and whether or not the voice allowance is unlimited, as detailed in Section 3.1. Constructing market-plan-level measures of mean data consumption is complicated by the fact that the aggregated plans in the choice set incorporate plans with different data limits. For example, the Orange 4 000 MB data limit plan in the choice set incorporates plans in the customer data with data limits ranging from 3 000 MB to 7 000 MB.

Since we use the mean data consumption in the data to discipline the predicted data consumption in our demand model, which is based on the data limit from the choice set, simply averaging the data consumption observed in the customer data could lead to biased estimates in the data consumption coefficients. For example, using the same 4 000 MB aggregated plan as before, if many customers in this category have plans with data limits above 4 000 MB, they may consume well above 4 000 MB without hitting their data limit. Simply averaging data consumption for this category might give mean data consumption above 4 000 MB, which our demand estimation would interpret as either being insensitive to download speeds (because they are willing to consume even at the very slow throttled speed) or heavily weight the amount of data consumed (because they are consuming large amounts of data despite the slow throttled speed). In fact, it might be that neither of those conclusions is consistent with consumers' data consumption decisions under their actual data limit.

In order to account for the fact that realized data consumption decisions reflect heterogeneous data limits within a single data limit category, we define (adjusted) mean data consumption as follows:[54]

$$\bar{x}_{jm} = \frac{1}{|\mathcal{I}_{jm}|} \sum_{i \in \mathcal{I}_j} \min \left\{ \frac{x_i}{\bar{x}_i}, 1 \right\} \bar{x}_j + \max \left\{ 0, x_i - \bar{x}_i \right\},$$

where $\mathcal{I}_{jm}$ is the set of consumers with plans that aggregate to j in market m, x_i is consumer i's data consumption, and $\bar{x}_i$ is the data limit of their plan. The value $\bar{x}_j$ is the data limit associated with the representative plan j. We separate these two terms rather than simply

[54]For contracts belonging to the group characterized by data limits of less than 500 MB, we impose that consumption cannot be greater than the data limit. For this category of contracts, add-on data packages are a common way of increasing one's data limit. Since we do not observe data package purchases, we simply assume that any consumer that consumed above the data limit did so with a purchased data package and that without one, she would have consumed as much as the data limit allowed. Our demand model reflects this, imposing that contracts in this category cannot consume above the data limit at a reduced speed (as they are able to do for high data limit contracts).

using the fraction of the data limit consumed times the representative plan's data limit because, conditional on bypassing the data limit, the data limit is irrelevant for further data consumption.

C.2.2 Socioeconomic Data

Socioeconomic characteristics are generated from the 2011 population census conducted by the French office of statistics (INSEE). These statistics include the deciles of income at the municipality level. Income is measured as the fiscal revenue of households living in a given municipality in 2011.

C.3 Quality Data

Quality measures are constructed using download speed test results provided by Ookla. Test results come from users who use Ookla's free Internet speed test, called "Speedtest," using a web browser or within an app. Using speed tests in France in the fourth quarter of 2015 yields 1 056 285 individual speed tests. Each speed test records the download speed, mobile network operator, and the user's location. We aggregate speed tests by averaging measured download speeds over tests for a given operator and geographic market, yielding an operator-market quality measure. An operator-market quality measure is, on average, an average of 284 test results. Note that our estimates rely on an instrument for these quality measures (see Section 4.1.3), alleviating concerns about attenuation bias.

C.4 Network Sharing

Network sharing occurs when a network operator shares a part or the whole of its network resources with a retail competitor. These resources can be passive network elements, such as antenna supports, masts, or active network elements, such as frequency bandwidths. Passive network sharing affects coverage differentiation but not necessarily quality differentiation. It typically consists of operators sharing the same tower and potentially the cost of electricity. In general, it is any agreement between MNOs that do not involve the sharing of available frequency bandwidths.

In contrast, under active network sharing (Radio Access Network-Sharing), operators cannot differentiate in terms of quality, defined as the frequency bandwidth available per customer. Typically, it consists of the sharing of frequency bands and the network elements involved in data transmission. Roaming agreements, whereby an operator's customers rely on the network of a host operator to communicate, is the highest level of active network sharing. It does not offer any possibility for quality or coverage differentiation.

Table 8 below presents the network sharing agreements reached between 2012 and 2015. These agreements apply to two types of areas according to their population density. “White Areas” or “Zones Blanches” correspond to areas where population density is so low that network deployment by several operators is not profitable. These areas, which are typically rural, are designated by the regulator and represent roughly 1% of the population and 10% of the national surface. Only ORG, SFR and BYT have invested in these areas.

The most widespread network technologies in the White Areas are 2G, EDGE and GPRS. [55] However, 3G technology has been recently deployed. As of the end of December 2015, half of ORG and BYT’s networks in these areas were covered by 3G, compared to 35% for SFR. In general, only one operator invests in a given White Area, and 64% of antennas in these areas are involved in a roaming agreement. Rival operators roam over the network of the only operator that invests in the area. As a result, there is no quality differentiation. For the remaining 36% of antennas, operators share passive network elements.

At the national level, FREE’s customers can roam over ORG’s 2G and 3G networks as long as there is no FREE antenna nearby. As a result, FREE cannot differentiate from ORG on 2G and 3G technologies, except when a FREE antenna is nearby its customer. In addition, FREE does not have access to networks in Zones Blanches where BYT or SFR is the leader. MVNOs have roaming agreements with their hosts and therefore cannot differentiate in terms of quality or coverage.

Our model focuses on high-density areas to avoid the need to explicitly model network sharing. During our period of study, the only active network sharing in such areas would have involved FREE’s customers receiving data from 2G and 3G infrastructure owned and operated by ORG. Meanwhile, ORG and FREE each owned and operated their own distinct 4G network infrastructure, At the margin, 4G investments were how firms were differentiating and competing in download speeds in 2015.

Table 8: Network sharing agreements 2012-2015

		FREE		ORG		SFR		BYT
Zone Blanche	Roaming: 64% of 2G & 3G antenna				↔			
	Passive sharing: 36% of antenna				↔			
Low Density	2G and 3G RAN-Sharing	✗		✗			↔	
	4G Roaming	✗		✗			→	
High Density		✗		✗		✗		✗
National	Passive sharing				↔			
	2G and 3G Roaming		→			✗		✗

Source: Summary from discussions with ORG’s experts.

Note: ↔: two-way (reciprocal) sharing, $A \rightarrow B$ one-way sharing hosted by operator B.

[55] EDGE and GPRS are suitable for low speed mobile data services.

D Supplementary Results (for online publication)

D.1 Demand Estimation Results

Table 9: Demand Parameter Estimates

Elasticity	Nesting Parameter	$\hat{\theta}_{p0}$	$\hat{\theta}_{pz}$	$\hat{\theta}_{v}$	$\hat{\theta}_{O}$	$\hat{\theta}_{d0}$	$\hat{\theta}_{dz}$	$\hat{\theta}_{c}$
−3.2	0.0	−0.343	−0.707	1.75	3.819	−1.308	0.311	1.164e−3
		(0.393)	(0.167)	(0.056)	(0.554)	(0.09)	(0.053)	(1.130e−4)
	0.5	−1.081	−0.689	0.872	2.881	−0.602	0.306	5.845e−4
		(0.463)	(0.21)	(0.064)	(0.399)	(0.091)	(0.053)	(8.839e−5)
	0.75	−1.809	−0.674	0.435	2.581	0.093	0.303	2.951e−4
		(0.661)	(0.325)	(0.097)	(0.298)	(0.095)	(0.063)	(6.792e−5)
	0.85	−2.326	−0.673	0.261	2.508	0.602	0.303	1.777e−4
		(0.97)	(0.502)	(0.115)	(0.257)	(0.097)	(0.083)	(5.778e−5)
−2.5	0.0	−0.549	−0.767	1.505	3.019	−0.741	0.312	6.667e−4
		(0.462)	(0.209)	(0.043)	(0.506)	(0.106)	(0.052)	(1.123e−4)
	0.5	−1.269	−0.758	0.755	2.542	−0.039	0.308	3.346e−4
		(0.553)	(0.267)	(0.064)	(0.37)	(0.142)	(0.055)	(8.215e−5)
	0.75	**−1.976**	**−0.753**	**0.378**	**2.435**	**0.653**	**0.307**	**1.687e−4**
		(0.822)	**(0.423)**	**(0.098)**	**(0.291)**	**(0.172)**	**(0.067)**	**(5.881e−5)**
	0.85	−1.357	−1.331	0.281	2.783	1.452	0.338	6.631e−5
		(0.969)	(0.406)	(0.043)	(0.367)	(0.096)	(0.053)	(6.299e−6)
−1.8	0.0	−0.756	−0.89	1.274	2.318	0.57	0.314	1.812e−4
		(0.605)	(0.3)	(0.032)	(0.471)	(0.512)	(0.053)	(1.125e−4)
	0.5	−1.452	−0.894	0.643	2.241	1.275	0.312	9.022e−5
		(0.719)	(0.375)	(0.061)	(0.349)	(0.673)	(0.057)	(7.275e−5)
	0.75	−2.158	−0.891	0.322	2.295	1.945	0.311	4.642e−5
		(1.054)	(0.579)	(0.094)	(0.272)	(0.805)	(0.072)	(4.674e−5)
	0.85	−1.925	−1.305	0.217	2.493	3.418	0.335	9.719e−6
		(1.295)	(0.605)	(0.064)	(0.315)	(0.639)	(0.06)	(7.505e−6)

The row in bold corresponds to the imputed elasticity and nesting parameter presented in the main text. Rather than estimating θ_c directly, we estimate $\log(\theta_c)$ since $\theta_c > 0$. The estimates in the final column for θ_c are the exponentiated values of our estimates for this parameter, in scientific notation, and we derive the associated standard errors using the Delta Method.

Demand parameter estimates are listed in Table 9 for a range of imputed price elasticities and nesting parameters. To interpret these estimates, we convert the parameter estimates into willingness to pay for certain contract characteristics. Consumers' willingness to pay varies considerably across income levels, as we allow the price and data consumption parameters (θ_p and θ_d, respectively) to vary by income, so we present these results across income percentiles.[56] We present tables capturing consumers' willingness to pay for higher data limits (Table 10), for an unlimited voice allowance (Table 11), and for higher download speeds (Table 12).

[56] Each percentile corresponds to the estimated willingness to pay for an individual with an income that is the average of that percentile across all markets in our sample. Specifically, the 10th percentile is 3 759 €, the 30th percentile is 8 705 €, the 50th percentile is 13 015 €, the 70th percentile is 18 101 €, and the 90th percentile is 28 096 €.

Table 10: Willingness to pay to go from 1 000 MB data plan to 4 000 MB plan

Elasticity	**Nesting Parameter**	10th %ile	30th %ile	50th %ile	70th %ile	90th %ile
−3.2	0.0	3.77 €	4.34 €	4.81 €	5.36 €	6.40 €
	0.5	3.86 €	4.40 €	4.86 €	5.39 €	6.38 €
	0.75	3.95 €	4.49 €	4.93 €	5.44 €	6.38 €
	0.85	3.98 €	4.51 €	4.95 €	5.46 €	6.38 €
−2.5	0.0	2.67 €	3.16 €	3.59 €	4.11 €	5.17 €
	0.5	2.70 €	3.19 €	3.62 €	4.13 €	5.18 €
	0.75	**2.74 €**	**3.23 €**	**3.65 €**	**4.17 €**	**5.21 €**
	0.85	0.84 €	1.31 €	1.85 €	2.71 €	5.54 €
−1.8	0.0	0.92 €	1.16 €	1.38 €	1.67 €	2.33 €
	0.5	0.91 €	1.15 €	1.38 €	1.67 €	2.35 €
	0.75	0.94 €	1.19 €	1.42 €	1.72 €	2.42 €
	0.85	0.20 €	0.31 €	0.44 €	0.63 €	1.26 €

Table 11: Willingness to pay for unlimited voice allowance

Elasticity	**Nesting Parameter**	10th %ile	30th %ile	50th %ile	70th %ile	90th %ile
−3.2	0.0	3.16 €	4.53 €	6.05 €	8.39 €	15.97 €
	0.5	3.27 €	4.65 €	6.16 €	8.48 €	15.87 €
	0.75	3.35 €	4.74 €	6.24 €	8.53 €	15.78 €
	0.85	3.38 €	4.77 €	6.27 €	8.57 €	15.81 €
−2.5	0.0	3.40 €	5.04 €	6.89 €	9.84 €	19.79 €
	0.5	3.49 €	5.15 €	7.02 €	9.97 €	19.89 €
	0.75	**3.55 €**	**5.22 €**	**7.10 €**	**10.07 €**	**20.01 €**
	0.85	1.73 €	3.43 €	5.91 €	10.95 €	36.82 €
−1.8	0.0	3.70 €	5.83 €	8.38 €	12.66 €	28.49 €
	0.5	3.75 €	5.93 €	8.54 €	12.92 €	29.19 €
	0.75	3.80 €	6.00 €	8.64 €	13.05 €	29.38 €
	0.85	2.34 €	4.58 €	7.79 €	14.27 €	46.87 €

Table 12: Willingness to pay for increase from 10 Mbps to 20 Mbps

Elasticity	**Nesting Parameter**	10th %ile	30th %ile	50th %ile	70th %ile	90th %ile
−3.2	0.0	2.90 €	3.59 €	4.20 €	4.97 €	6.67 €
	0.5	2.98 €	3.66 €	4.27 €	5.02 €	6.65 €
	0.75	3.07 €	3.75 €	4.34 €	5.08 €	6.66 €
	0.85	3.10 €	3.77 €	4.37 €	5.10 €	6.67 €
−2.5	0.0	2.06 €	2.63 €	3.15 €	3.83 €	5.41 €
	0.5	2.10 €	2.66 €	3.19 €	3.86 €	5.43 €
	0.75	**2.13 €**	**2.70 €**	**3.23 €**	**3.90 €**	**5.47 €**
	0.85	0.61 €	1.04 €	1.57 €	2.46 €	5.70 €
−1.8	0.0	0.71 €	0.97 €	1.22 €	1.56 €	2.45 €
	0.5	0.71 €	0.96 €	1.22 €	1.57 €	2.48 €
	0.75	0.74 €	1.00 €	1.26 €	1.62 €	2.55 €
	0.85	0.15 €	0.25 €	0.38 €	0.58 €	1.31 €

Table 10 presents consumers' willingness to pay for an increase from a 1 000 MB plan to a 4 000 MB plan, with quality equal to the median download speed observed in our data (24.3 Mbps). Higher income consumers are willing to pay considerably more for this upgrade than are lower income consumers. From the estimates corresponding to our preferred imputation (the row in bold), a consumer with an income equal to the 90th percentile would be willing to pay 5.21 € for the upgrade, while a consumer with an income equal to the 10th percentile would only be willing to pay 2.74 €. These differences reflect that the estimated price parameter is decreasing in magnitude in income while the data consumption parameter is increasing. Estimates of willingness to pay are pretty stable across choices of the nesting parameter, while unsurprisingly vary in levels across imputed price elasticities. Patterns across income levels are broadly consistent across imputed parameters, however.

Table 11 presents willingness to pay for an unlimited voice allowance. From the estimates corresponding to our preferred imputation (the row in bold), a consumer with an income equal to the median would be willing to pay 7.10 €. Across imputed parameters, as with the increase in the data limit, higher income consumers are willing to pay much higher prices for unlimited voice allowances than are lower income consumers.

Table 12 presents willingness to pay for an increase in download speeds from 10 Mbps to 20 Mbps on a 10 000 MB plan. Results are similar to the estimated willingness to pay for an increase in the data limit from 1 000 MB to 4 000 MB (Table 10). Using the preferred imputations (the row in bold), a consumer with an income equal to the 90th percentile would be willing to pay 5.47 € for the faster download speed, while a consumer with an income equal to the 10th percentile would only be willing to pay 2.13 €.

D.2 Cost Estimation Results

Tables 13 and 14 present per-user and per-tower cost estimates, respectively, across a range of imputed price elasticities and nesting parameters. These estimates are recovered by inverting prices and radii, as described in Section 4.2 in the main text. Table 13 presents the estimated per-user costs, averaged across products with similar data limits, and Table 14 presents estimated costs per tower for each MNO, averaged across markets.

Estimated per-user costs increase considerably in the size of the data limit. For our preferred elasticity and nesting parameter, for example, small data limit plans (those with data limits less than 1 000 MB) have an average per-user cost of 5.50 €, medium-sized data limit plans (between 1 000 and 5 000 MB) an average of 9.56 €, and large data limit plans (over 5 000 MB) an average of 18.18 €. These patterns hold across different imputations of the elasticity and nesting parameter. We present these cost estimates by data limit category (rather than firm) because these categories account for most of the variation in per-user costs across products.

Table 13: Per-user cost estimates

Elasticity	**Nesting Parameter**	$\bar{d} < 1\,000$ (in €)	$1\,000 \leq \bar{d} < 5\,000$ (in €)	$\bar{d} \geq 5\,000$ (in €)
-3.2	0.0	6.77	9.11	15.47
		(0.33)	(0.81)	(1.68)
	0.5	6.71	9.11	15.49
		(0.28)	(1.07)	(2.18)
	0.75	6.71	9.16	15.63
		(0.35)	(1.57)	(3.30)
	0.85	6.73	9.20	15.75
		(0.57)	(2.24)	(4.89)
$\mathbf{-2.5}$	0.0	5.52	9.40	18.03
		(0.57)	(0.72)	(1.68)
	0.5	5.48	9.47	18.06
		(0.57)	(0.89)	(2.11)
	0.75	**5.50**	**9.56**	**18.18**
		(0.84)	**(1.27)**	**(3.37)**
	0.85	6.25	10.21	12.88
		(0.75)	(0.92)	(4.43)
-1.8	0.0	3.32	8.44	18.75
		(1.23)	(0.89)	(2.50)
	0.5	3.32	8.54	18.68
		(1.39)	(1.10)	(3.33)
	0.75	3.35	8.64	18.80
		(2.12)	(1.69)	(5.74)
	0.85	4.34	8.85	14.48
		(1.77)	(1.45)	(6.57)

Values are the estimated average per-user cost, where the average is taken across all products in the data limit range of the corresponding column. Values in parentheses are the average standard errors. The row in bold corresponds to the imputed elasticity and nesting parameter presented in the main text.

Estimated per-base station costs are similar among each of the four MNOs. Converting monthly estimates to the sunk cost of investment (see the footnote attached to Table 14 for details), the estimated cost per base station for Orange for our preferred imputations is 143 000 €. Per-base station costs do vary across markets. For Orange, the estimated standard deviation in the cost per base station across markets is 45 000 €, reflecting differences in land acquisition costs, labor costs, etc.

D.3 Counterfactual Results

This section considers the robustness of our counterfactual results to different price elasticities and nesting parameters. In this section, we present results for different counterfactual exercises described in Section 6 in the main text for the same range of elasticities and nesting parameters as those used in sections D.1 and D.2 above.

Endogenous variables such as prices, investment, and download speeds are broadly quite

Table 14: Per-base station cost estimates

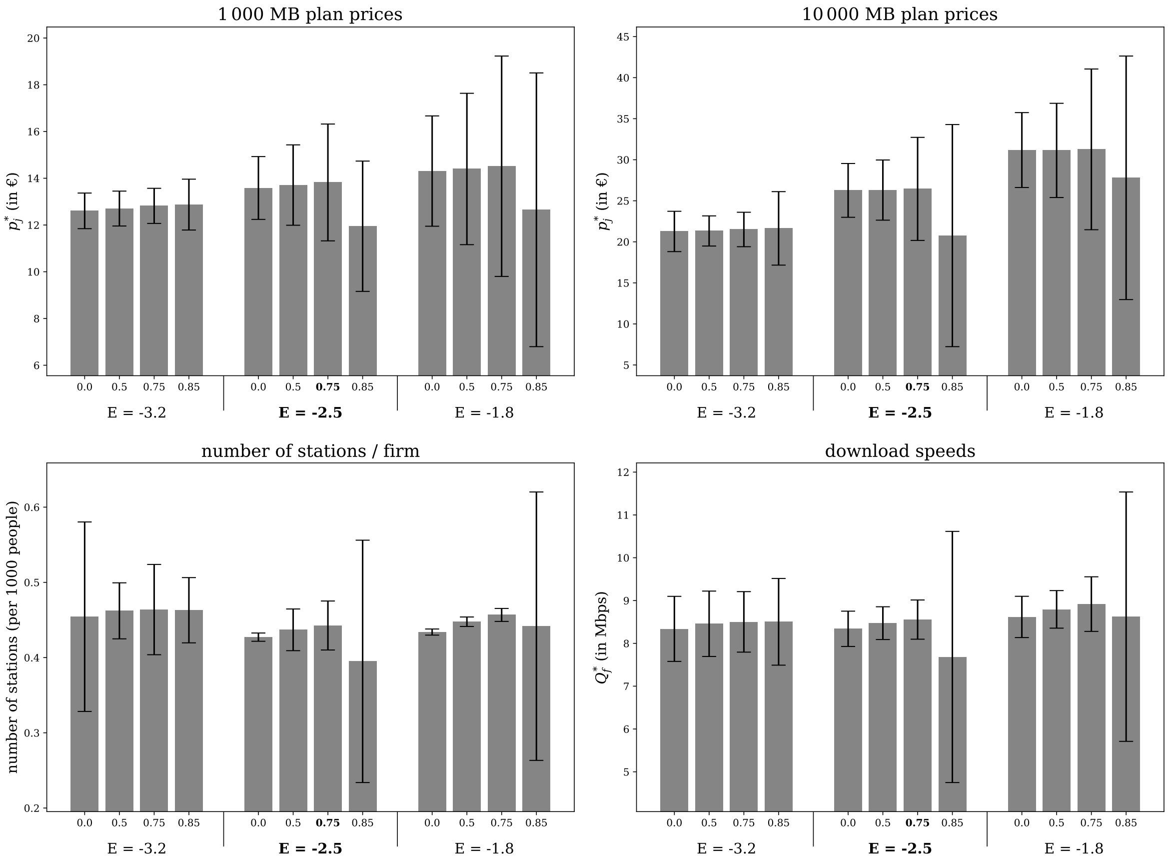
Elasticity	Nesting Parameter	Orange (in €)	SFR (in €)	Free (in €)	Bouygues (in €)
−3.2	0.0	190 006	157 812	128 021	202 959
		(64 063)	(57 687)	(44 025)	(69 802)
	0.5	189 626	155 995	124 536	201 442
		(63 663)	(56 053)	(42 612)	(67 641)
	0.75	189 690	155 822	125 641	200 898
		(63 340)	(55 085)	(43 295)	(65 587)
	0.85	189 605	155 488	124 612	200 567
		(63 074)	(54 620)	(43 102)	(64 666)
−2.5	0.0	143 342	106 168	101 042	150 937
		(45 540)	(32 809)	(31 459)	(48 421)
	0.5	142 949	104 070	95 653	149 515
		(45 118)	(31 671)	(29 666)	(47 212)
	0.75	**142 954**	**103 026**	**92 995**	**148 856**
		(44 815)	**(31 003)**	**(28 910)**	**(46 373)**
	0.85	106 447	47 780	8 908	107 670
		(32 973)	(17 340)	(3 507)	(40 493)
−1.8	0.0	55 297	36 299	39 716	57 597
		(17 435)	(10 119)	(11 821)	(18 057)
	0.5	54 707	34 979	36 405	56 530
		(17 141)	(9 662)	(10 829)	(17 578)
	0.75	55 696	35 163	36 070	57 257
		(17 345)	(9 653)	(10 758)	(17 707)
	0.85	20 984	10 464	4 276	21 231
		(6 560)	(3 156)	(1 333)	(6 995)

We estimate base station costs using monthly profits. Estimates presented here are in per base station terms rather than per base station-units of bandwidth terms. To create per base station costs, we use the per base station-units of bandwidth costs we recover from our estimates and multiply them by 75. This corresponds to the estimated cost of a base station operating 75 MHz of bandwidth, which is similar to the average amount of bandwidth per firm across markets. To recover the cost of long-lived base stations from our estimates based on monthly profits, we assume a monthly discount rate of 0.5%. The above results are therefore $\frac{1}{1-0.995} = 200$ times the per-base station costs we recover. Values in parentheses are standard deviations of the distribution of estimated costs across markets (not standard errors in the estimates). The row in bold corresponds to the imputed elasticity and nesting parameter presented in the main text.

similar across elasticities and nesting parameters. Figure 18 plots these endogenous variables in the four-firm symmetric equilibrium for different imputations. Prices for the high data limit plan increase with a less elastic imputed elasticity (the price is 21.50 € for $E = -3.2$ and 31.26 € for $E = -1.8$ for $\sigma = 0.75$), but prices for the low data limit plan are nearly the same across elasticities (12.82 € versus 14.51 € for the same elasticities). Investment and download speeds follow a similar pattern to that of the prices for the low data limit, increasing only a little as we impute a less elastic elasticity.

The relationship between the number of symmetric firms and welfare, however, displays a pattern that is more dependent on the imputed elasticity. Figure 19 plots the relationship between the number of symmetric firms and consumer, producer, and total surplus for different

Figure 18: Counterfactual prices and qualities across imputations



Each subplot corresponds to a particular variable in the four symmetric firm-equilibrium. Along the x-axis of each subplot, the bottom row corresponds to an imputed price elasticity, and the top row corresponds to an imputed nesting parameter. The imputations in bold correspond to those presented in the main text. Error bars represent 95% confidence intervals.

elasticities (rows) and nesting parameters (individual lines). The optimal number of firms from the perspective of consumer or total surplus varies considerably based on the imputed elasticity. The number of symmetric firms that maximizes consumer surplus at $\sigma = 0.75$ is 2 for $E = -3.2$, 6 for $E = -2.5$, and 9 for $E = -1.8$, and the number that maximizes total surplus follows a similar pattern (2, 3, and 9, respectively).[57] The nesting parameter does not appear to have as much of an impact on the optimal number. While these results are quite sensitive to the choice of the imputed elasticity, Bourreau, Sun and Verboven (2021), also studying the French mobile telecommunications industry, finds an elasticity that corresponds to about -2.5, making it a sensible baseline.

Bandwidth derivatives, which capture the value of marginal bandwidth (see Section 6.2), are

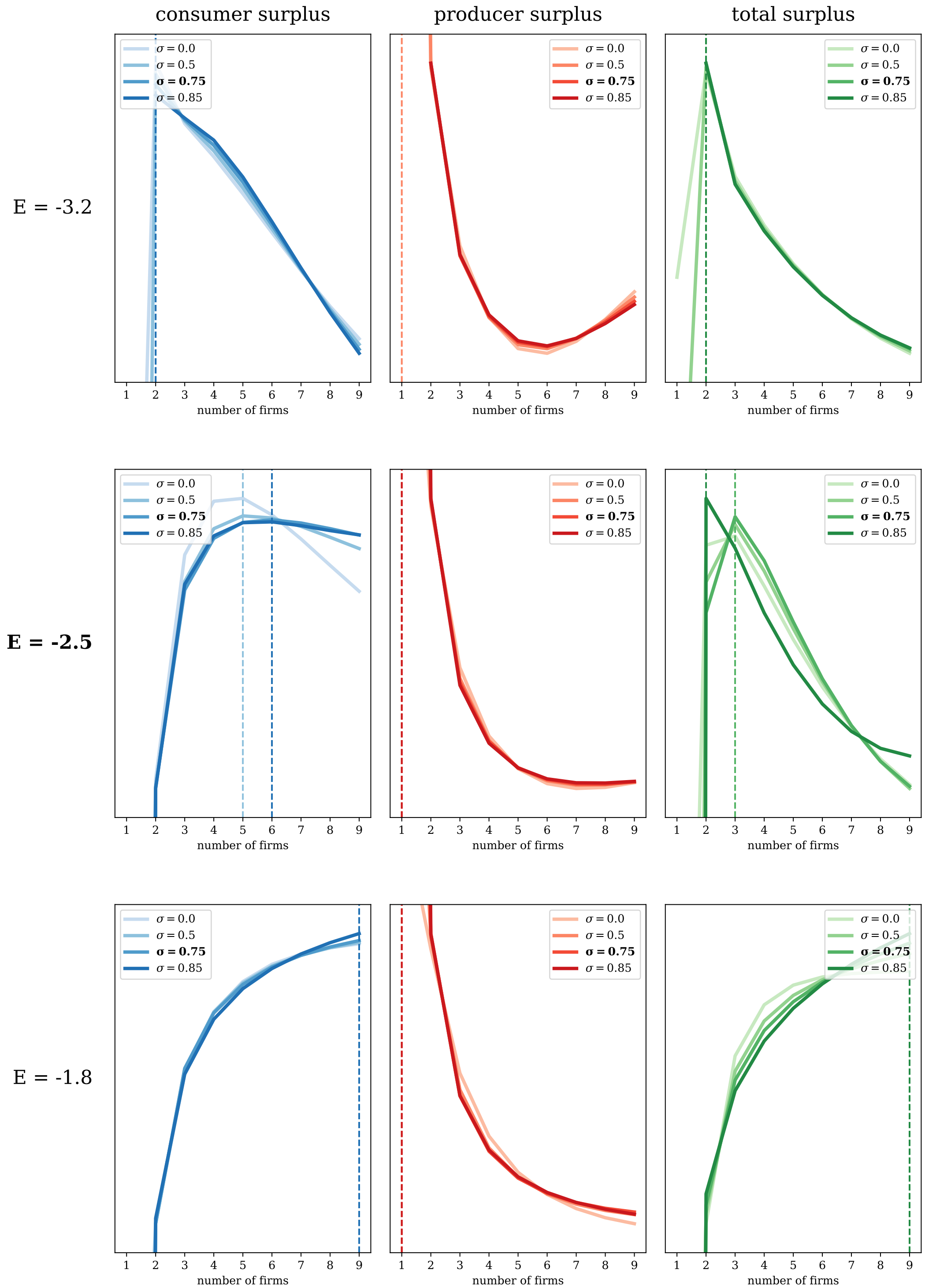
[57] Note that 9 firms is the miaxmum number that we simulate, but the maximum may actually occur at a higher number.

responsive to both the elasticity imputed and the specification of the cost function. Figure 20 presents bandwidth derivatives (analogous to Figure 11) for four *ex ante* symmetric firms. The columns correspond to either a fixed or a bandwidth cost function specification (columns). The fixed cost specification assumes that base station costs are fixed and do not vary by the amount of bandwidth operated, while the bandwidth cost specification assumes that base station costs scale with bandwidth (as in equation 18 and the results presented in the main text). Within each subplot is the estimated derivative for a range of elasticity and nesting parameter imputations. For each derivative and cost function, the magnitude of the derivative is decreasing as we make the imputed elasticity less elastic.

The value of most interest, the ratio of marginal own-profits and marginal consumer surplus, however, is less sensitive to the imputed price elasticity. Using $\sigma = 0.75$, for the bandwidth cost specification, the ratio $\frac{\partial CS}{\partial b} / \frac{\partial \Pi_f}{\partial b_f}$ is, from most elastic to least, 4.6, 5.3, and 6.1. This value does, however, vary in levels depending on which cost specification we use. The same value as before but for the fixed cost specification yields smaller ratios of, again from most elastic to least, 3.6, 4.0, and 4.4.

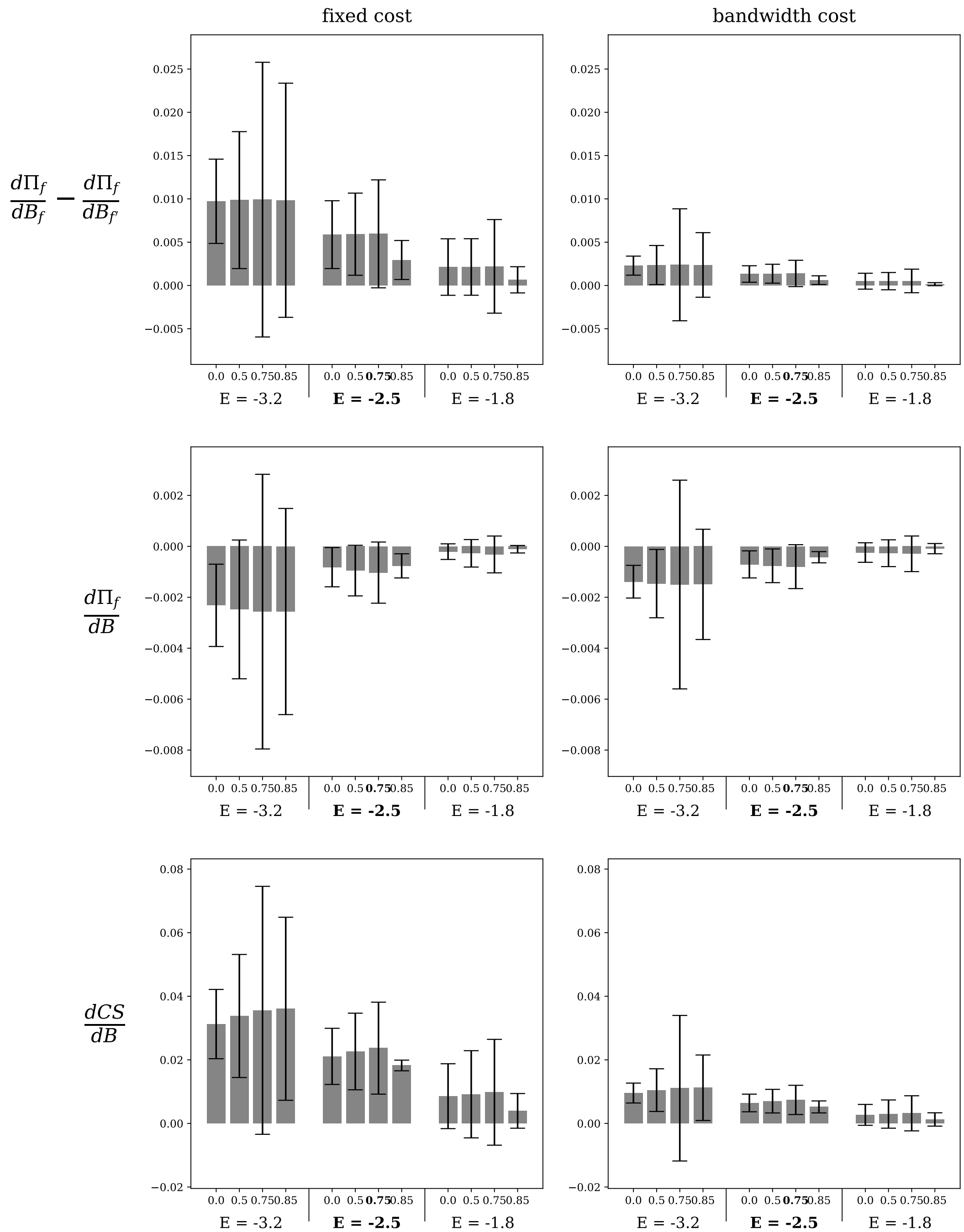
Welfare differences between allocating spectrum to a new firm versus existing firms is somewhat sensitive to the imputed price elasticity but not so sensitive to the imputed nesting parameter or cost specification. Figure 21 presents the impact on consumer, producer, and total surplus relative to the three-firm, original amount of bandwidth equilibrium for different imputed price elasticities (rows), cost specifications (columns), and nesting parameters (groups of bars within subplots). Which equilibrium (allocating to incumbent firms, denoted "3" in the graph, or allocating to an entrant, denoted "4") maximizes a welfare measure is robust to both cost specifications and to all of the nesting parameters we consider. It does, however, change based on whether we use an elastic or inelastic imputed elasticity. For the most elastic one that we consider, allocating new spectrum to incumbent firms is better from both a consumer surplus perspective and a total surplus one. For the least elastic one, the reverse is true; allocating to an entrant is better from both perspectives. For our baseline elasticity, we get the tension presented in the main text; allocating to an entrant is better for consumer surplus while to incumbents is better for total surplus.

Figure 19: Welfare by number of firms across imputations



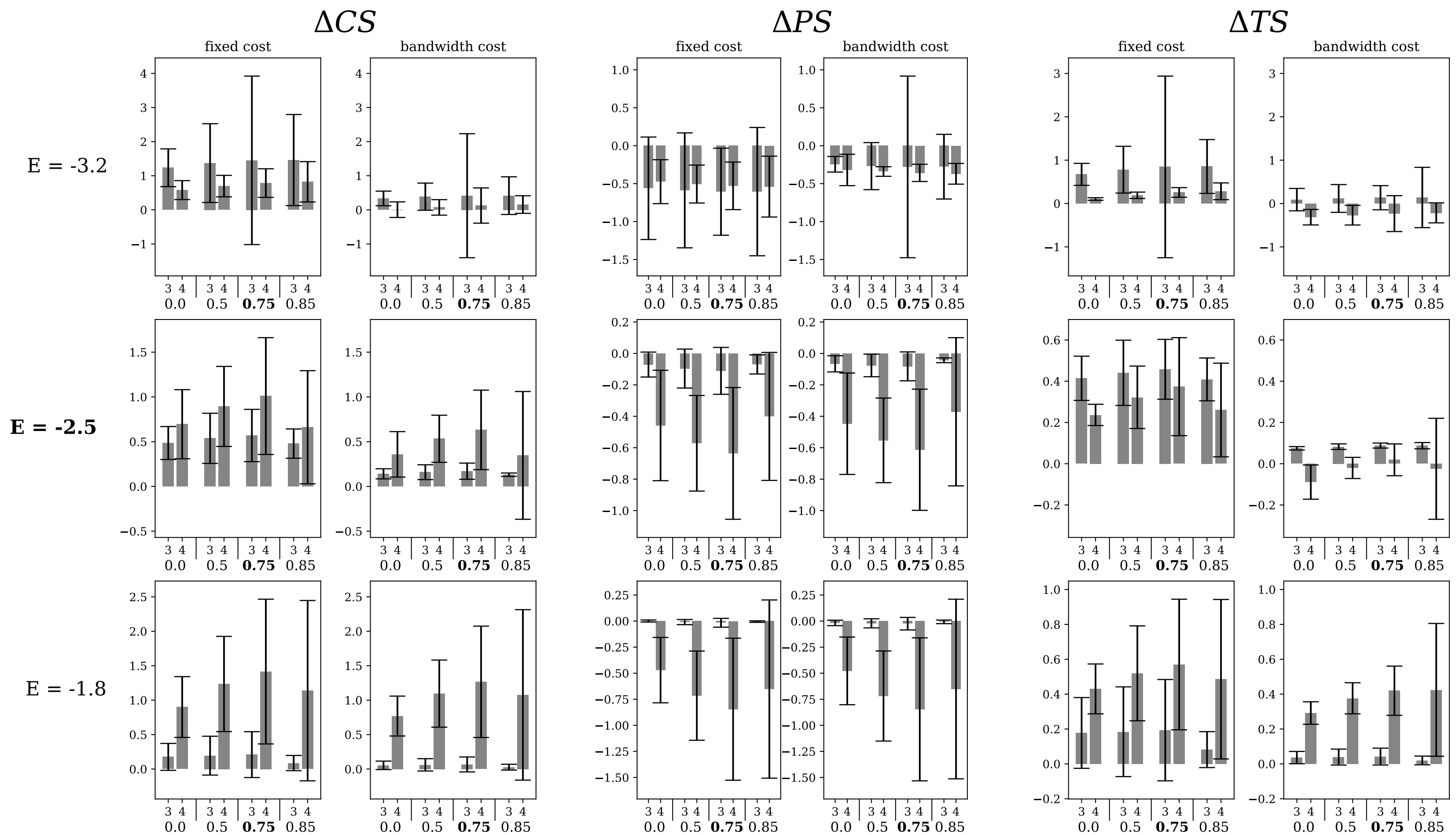
Columns correspond to consumer, producer, and total surplus, and rows correspond to an imputed price elasticity. The x-axis of each subplot represents the number of symmetric firms in the simulated market, and within each subplot, each line corresponds to an imputed nesting parameter. Dashed lines represent the number of symmetric firms that maximizes the welfare measure in the corresponding column. The imputations in bold correspond to those presented in the main text.

Figure 20: Bandwidth derivatives across imputations



Rows correspond to the marginal contributions of bandwidth allocations for the equilibrium with four *ex ante* symmetric firms. Columns correspond to a cost specification in which the cost comes from the number of base stations ("fixed cost") and one in which the cost comes from the amount of bandwidth operated ("bandwidth cost"). Along the x-axis of each subplot, the bottom row corresponds to an imputed price elasticity, and the top row corresponds to an imputed nesting parameter. The imputations in bold correspond to those presented in the main text. Error bars represent 95% confidence intervals.

Figure 21: Welfare impact of allocating spectrum within the industry across imputations



Columns correspond to the changes in consumer, producer, and total surplus that result from allocating additional bandwidth to the market relative to the three symmetric firm-equilibrium without the additional bandwidth. Rows correspond to imputed price elasticities. Within each column capturing a change in surplus, sub-columns correspond to a cost specification in which the cost comes from the number of base stations ("fixed cost") and one in which the cost comes from the amount of bandwidth operated ("bandwidth cost"). Along the x-axis of each subplot, the bottom row corresponds to an imputed nesting parameter, and the top row corresponds to an allocation of the additional bandwidth, either allocating 33% more to each firm ("3") or adding an additional firm with the same amount of bandwidth ("4"). The imputations in bold correspond to those presented in the main text. Error bars represent 95% confidence intervals.

Table 15: Notation

Symbol	Description
f	indexes firms
i	indexes consumers
j	indexes mobile phone plans
$\mathcal{J}$	set of mobile phone plans
ℓ	a location
m	indexes markets (municipality)
γ_m	data transmission efficiency in market m
ε_{ij}	idiosyncratic, consumer-plan-level demand shock
θ	demand parameters
θ_{pi}	price coefficient
θ_{p0}	parameter controlling the mean of the price coefficient
θ_{pz}	parameter controlling the heterogeneity in the price coefficient
θ_v	coefficient on dummy for unlimited voice
θ_O	coefficient on dummy for Orange plans
θ_c	opportunity cost of time spent downloading data coefficient
θ_{di}	parameter of exponential distribution that defines distribution from which a consumer's utility of data consumption is drawn
θ_{d0}	parameter controlling the mean of θ_{di}
θ_{dz}	parameter controlling the heterogeneity in θ_{di}
ϑ_i	random shock to consumer's utility of data consumption, distributed exponentially with parameter θ_{di}
ξ_{jm}	market-level demand shock
σ	nesting parameter
B_{fm}	bandwidth (in Megahertz)
c^u_j	cost per user
c^s_{fm}	cost per base station and unit of bandwidth
$\overline{d}_j$	data consumption limit of phone plan j
D_m	population density
F	used for CDFs
$I_\ell\left(R_{fm}\right)$	interference power at location ℓ when cell radius is R_{fm}
N_{fm}	number of base stations for firm f in market m
P_j	price of phone plan j
$q_{m\ell}$	data transmission speed at location ℓ in municipality m (in Mbits/second)
$\overline{Q}_{fm}$	channel capacity (in Mbits/second)
Q_{fm}	download speed (in Mbits/second)
Q^L	throttled download speed (in Mbits/second)
Q^D_{fm}	demand requests (in Mbits/second)
R_{fm}	radius of area served by one base station (in km)
s_{jm}	market share
$\boldsymbol{s}$	vector of market shares
S_ℓ	signal power at location ℓ
u	utility of a phone plan
w	utility from data consumption over course of month
x	monthly data consumption