Industrial policies for multi-stage production: The battle for battery-powered vehicles*

Keith Head† Thierry Mayer‡ Marc Melitz§ Chenying Yang¶

February 12, 2026

Download latest draft

Online Appendix

Abstract

We model a multi-stage supply chain for EVs from battery production to vehicle distribution. Given industrial policies, firms select where to open facilities at each stage. This is a difficult combinatorial choice problem that we solve with a fast mixed integer linear programming formulation. We estimate the variable and fixed costs parameters using SMM. Counterfactual simulations reveal a tension between boosting EV adoption and promoting domestic supply chains. Due to increasing returns, even unconditional subsidies raise the number of factories in the subsidizing region—by about 16% for EVs and 7% for cells in North America, and even more in Europe. Theoretically, local assembly requirements can push down delivered marginal costs relative to unconditional subsidies. Empirically, local content requirements quadruple the expansion of cell factories in America, but they drive up costs and reduce subsidy uptake, undoing more than half of the EV adoption stimulus coming from pure buyer subsidies.

*We thank Rocco Machiavello, Rowan Shi, Felix Tintelnot, Alastair Fraser, Pablo Fajgelbaum, and Jin Liu for discussions of our paper at conferences, as well as participants at seminars for many valuable suggestions. We also thank Eileen Bian and Adrian Kulesza for excellent research assistance. All four authors acknowledge internal grants from our respective institutions that helped to finance the data purchased in this project. There are no conflicts of interest to disclose.

†University of British Columbia and CEPR

‡Sciences Po, CEPII and CEPR

§Harvard University, NBER, and CEPR

¶Singapore Management University

1 Introduction

Industrial policies are increasingly being used to reshape the plant location patterns of industries characterized by multi-stage production and economies of scale. Such industries feature complex spatial interdependencies: At any given stage, different locations substitute for each other. However, upstream and downstream production sites are complementary. Costs of production differ, and final demand is spread in a highly uneven manner. These factors imply that there is no obvious way to structure the firm's value chain across space. Furthermore, industrial policies such as subsidies or trade restrictions induce firms to reorganize their global supply chains in complex ways. These challenges are relevant for industries such as semiconductors, solar panels, aluminum, and electric vehicles, all of which have been the subject of recent policy interventions.

This paper develops a quantitative method to analyze those industrial policies for multi-stage production and applies it to the battery electric vehicle (hereafter, EV) industry. Firms decide where to build plants, where to source inputs at each production stage, and where to sell in the final distribution stage. Characterizing this allocation of production and distribution stages across space—the firm's global supply chain—is a challenging computational problem for even moderate numbers of alternative locations. Even with just one stage of production and fixed distribution, the facility location problem (as it is referred to in the operations research literature) is already known to be an NP-hard problem. Loosely speaking, this means that there are no algorithms guaranteed to solve it in polynomial time, or, put more simply, the problem-solving time blows up as the number of locations increases. The difficulty is compounded under multi-stage production because the optimal location of a given stage is not only a function of the marginal cost and fixed cost at that production stage, but is also shaped by the proximity of that location to the desired locations of production of the upstream and downstream stages. One contribution of this paper is to adapt techniques from operations research to solve for the geography of global supply chains with fixed costs.[1] We extend the multi-stage production cost minimization problem to allow for endogenous demand and market entry by multi-product firms.

The application to EVs follows from policy relevance and data availability. The 2022 US Inflation Reduction Act (IRA) awards consumer and production subsidies that do not extend to Asian and European-made cars. Canada promised to match US IRA production subsidies for factories that Volkswagen and Stellantis are constructing in Ontario. Deputy Prime Minister Freeland defended the Canadian government's decision to spend roughly $30 billion on production subsidies, saying in July 2023, "Our government is absolutely determined that Canada gets its fair share of those green jobs." In 2024, the European Union imposed tariffs on Chinese EV imports and France limited its EV subsidy to cars with batteries from countries with a clean energy mix. With large fixed costs for each new facility across production stages, these industrial policies have the potential to induce a complex process of production relocations across those production stages.

Estimation of the model proceeds in two stages. Data constructed by the S&P Mobility con-

[1] "Squeezing" algorithms following the contribution of Jia (2008) cannot accommodate our multi-stage framework with fixed costs. We discuss this in further detail below.

sultancy list the configuration of the battery electric vehicles sold in all the major markets of the world. From this data, we observe for every car model sold in a given country, the factory where that car was assembled. We call this the vehicle sourcing decision. The data further reports the location where the cells were manufactured, which we call the cell sourcing decision. Our focus on cells comes from the fact that they represent around three quarters of the overall battery cost (AlixPartners, 2021). Conditional on the location of the cell and assembly facilities, the determinants of these sourcing decisions can be estimated using nested discrete choice. We take advantage of this feature to reduce the number of parameters to be estimated in the computationally demanding second-stage estimation. The first stage provides variable costs as a parametric function of dyadic variables (borders, distance, trade agreements and tariffs) along each potential path (cell-location, assembly-location, consumer country).

The first stage estimation takes the open facilities as given, so fixed costs must be obtained in a second stage that solves the firms' full optimization problem—including the facility activation decisions. We estimate the parameters governing the distribution of fixed costs via simulated method of moments (SMM) solved using a constrained optimization approach, mathematical program with equilibrium constraints (MPEC) to account for endogenous market demand. The estimated fixed costs increase in distance to headquarters and the quality of the vehicle, and are much lower in Asia (China, Japan, Korea) than in North America or Europe.

With the complete set of parameter estimates, we turn to the primary objective of the paper: quantifying the impact of industrial policies on the location of the value chain and on the extent of clean vehicle adoption. We conduct counterfactuals on the main types of policy that governments have contemplated or implemented in recent years. Our main focus is on consumer subsidies that impose requirements on the location of different production stages. Given the large fixed costs, subsidies can induce changes in production locations that reduce the delivered cost to consumers, excluding the subsidy itself. When the new open facility is in the subsidizing region, this potentially delivers a "win-win" outcome for industrial policy: lower costs and higher employment. In some cases, adding local content requirements are needed to induce the relocation of a facility to the subsidizing region. Thus, a protectionist element can theoretically deliver such a win-win outcome. However, one of our key takeaways is that, empirically, those requirements do not deliver cost gains relative to unconditional subsidies and therefore impede EV adoption.[2]

In both North America and Europe, we find that the most effective way to expand EV purchases is to offer consumers the subsidy regardless of where the car and its battery is produced. For North America, a 20% subsidy generates a 5% variable cost reduction generated by production relocation choices. If the subsidy requires North American production, the EV adoption stimulus declines from 86% to 71% (assembly only) to 32% (assembly and cells). This latter conditioning drastically affects the associated variable cost change, reversing the cost reduction into a 2% cost *increase*. It also delivers a mixed protectionist outcome for North America. Although it increases

[2]In appendix A.2, we develop a simple theoretical example highlighting how subsidies and content requirements can generate those cost savings.

the number of cell plants by 20%, this is offset by 7% fewer assembly plants—relative to the unconditional subsidy. On the other hand, the same policy for Europe achieves an unambiguous protectionist goal of raising both cell and vehicle production (21% for cells and 7% for vehicles).

The paper is organized as follows. Section 2 positions this paper within the literature on sourcing across the value chain as well as the emerging work on EVs. Section 3 describes the data on the global EV industry we use, and illustrates important features of the two major stages of production of this sector (cell manufacturing and vehicle assembly). We document a set of key facts about the EV industry that inform core assumptions in the model in section 4. Section 5 presents the model of multi-stage production market entry, and equilibrium. Section 6 estimates the parameters governing sourcing conditional on existing plants. The method of obtaining variable profits associated with each potential path is developed in section 7. We use simulated method of moments combined with MPEC to estimate the remaining set of structural parameters in section 8. With the full set of parameters in hand, section 9 reaches the end goal of simulating the effects of a menu of industrial policies drawn from recent approaches taken in North America and Europe. Section 10 concludes.

2 Literature

This paper is situated at the confluence of three streams in the literature. The first is the modelling of global value chains (GVCs) which we refer to in this paper as multi-stage production. The second literature considers the role of interdependencies in sourcing/location choices at a given stage of production. The third literature evaluates recent industrial policies directed at the EV industry.

Global value chains with constant returns: The GVC literature starting with Yi (2003, 2010) and more recently Antràs and De Gortari (2020), Tyazhelnikov (2022), and Johnson and Moxnes (2023) tackles one type of interdependence: the presence of trade costs makes production stages interdependent such that simple stage-by-stage minimization of costs does not solve the global problem. These papers consider multi-stage production with transport costs and comparative advantage, developing recursive methods where cost minimization at each stage accounts for optimal choices made at all preceding stages. The methods in these papers assume constant returns. However, location choices typically involve opening facilities, which require fixed costs and generate increasing returns to scale. Non-constant returns to scale induce a second type of interdependency between location decisions within each stage.

Plant location/sourcing with interdependencies: An early analytical treatment of this second type of interdependence is Yeaple (2003), who focuses on a three-country case, where two "Northern" countries trade-off horizontal and vertical motives for foreign direct investment. Tintelnot (2017) pioneers work on multinational location choice with many countries, heterogeneous firms,

and fixed costs. Antràs et al. (2017) model global sourcing decisions when there is a fixed cost of adding a new supplier country. In the short run, yet another form of interdependence can be important. Castro-Vincenzi (2024) shows how to solve single-stage location models with capacity constraints. This is important in the empirical application of his paper: floods that shut down factories. Our paper is oriented towards long run impacts and does not incorporate capacity constraints. We will present evidence to justify this approach in our context.

Plant location with fixed costs is a hard combinatorial problem as the number of potential locations (L) grows. Even with just a single production stage, brute force requires evaluating 2^L alternatives. Researchers have developed two approaches to this computational challenge. One path, taken by Oberfield et al. (2024), uses a limit solution where firms choose a density of plants over continuous space rather than making discrete plant location decisions. This limit approach is particularly relevant for retail/services where firms have many "plants" serving small catchment areas. The other path, pioneered by Jia (2008), uses supermodularity of the profit function to reduce the computational burden. Antràs et al. (2017) point out that the global sourcing problem can be either supermodular or submodular depending on the relative magnitudes of the key demand and supply parameters. With σ denoting the elasticity of substitution between varieties in the demand function, and θ denoting the Fréchet shape parameter on the supply side, supermodularity obtains if and only if $\sigma > 1 + \theta$. This restriction holds in Antràs et al. (2017): $\sigma = 3.9$, $\theta = 1.8$. However, Arkolakis et al. (2023) find $\sigma = 4$, $\theta = 4.5$ implying submodularity. They extend the Jia (2008) method to exploit submodularity with a "generalized squeezing" algorithm. Their solution covers cases that are either sub- or super-modular *globally*.

Our structure requires a different approach because it combines both submodularity and supermodularity forces: The submodularity mechanism is that plants at the same stage substitute for each other. The supermodularity mechanism is that upstream production facilities complement downstream ones and vice-versa. Furthermore, activated distribution facilities (product entry) increase the profitability of adding production plants. Thus, there is no realistic way of imposing either global sub- or supermodularity to our application. A critical advantage of our approach is that we do not need either one to hold.

Combining *multi-stage production with fixed costs* is a challenging problem and there have so far been only a few attempts. Antràs et al. (2024), de Gortari (2020), and Castro-Vincenzi et al. (2025) stand out in this respect. Antràs et al. (2024) extend Antràs et al. (2017) to allow for many markets to be served by the firm with its potentially many plants. Firms choose three global strategies associated with three *firm-level fixed costs:* one for assembly plants, one for sourcing plants and one for marketing. This enables rich interdependencies *between* each of those margins. However, the authors make strong parametric assumptions that restrict the firm-level problem to be supermodular.

A note by de Gortari (2020) proposes an algorithm for solving GVC models with fixed costs. The proposed solution is elegant and fast, but limits the analysis to a single final good sold in a single destination, reducing considerably the extent of interdependencies. In our framework with

multi-market and multi-product firms, the facility fixed costs make it impossible to separate the computation of the equilibrium by product-market: Those problems are all interconnected and must be solved globally for the firm, along with the industry equilibrium in each market. Castro-Vincenzi et al. (2025) also develop a model with multiple markets and products. They extend the Arkolakis et al. (2023) squeezing algorithm for problems that are neither globally sub- nor supermodular. The squeezing algorithm then leads to bounds on the firm's optimal choices. In their application to the auto industry, they recover bounds for the fixed location and market entry costs.

Our methodological contribution to this combinatorial discrete choice literature is to adapt mixed integer linear programming (MILP) methods from operations research to obtain unique solutions to global supply chain location choices that combine sub- and super-modularity forces. More precisely, we extend the Multi-level Uncapacitated Facility Location Problem (MUFLP), (Ortiz-Astorquiza et al., 2018, provide a comprehensive survey) to allow for endogenous demand, multi-product firms and market entry.

Comparison to recent work on industrial policy towards EVs: Barwick et al. (2024) document that EV-oriented industrial policies now account for almost half of all industrial policies worldwide. Two recent working papers examine the application of industrial policies to the EV industry: Allcott et al. (2024) and Barwick et al. (2025). While there are several areas where our papers overlap with these two, the crucial difference is our focus on the endogenous location of plants.

Allcott et al. (2024) stands out for its careful treatment of the institutional features of the IRA, including the leasing loophole and the detailed treatment of subsidy eligibility over time. The policy simulations in our paper are inspired by the IRA provisions of making subsidies contingent on the value chain, but are not intended to evaluate the IRA's effects. Barwick et al. (2025) investigate how subsidies (consumption and production related) affect learning-by-doing in the EV battery industry in a model where oligopolistic EV assemblers and battery manufacturers bargain over the final battery prices. This results in an estimation equation that accounts for both vehicle and battery equilibrium markups (obtained from the bargaining equilibrium) in the formation of vehicle's price, together with marginal costs. In our framework, producers maximize total profits over the whole value chain. The implicit assumption is that bargaining between cell producers and assemblers leads to efficient production decisions where their relative bargaining power determines lump-sum transfers.[3] This joint optimization determines the endogenous location choices for both battery production and vehicle assembly—whereas Barwick et al. (2025) take facility locations as given. In our case, industrial policy potentially reduces marginal costs via adding plants closer to downstream purchasers, rather than through achieving cost reductions via learning-by-doing as in their paper.

[3] In a few important cases such as BYD and Tesla, cell production is integrated by the assembler. In many other cases, cell production by a separate firm is set up as a joint-venture with the assembler. We provide additional supporting evidence for this simplifying assumption of joint optimization by cell producers and assemblers in section 4.2.

3 Description of the Global EV Industry

This section describes our data and presents some useful facts about the EV industry.

3.1 Data on Batteries and EVs

Our analysis makes use of three data sets provided by the consultancy IHS-Markit, which was recently integrated within S&P Mobility.

The first one is the **High Voltage Battery (HVB)** dataset. For each electric vehicle produced in a given assembly plant, HVB provides the source of the battery cells (production plant location and supplier). Additional details are provided regarding the shape of the cell (prismatic, pouch or cylinder), the detailed chemistry (e.g. NCM523, NCM811), and the battery capacity in kilowatt-hours. We aggregate the detailed chemistry into two main material categories: NMC (Nickel-Manganese-Cobalt) and LFP (Lithium-Ferrous-Phosphate). We calculate the power of the cells as the ratio of battery capacity to number of cells. For each battery specification, HVB provides the total volumes of cars by assembly location starting in 2015. This is production-based data that do not specify where the cars are sold.

The second source is the **Global Sales** dataset, which is an updated version of the data used in Head and Mayer (2019). For every light vehicle sold in 75 destination countries, we observe the vehicle model and assembly plant. For each model, Global Sales records the brand (e.g. Chevrolet; Audi), the brand owner (e.g. GM; VW), and the production platform (e.g. "GEN III" for Tesla Model 3 and Model Y; "MEB" for VW ID.4, ID.3, Audi Q4, and Skoda Enyaq). Platforms establish the architecture of core vehicle systems (e.g. steering, chassis) and dictate the placement of their components. They require a specific organization of the assembly line and its machine tools.

Some models are produced both as EVs but also with internal combustion engines (ICE).[4] For those models, Global Sales does not provide the sales destinations by fuel type. In those cases, we use a third dataset to recover the EV production-destination flows. This third source is **New Registration and Prices (NRP)** dataset. This is a destination-based dataset reporting sales (registrations) for 24 major markets by detailed model characteristic. The latter includes the fuel type, which we use to recover the production-destination flows for all EVs. Another characteristic is the vehicle price, which we also use in our analysis.

Combining these three sources, we construct a novel type of micro-level GVC dataset that tracks the production paths from battery cells to assembly to final sales destination for all EV models sold in 24 major markets from 2015 to 2023. We geocode all production locations across those two stages by latitude and longitude, which we use to compute distances for every production path from battery cell to assembly and from assembly to final destination. For the latter, we use major cities as destinations and compute a population weighted average of the distances. We add a binary Regional Trade Agreement (RTA) indicator from the CEPII gravity data and tariffs

[4]For example, the Hyundai Kona is produced with many fuel types: Gas, Diesel, Hybrid, Plug-in hybrid, and EV.

from World Integrated Trade Solution (WITS) for the two segments along every path.[5]

3.2 The Geography of the Global EV Industry

We use the above datasets to describe the basic features of where EVs are purchased and the spatial organization of their production.

Table 1: The 10 top EV markets in 2023

Rank	Country	# Models	# Firms	Sales (000′)	EV share (%)
1	China	206	38	5353.0	21.0
2	United States	57	16	1264.3	8.1
3	Germany	124	25	625.8	20.2
4	United Kingdom	104	20	389.7	17.3
5	France	108	20	383.6	17.9
6	South Korea	48	10	181.8	10.7
7	Canada	53	15	160.3	9.4
8	Netherlands	120	24	147.3	33.3
9	Sweden	111	23	143.7	43.1
10	Norway	113	26	132.0	83.7
11	Rest Of World	40	12	1225.2	7.4

Note: Rest of World row reports averages across 64 countries for numbers of models and EV share and the sum for sales.

Table 1 shows the top 10 EV markets in 2023. China is by far the largest market in all dimensions except the share of EVs. Relative to the US, it has more than double the number of active firms and almost four times as many models sold in the market. Major ICE vehicle-making countries, such as the US and Canada, have EV shares below 10%, whereas countries with high EV shares—Norway, Sweden, the Netherlands, and China—offered early aggressive inducements for buyers to switch to EVs. Our model captures both heterogeneous market sizes and various pro-EV policies.

3.3 Investment Costs

The development of the EV industry has generated staggering levels of capital investment in new plants along the value chain. There are no official statistics on these investments, but we have collected hundreds of news reports for those investments. Unfortunately, most articles tend to refer to battery plants without differentiating between cell and pack plants.[6] Also, some reports lump investments in the assembly factories together with the battery investments.

Table 2 reports that an average EV assembly plant involves an investment of $660 million, while an average battery plant requires $1.85 billion. There is substantial heterogeneity as cap-

[5]We correct for missing data following the procedures in Head and Mayer (2019).

[6]The most common configuration of an EV battery is that cells are assembled into modules, themselves grouped into a pack, which includes a battery management system notably regulating the cooling system. The pack is what is usually referred to as the battery.

Table 2: Investment costs for battery and EV plants

Stage	Mean Cost ($US bn)	Std. Dev.	Std. Dev (logs)	Article Count
Cells	1.85	2.02	1.50	160
Vehicles	0.66	1.03	1.62	281
Vehicles + Cells	3.88	3.37	1.11	16

Source: Authors' calculations based on collected news articles.

tured by standard deviations that are similar to the means. The standard deviations in logarithms for both cells and vehicle assembly plants serve as targeted moments in our simulated method of moments estimation in section 8.

We take these massive investment costs as evidence for indivisible fixed costs that need to be incorporated into the model because they are core features of this industry. Firms face a fundamental dilemma regarding activating more plants or incurring higher distance and border costs. The fixed costs relevant for the static model described in the next section differ from these investment costs in two main respects. First, the investments lead to long-lived plants, so we need to multiply by the capital cost (interest plus depreciation) to obtain annual fixed costs. Second, governments have often given large subsidies to these firms. From the firm's point of view, the relevant fixed costs are the ones net of subsidies.

3.4 Location of Plants and Sourcing Along the Value Chain

Figures 1 and 2 show the impressive growth in the numbers of plants for cells and vehicles from 2015 to 2023—in all the main regions. The size of plant symbols corresponds to total output in gigawatt-hours for cells and total sales for vehicles. East Asia remains dominant in cell production but the number of plants in the US and Europe grow by factors of three and two respectively and the capacity in GWh rises by two orders of magnitude. While the growth of cells is impressive, it is also worth noting that many countries with multiple vehicle plants in Figure 2 (Italy, Spain, Portugal, and Turkey) lack local cell production.

The maps show the geographic dispersion of plants, but they do not indicate trade relationships between them. Figure 3 provides insight into the distances that components travel. By showing medians and means we see the huge skewness in the distribution of distances that the output of a stage travels before becoming an input. Comparing the two panels of the figure also shows that the distance traveled fell substantially between 2015 and 2023, reflecting the fact that the objectives for new plants are primarily driven by the desire to serve local/regional markets.

Figure 1: Cell plants in the major regions

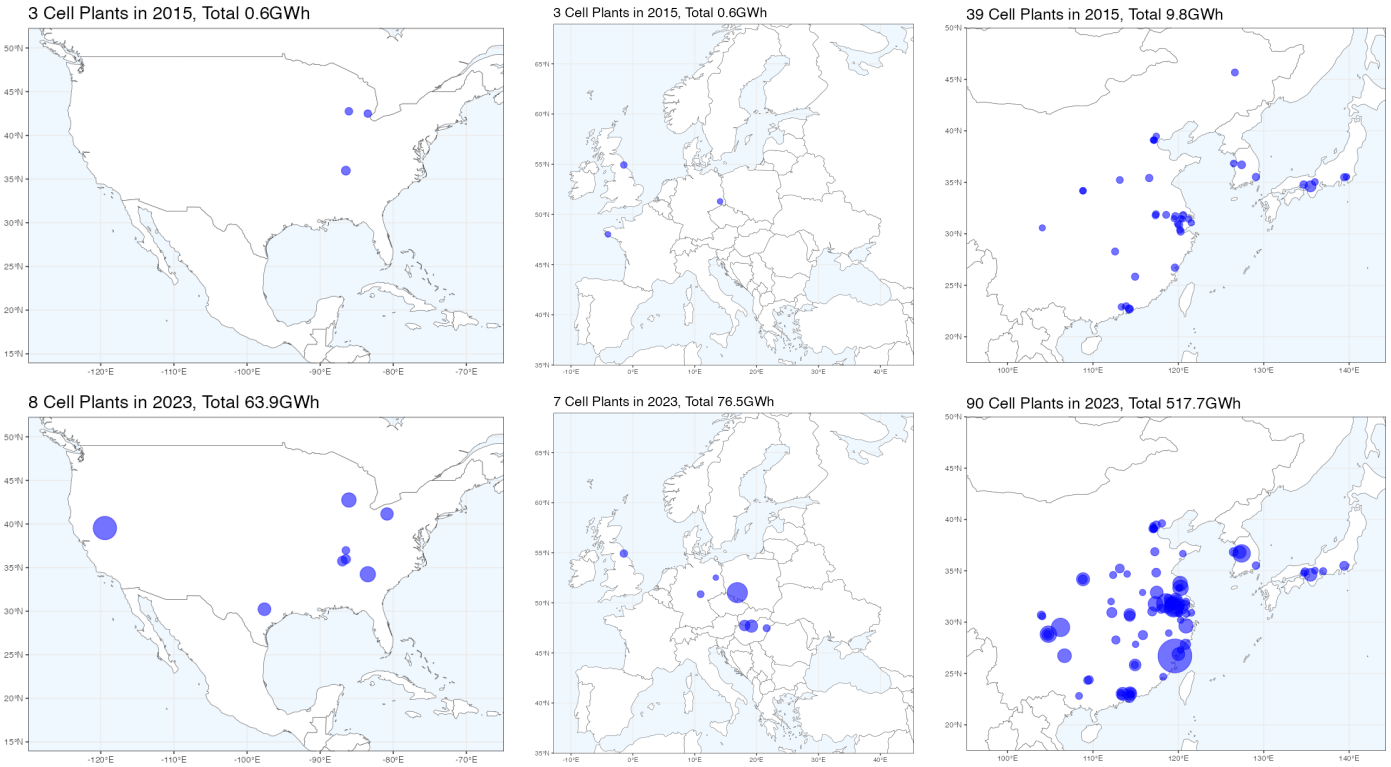


Figure 2: EV plants in the major regions

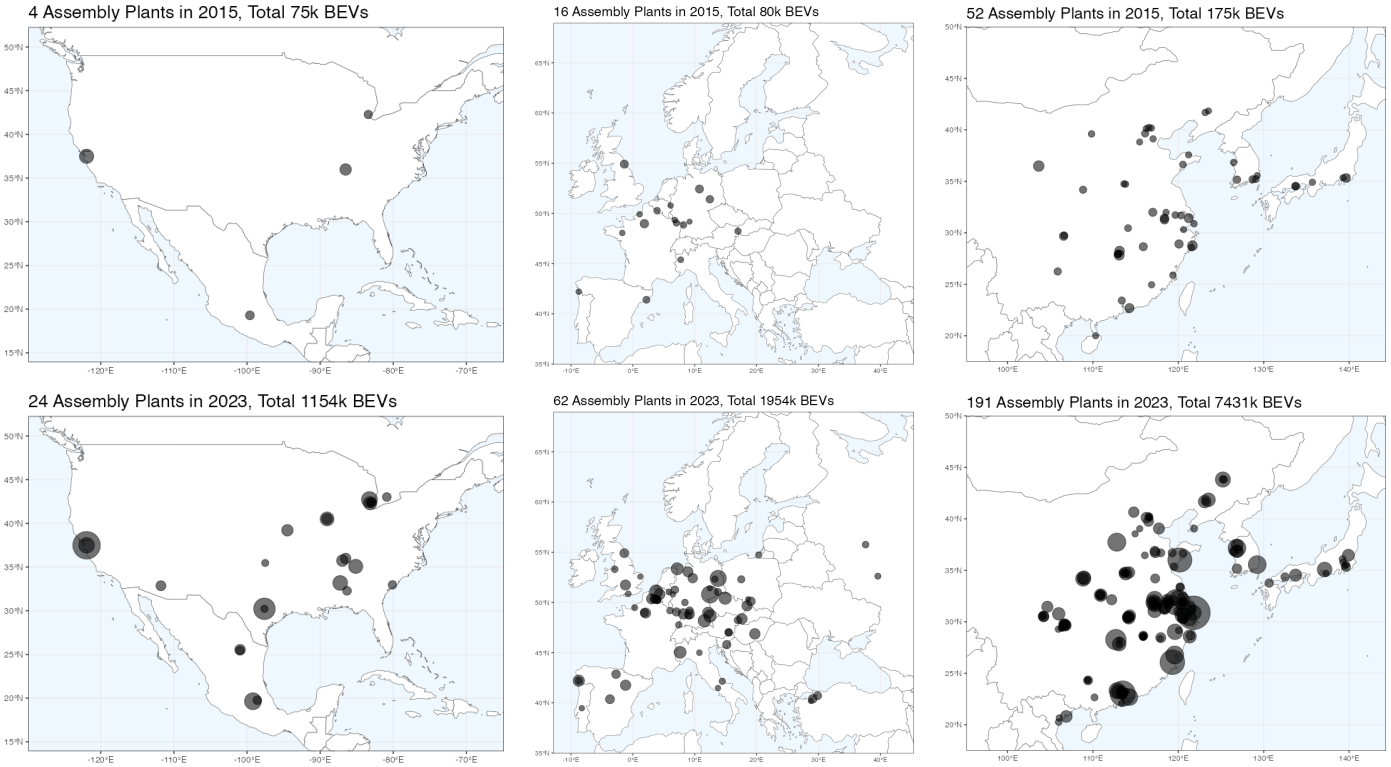
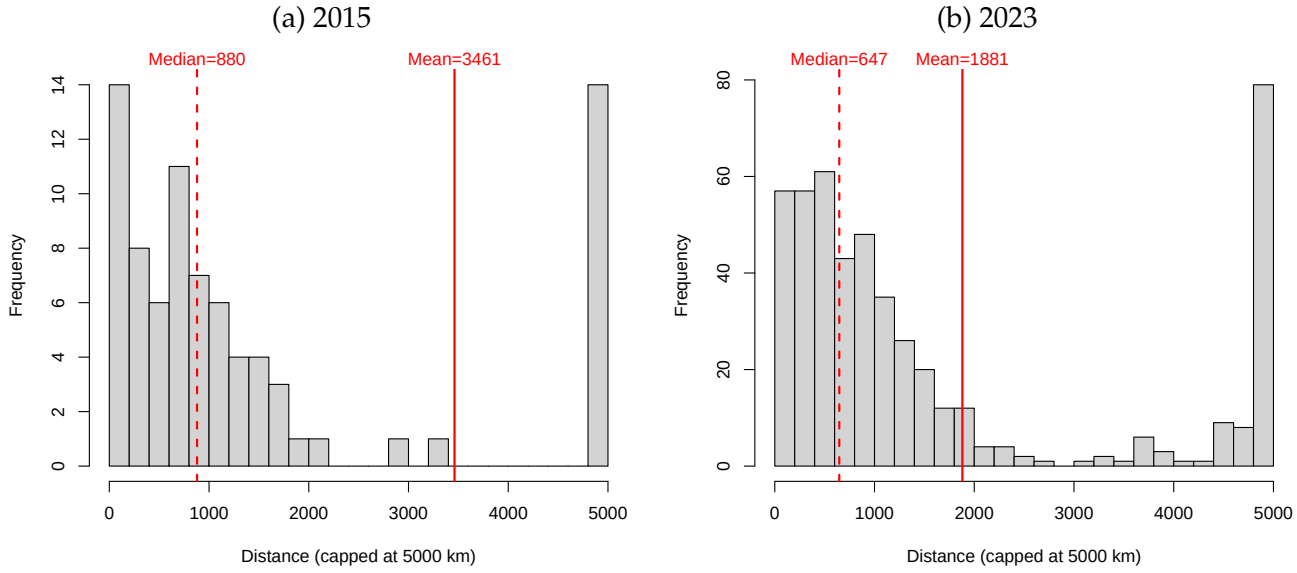


Figure 3: Distances between cell manufacture and vehicle assembly are shortening



4 Facts Guiding the Model

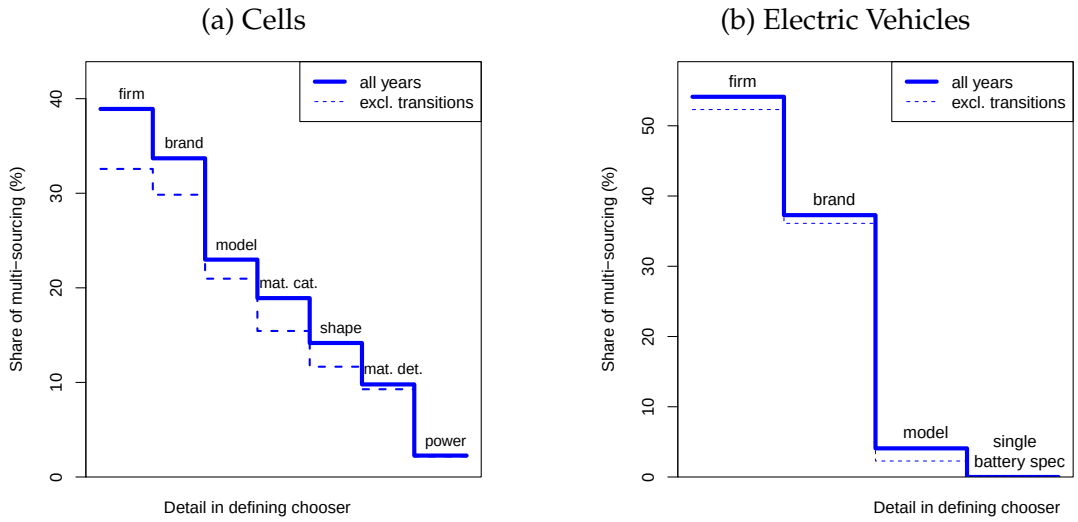
We now turn to key features of the data which determine modeling choices made in the next sections.

4.1 Single Sourcing of Cells and Vehicles

In our model, we assume constant marginal costs of production in both stages, with the increasing returns induced by fixed costs for uncapacitated production facilities. This implies that firms choose a single optimal path for a given EV model sold in a given market. In other words, conditional on the set of activated plants, our framework features single-sourcing. On the other hand, decreasing returns generated by either increasing marginal costs or facility capacity constraints would induce firms to multi-source: obtaining a model sold in a given market from more than one assembly plant, or obtaining a battery used at a given assembly location from more than one cell factory.

Figure 4 shows that when we disaggregate the data to the car model and battery type level, single-sourcing almost always holds. Data availability at this level of disaggregation is crucial for assessing this pattern, as more aggregated data would appear to suggest multi-sourcing. Consider the case of battery cells in panel (a): An assembly plant may purchase one type of cell from one plant and a different type from another plant. This apparent share of multi-sourcing for cells is about 38% at the most aggregated level of a firm-assembly plant purchaser. However, this multi-sourcing share drops as the data is disaggregated, first by the vehicle type that the battery is purchased for (car brand and model), then by the battery type (shape, material, and power). At the most disaggregated level, the share of multi-sourcing is 3%. The dashed lines drop multi-sourcing

Figure 4: Multi-sourcing is rare for both cells and EVs (2022)



Note: The shares represent the number of choosers which decide to source from multiple options. In panel (a), the chooser goes from a firm-assembly plant-year to a brand-model-battery specification (shape, material and power)-assembly plant-year combination. The options are the different cell plants. In panel (b), the chooser goes from a firm-destination-year to a brand-model-destination-year combination. The options are the different assembly plants.

that is transitory, which we define as lasting less than 3 years.

A similar pattern holds for the sourcing of vehicles for delivery to a given market in panel (b). A destination market often sources vehicles produced by a given firm from multiple locations. Indeed, this happens in more than half of cases. However, this multi-sourcing share drops to about 4% for the sourcing of a particular car model.[7] This residual multi-sourcing hides further model details that we do not fully observe in our flow data of vehicles from assembly to final destination. For example, EV models are also differentiated by their battery type. Using our most detailed measure of battery type, we find that car models with a single battery specification are never multi-sourced.[8]

This single-sourcing pattern is very important for thinking about how to formalize the firm's profit maximization problem. It is common in the supply-chain literature reviewed in Antràs and Chor (2022)—based on more limited/aggregated data—to model part differentiation at the level of the supplying countries and broad product categories. This entails a very different modeling approach that captures those sourcing shares. Instead, firms in our model choose the lowest cost source for a narrowly defined component or final vehicle—as we observe in our more disaggre-

[7] In this respect, EVs do not differ notably from traditional vehicles. Head and Mayer (2019) find multi-sourcing (measured in terms of countries of origin rather than plants) for 2.3% of all model-market years from 2000 to 2016.

[8] In appendix figure D3, we show an example of apparent multi-sourcing that is related to models differentiated by battery type. Panel (a) shows the sourcing of VW ID.4s sold in Sweden from two different German assembly plants. This multi-sourcing arises because the ID.4 assembled in Mosel use pouch cells whereas those assembled in Emden use prismatic cells.

gated data.

There is another explanation for multi-sourcing which is also important to consider in setting up the firm's profit maximization problem: capacity constraints. We might see multi-sourcing not because of the love of variety common in trade models but simply because the preferred source is unable to meet the demand. In appendix D.1, we present some examples suggesting that persistent capacity constraints do not seem to be a major concern in our case. Between 2015 and 2023, appendix figure D4 shows that the Tesla factory in Fremont multiplied its output by 10 (from about 50 thousand vehicles to around 500). On top of that, the Shanghai factory was able to add more than a 150 thousand vehicles on its first year in 2020, and reached 950 thousand in 2023.

Continuing with the Tesla example, figure D2 in the same appendix shows the sourcing volumes of the four Tesla models sold in Germany. The Model 3 is sourced both from Fremont and Shanghai in 2020, but this turns out to be transitory: Fremont was the only available source until 2019, when the Shanghai assembly facility opens. Shanghai then transitions to become the sole supplier for the Model 3 to Germany starting in 2021. The story is very similar with the Model Y, which is uniquely sourced from Shanghai in 2021. When the Berlin factory opens in 2022, sourcing for the Model Y transitions to Berlin. In 2023, Berlin becomes the dominant supplier; and sourcing from Shanghai ceases in 2024.

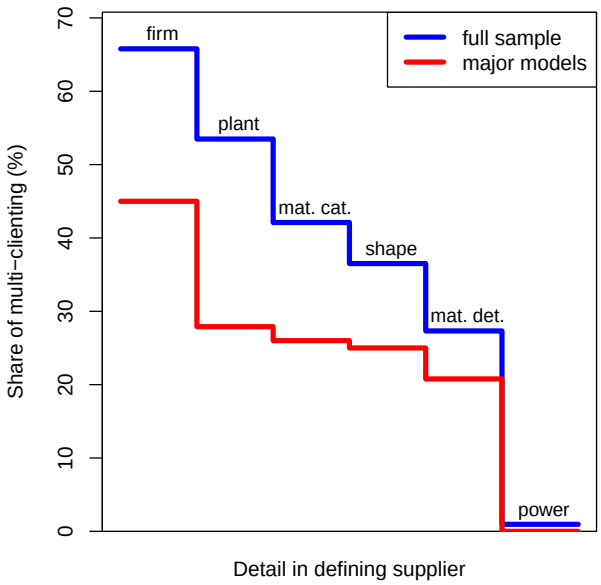
Figure D3 shows that this pattern is not particular to Tesla. Panel (b) shows the case of Volkswagen ID.4 SUV. In 2021, the United States and Canada both uniquely source that vehicle from the German factory in Mosel. When the American plant in Chattanooga opens in 2022, sourcing to both North American countries transitions to that factory. By 2023, sourcing from Germany ceases. In line with these data patterns, we employ the uncapacitated version of the mixed integer linear programming model for facility location problems in this paper.

4.2 Single Clients for Cell Plants

As in Antràs and De Gortari (2020), we model a single "lead" firm that optimizes over the whole supply chain and chooses the lowest cost paths along with the endogenous set of open production facilities. These "lead" firms often take the form of a joint venture between an assembler and cell producer when the two are not vertically integrated. This rules out the case of a cell producer paying a single fixed cost to produce battery cells for multiple assemblers. We show that such multi-clienting is exceedingly rare once we condition on all observable characteristics of the battery cell.

Figure 5 shows the share of cell plants with multiple clients as we increase the detail of the battery characteristics. The first step of the blue curve shows that more than 65% of cell plants serve more than one assembly firm. This share falls by more than 10 p.p. when the client is an assembly plant. We then separate these relationships by battery cell characteristics. At the detailed material and shape level of the cell, the share of multi-clienting is around 30%. The largest drop occurs when we differentiate by the power of the cell: the share of multi-clienting becomes negligible. For the major models that we will use in our estimation, the share of multi-clienting

Figure 5: Single clients served by a battery plant for a type of battery (2022)



Note: The shares represent the number of battery plants (for a particular battery material, shape and power) that supply multiple carmakers.

is even lower (red curve). This suggests that in almost all cases battery manufacturers produce a specific cell for a single car assembler, rather than supplying a generic type of cell to several buyers.

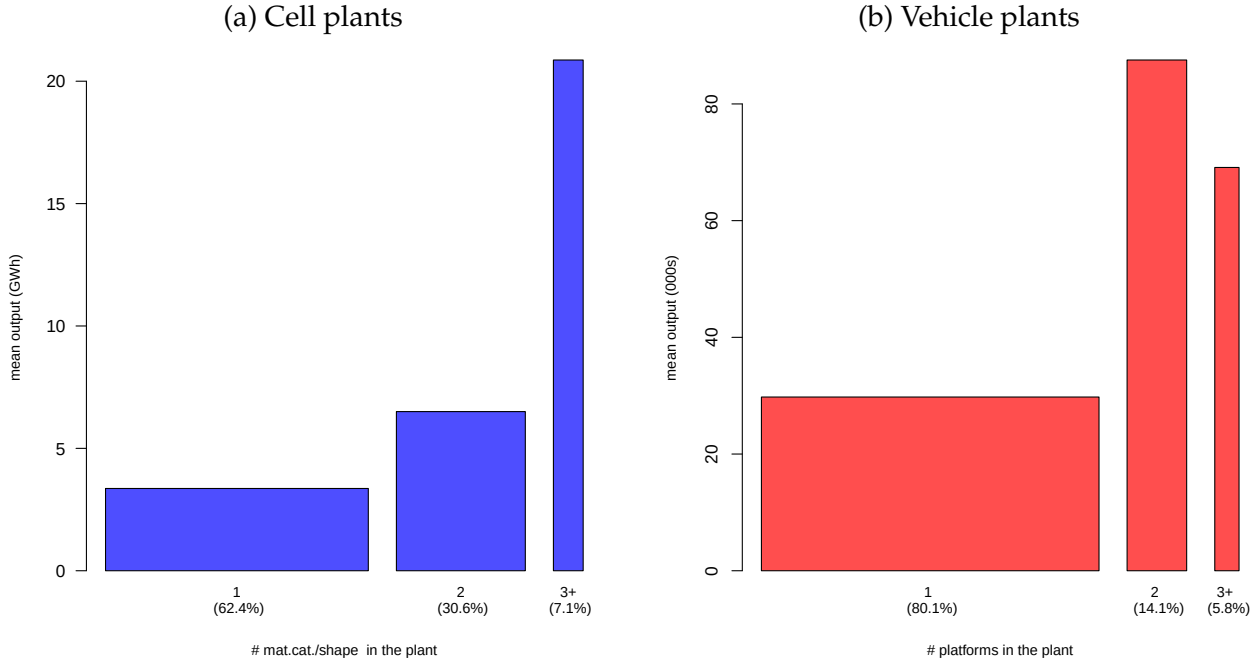
4.3 Aggregation Levels for Fixed Costs

Our integer programming approach for the decision to open facilities allows substantial flexibility in defining the *scope* of a facility that incurs a fixed cost. For vehicle assembly, the fixed cost could be incurred at different aggregation levels, depending on how narrowly the vehicle type is specified. And similarly for cell producers, the fixed cost could be incurred at different aggregation levels for battery types. Guided by the data, we define a fixed cost group at the *platform* level for vehicles and at the combination of *material category and shape* for cells.

Figure 6 shows the 2022 relationship between the size of cell or vehicle plant and the number of groups it produces. In panel (a) we see that almost two thirds of the cell plants produce a single shape/material category. Plants that make two different groups are half as frequent and on average about twice as large. This is what we would expect if adding a second group entailed an additional fixed cost. Plants that produce three or more groups are rare and substantially larger. A similar pattern is shown in panel (b) for vehicles. Only one in five assembly plants uses more than one platform, and they tend to be more than twice as large as the single-platform plants. Taken as a whole, figure 6 suggests that adding a new product group entails large fixed costs at both stages,

which requires a larger scale of production to be profitable.

Figure 6: Size of plant and count of groups (2022)



Note: The share shown in parentheses along the horizontal axis is the share of plants producing the respective number of product groups.

5 Multi-product, Multi-stage, Multinational Production

Guided by the facts in the previous section, we now specify a model of multinational firms' plant location and sourcing decisions. A firm is characterized by a set of models (varieties), $\{m : m \in M_f\}$, and a headquarters location. The models have a vector of "appeal" in each market n. While the firm chooses which models to offer where, they cannot add new models. A global value chain for a given model has levels of production $\{1, \ldots, K\}$, where level 1 is the furthest production stage from consumers and level K the closest stage. The set of potential locations where firms can open facilities, $\mathbf{L}$ is partitioned into K levels: $\mathbf{L} \equiv \{L_1, \ldots, L_K\}$. The index ℓ_k represents a particular production location at level k. A complete path of production (or value chain) is denoted $\boldsymbol{\ell} \equiv \{\ell_1, \ell_2, \ldots, \ell_K\}$. Final consumption happens at level $K+1$, with the location of final consumers (which we also refer to as markets) denoted $n \in N$. Therefore, a path for market n for model m is given by $\boldsymbol{\ell}_{mn} \equiv \{\ell_1, \ldots, \ell_K\}_{mn}$.

5.1 The Firm's Problem

Our model has two types of firm decisions. The first is an *activation* problem: a firm chooses where to open facilities at each level from $k = 1$ to K and which models to offer in which markets (level

$K+1$). The sets of nodes at each stage that paths can follow are given by $\boldsymbol{\mathcal{L}}_f = \{\mathcal{L}_1(f), \ldots, \mathcal{L}_K(f)\}$. Because of fixed costs, firms do not open facilities in every element of L_k nor serve all markets in N. The payment of fixed costs activates a location for groups of models that share common characteristics. The grouping structure at each stage is given by the matrices Γ^k, where $\Gamma^k_{mg_k} = 1$ if model m belongs to group g_k at stage k, and zero otherwise. Importantly, each model is mapped to a unique group at each stage, so $\sum_{g_k \in G_k} \Gamma^k_{mg_k} = 1$ for every model. Firms can add models at a given location without additional fixed costs, so long as they share a common group.[9] Conditional on active facilities and markets, the firm takes chooses *paths* for each car model from cells to final consumers. This second decision is an *assignment* problem. The chosen path $\boldsymbol{\ell}_{mn}$ for model m sold in market n is constrained by $\boldsymbol{\mathcal{L}}_f$.

Figure 7: Schematic of a supply chain with $K = 2$

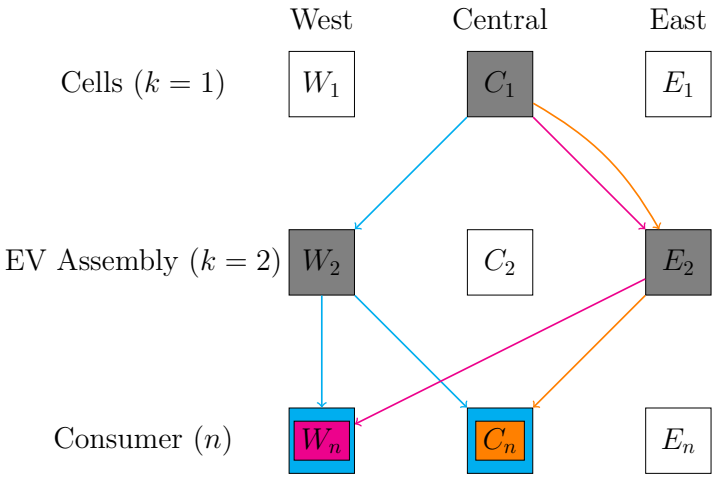


Figure 7 depicts a simplified version of our model to illustrate the decisions we are modeling. As in our application, we set $K = 2$ stages of production for cell manufacturing and vehicle assembly, as well as a final stage of consumer market entry.[10] The schematic portrays a single firm producing three models. All three models share a common stage 1 group. At stage 2, the Magenta and Orange models have a common group, but the Cyan model is in a different group. There are three potential locations: East, Central, and West.

In figure 7, Cyan is offered in West and Central, while Magenta and Orange are only offered in Central. No model was profitable to offer in East. The firm simultaneously decides the locations of facilities along the value chain. In this example, only one location (Central) has an active cell plant; there are two assembly plants (Central is not activated at $k = 2$). Once a firm activates a production facility, it can be used by more than one model, as we see for E_2. The decisions of

[9] We describe the groupings for our empirical application in section 4.3; but our solution methods apply to arbitrary groupings of models at each stage of production.

[10] Our solution method can handle additional production stages subject to the memory requirement associated with keeping track of every potential path assignment. Another important consideration is speed, given that our simulations methods require us to run the optimization algorithm millions of times.

which plants and markets to activate depend on quantities, which are in turn determined by the marginal costs incurred at each node and along each edge.

The firm's problem is to maximize variable profits net of activation costs. Our method relies on the following assumption, which ensures that the firm's variable profits are determined by its chosen path.

Assumption 1. *The choice of path $\boldsymbol{\ell}_{mn}$, combined with a market-level aggregator, A_n, are sufficient to determine the variable profits of model m in market n.*

Here the path enters profits via the variable cost function, $c(\boldsymbol{\ell}_{mn})$, so variable profits can be expressed as $\pi(c(\boldsymbol{\ell}_{mn}), A_n)$. Later, we add a model-market mn demand shifter, but its impact on variable profits will be isomorphic to a cost shifter. Aggregators, often specified as price indexes, exist for a large set of demand systems (including CES and logit) under monopolistic competition.

For a chosen path $\boldsymbol{\ell}_{mn}$, the firm also incurs fixed costs $\phi_{fg_k\ell_k}$ in location ℓ_k for a plant to produce type g_k inputs (or final assembly for $k = K$). We denote this facility location choice with a binary variable $y_{fg_k\ell_k} \in \{0, 1\}$. Let ϕ_{mn} be the fixed market entry cost of selling model m in market n. This includes the cost of marketing, sales training, and space for display. The firm sells all models for which the variable profit $\pi(c(\boldsymbol{\ell}_{mn}), A_n)$ exceeds this cost. We denote these market entry choices with binary variables $z_{mn} \in \{0, 1\}$. The firm's total profit sums variable profits over all of its models m and served markets n, and subtracts the fixed costs of activating facilities and entering markets. In the following section, we describe how the firm maximizes global profits by determining all the binary $y_{fg_k\ell_k}$ facility location choices and z_{mn} market entry choices, while assigning paths $x_{mn\boldsymbol{\ell}}$ through the open facilities for all activated markets.

5.2 Mixed Integer Linear Programming Formulation

The firm's problem described in the previous subsection can be expressed as a linear optimization problem with integer constraints, for which highly efficient algorithms exist. The operations research framework closest to ours is the Multi-level Uncapacitated Facility Location Problem. We add three features to that framework: multi-product firms, endogenous choice of market entry, and endogenous quantities. This leads to the following mixed integer linear programming formulation:

$$\begin{aligned}\max_{\mathbf{x},\mathbf{y},\mathbf{z}} \quad & \sum_{m\in M_f}\sum_{n\in N}\sum_{\ell_1\in L_1}\sum_{\ell_2\in L_2}\pi(c(\boldsymbol{\ell}_{mn}), A_n)x_{mn\ell_1\ell_2} \\ & - \sum_{g_1\in G_1}\sum_{\ell_1\in L_1}\phi_{fg_1\ell_1}y_{fg_1\ell_1} - \sum_{g_2\in G_2}\sum_{\ell_2\in L_2}\phi_{fg_2\ell_2}y_{fg_2\ell_2} - \sum_{m\in M_f}\sum_{n}\phi_{mn}z_{mn} \qquad (1)\end{aligned}$$

subject to

$$\sum_{\ell_1 \in L_1} \sum_{\ell_2 \in L_2} x_{mn\ell_1\ell_2} \quad \leq \quad z_{mn}, \;\; n \in N, m \in M_f \tag{2}$$

$$\sum_{\ell_1 \in L_1} x_{mn\ell_1\ell_2} \quad \leq \quad \sum_{g_2 \in G_2} \Gamma^2_{mg_2} y_{fg_2\ell_2}, \;\; n \in N, m \in \mathcal{M}_f, \ell_2 \in L_2 \tag{3}$$

$$\sum_{\ell_2 \in L_2} x_{mn\ell_1\ell_2} \quad \leq \quad \sum_{g_1 \in G_1} \Gamma^1_{mg_1} y_{fg_1\ell_1}, \;\; n \in N, m \in \mathcal{M}_f, \ell_1 \in L_1 \tag{4}$$

$$x_{mn\ell_1\ell_2} \;\; \geq 0, \qquad y_{fg_1\ell_1} \in \{0,1\}, \quad y_{fg_2\ell_2} \in \{0,1\} \quad z_{mn} \in \{0,1\}. \tag{5}$$

The objective function (1) is the firm's total profits. It sums the variable profits for every selected path $x_{mn\boldsymbol{\ell}} = 1$ and subtracts the fixed cost of all activated facilities $y_{fg_k\ell_k} = 1$ and markets $z_{mn} = 1$. The $M_f \times N$ inequalities defined in (2) are *activity constraints* governing market entry: they ensure that paths are assigned only to models that are chosen to enter the market. At the optimum, if model m is sold in market n, a single $x_{mn\ell_1\ell_2}$ path will be selected and the left-hand-side sum is one; otherwise, all paths equal zero and the sum is zero.[11]

The $M_f \times N \times L_2$ inequalities contained in (3) and the $M_f \times N \times L_1$ contained in (4) are *activity constraints* governing facility activation: they ensure that a facility must be active to be used in supply chains for its corresponding group of cell ($k = 1$) and vehicle ($k = 2$) types. Again, at the optimum, single-sourcing implies that the left-hand-side sum at each stage equals one for activated facilities along the chosen path and zero otherwise. Unlike the facility activation (y) and market entry (z) indicators, the path assignments $x_{mn\ell_1\ell_2}$ do not need to be binary variables (equation 5). Provided the x variables are constrained to be non-negative, they will be binary at the optimum.[12]

The brute force alternative to MILP would involve $2^{|G_1|\times|L_1|+|G_2|\times|L_2|}$ combinations of facility choices. At the median levels of G_k (2 cell types and 4 vehicle platforms) and L_k (12 cell locations and 15 assembly locations) across the firms and locations used in our estimation, this represents 1.93×10^{25} possible combinations. Our MILP approach solves this problem in about one second. Head et al. (2026) explain why this approach is computationally efficient. We detail computation times by number of locations and stages, and also compare performance with the "squeezing" method of Arkolakis et al. (2023).

5.3 Roadmap to Estimation of the Model

While the MUFLP framework is quite general, specific modeling choices are required to estimate parameters and conduct realistic policy simulations. We divide the estimation into three steps. In section 6, we use the assignment (or sourcing) decisions in the data—which are conditional

[11] In the standard MUFLP without endogenous market entry, equation (2) becomes an equality set to 1, called the flow-conservation constraint. It becomes an inequality here because firms only serve markets when it is profitable.

[12] Intuitively, under constant marginal costs with perfectly substitutable inputs across plants and no capacity constraints, it makes no sense to source from a high delivered-cost plant when a lower-cost plant is available. Formally the assignment problem is *totally unimodular*.

on open markets and facilities—to estimate the bilateral variable costs $\tau^k_{\ell_k \ell_{k+1}}$ for all locations at each level of production. Section 7 shows how to use these parameters to recover variable profits in levels. In section 8, we estimate all the remaining parameters using simulated method of moments matching those generated from the MUFLP solution to their counterparts in data, subject to market equilibrium as constraints. These parameters include those governing the distribution of fixed costs, which we describe later in that section.

6 Estimation of Path Costs via Sourcing Regressions

Since our model satisfies Assumption 1, path costs and the market-level aggregator (A_n) are sufficient to determine variable profits for each model in each market. Conditional on the facility and entry vectors, y and z, total profit maximization subsumes cost minimization through the choice of x, which we refer to as the sourcing decision. In this section, we use the data on sourcing decisions to estimate costs on each "edge" (origin-destination pair) along the path from cell manufacturing to EV consumer. Specifically, we estimate the impacts of distance, borders, trade agreements and tariffs on variable costs.

6.1 Path Costs

We begin by specifying functional forms for the costs that will determine assignment and location decisions. Any production function that characterizes total path cost as a function of the chosen path can be used for the firm's optimization problem. Leontief and Cobb-Douglas production functions yield cost functions that are readily suitable for estimation via discrete choice sourcing regressions because of their separability. These cases can be viewed as first-order approximations to more general cost functions (or log cost in the case of Cobb-Douglas). We follow Yi (2003), Yi (2010), Antràs and De Gortari (2020), and Johnson and Moxnes (2023) in using Cobb-Douglas across stages. We prefer Cobb-Douglas because it expresses cost differences as percentages, making it well-suited for ad valorem duties, whereas Leontief's additive structure is more compatible with specific duties.[13]

The variable costs for a particular chain depend on three key aspects: (i) a location-stage production cost shifter, $w^k_{\ell_k}$, comprising wages and other local material costs, (ii) trade costs $\tau_{\ell_1 \ell_2}$ between cell production and vehicle assembly, and $\tau_{\ell_2 n}$ between assembly and consumers, (iii) model-specific cost shocks attributable to bilateral edges $\varepsilon^1_{m\ell_1\ell_2}$ and $\varepsilon^2_{m\ell_2 n}$.[14]

[13] Leontief might seem the natural choice when combining stages of production. For example, each car needs two axles and four wheels. However, Cobb-Douglas allows labor at an upstream stage to substitute for labor at a downstream stage. One justification is in terms of subassemblies. In our case, cells are usually combined into modules that are then combined into packs. The module and pack production could occur at the cell factory, at an intermediate location, or at the vehicle factory. The former cases would increase labor input upstream relative to the last case which requires more labor at the assembly factory. Cobb-Douglas also implies constant cost shares across stages. This is appealing for EVs because as the size, power and electric range rise, battery capacity requirements increase proportionately.

[14] The subscript m appears in the shock term and explains why sourcing decisions vary across models. We will provide parametric assumptions for the ε later when we estimate these cost equations using discrete choice.

The marginal cost of cells (produced in ℓ_1, delivered to ℓ_2) and of vehicles (assembled in ℓ_2, delivered to n) are given by

$$\text{Cells:} \quad c^1_{m\ell_1\ell_2} = w^1_{\ell_1} \tau^1_{\ell_1\ell_2} \varepsilon^1_{m\ell_1\ell_2}, \tag{6}$$

$$\text{Vehicles:} \quad c(\boldsymbol{\ell}_{mn}) = \left(w^2_{\ell_2}\right)^{\alpha_{22}} \left(c^1_{m\ell_1\ell_2}\right)^{\alpha_{12}} \tau^2_{\ell_2 n} \varepsilon^2_{m\ell_2 n}, \tag{7}$$

where $\alpha_{k2} \in (0,1)$ denotes the cost share of stage $k = 1, 2$ factors in final assembly ($k = 2$). The above costs include those incurred at production facilities and along transport edges. Costs at the final consumption stage—sales taxes, buyer subsidies, variable costs of distribution—can be incorporated in $\tau^2_{\ell_2 n}$.

We can also extend these marginal cost determinants to include costs of additional inputs, other than battery cells and vehicle assembly. Assuming negligible trade costs of these inputs to cell or assembly facilities, their costs enter as additive constant in the log-cost function and do not affect the sourcing regressions in the next section. We therefore suppress them in the notation for simplicity.

Substituting upstream cost (6) into the next stage's cost equation (7), the log cost of the complete production path is

$$\ln c(\boldsymbol{\ell}_{mn}) = \alpha_{22} \ln w^2_{\ell_2} + \ln \tau^2_{\ell_2 n} + \alpha_{12} (\ln w^1_{\ell_1} + \ln \tau^1_{\ell_1\ell_2}) + u(\boldsymbol{\ell}_{mn}), \tag{8}$$

where $u(\boldsymbol{\ell}_{mn}) = \alpha_{12} \ln \varepsilon^1_{m\ell_1\ell_2} + \ln \varepsilon^2_{m\ell_2 n}$ is the structural interpretation of the error term. The optimal sourcing decisions, x, minimize log costs conditional on the open facilities at both production stages ($y^1_{fg_1\ell_1}$ and $y^2_{fg_2\ell_2}$) and served markets (z_{mn}).

6.2 Conditional Choice Estimation of Variable Costs

The probability of choosing path $\boldsymbol{\ell}_{mn}$ is given by the product of the conditional probabilities at each stage:

$$\mathbb{P}(\boldsymbol{\ell}_{mn}) = \mathbb{P}^1_{\ell_1|\ell_2}(m) \times \mathbb{P}^2_{\ell_2|n}(m). \tag{9}$$

Decomposing the probability into these logarithmically separable components greatly facilitates estimation. There are several ways to obtain a tractable formulation of the choice probabilities at each stage. The best known method is the nested logit approach by McFadden (1978), modeling $u(\boldsymbol{\ell}_{mn})$ as single shock with a generalized extreme value (GEV) distribution that allows for correlated errors. Here we follow the presentation by Anderson et al. (1992, chapter 2), where the correlation is driven by two parameters, θ_1 and θ_2, one for each production stage. Intuitively, larger values of θ_k implies that alternatives within a stage k are closer substitutes. If $\theta_1 = \theta_2$, the error term $-u(\boldsymbol{\ell}_{mn})$ is a Gumbel shock with scale parameter $1/\theta_2$, as in Antràs and De Gortari (2020) and Allen and Arkolakis (2022).

Conditional on its set of level-1 facilities, $\mathcal{L}_1(f)$, the firm chooses, for each model, the location that minimizes the cost of serving ℓ_2. We can therefore combine equation (6) with our GEV

assumption to derive the conditional probability of ℓ_1 being the cost-minimizing location of supplying to ℓ_2 as:

$$\mathbb{P}^1_{\ell_1|\ell_2}(m) = \frac{\left(w^1_{\ell_1}\tau^1_{\ell_1\ell_2}\right)^{-\theta_1}}{\Phi^1_{\ell_2}(m)}, \text{ with } \Phi^1_{\ell_2}(m) = \sum_{g_1 \in G_1} \Gamma^1_{mg_1} \sum_{\ell_1 \in L_1} y_{fg_1\ell_1} \left(w^1_{\ell_1}\tau^1_{\ell_1\ell_2}\right)^{-\theta_1}, \tag{10}$$

where $\Gamma^1_{mg_1} = 1$ if model m uses group g_1 for cells, and $y_{fg_1\ell_1} = 1$ if location ℓ_1 has activated this group g_1. Our estimating equation is

$$\mathbb{P}^1_{\ell_1|\ell_2}(m) = \exp\left[\text{FE}^1_{\ell_1} + \text{FE}^1_{\ell_2}(m) + \beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right], \tag{11}$$

where $D_{\ell_1\ell_2}$ is a measure of bilateral "distance" (any observable bilateral friction), $t^1_{\ell_1\ell_2}$ is the bilateral tariff for stage 1, with β^1_D and β^1_t being their respective regression coefficients. The term $\text{FE}^1_{\ell_1}$ is a source ℓ_1 fixed effect. The second fixed effect is specific to the *chooser* and is given by

$$\text{FE}^1_{\ell_2}(m) = -\ln\left\{\sum_{g_1 \in G_1} \Gamma^1_{mg_1} \sum_{\ell_1 \in L_1} y_{fg_1\ell_1} \exp\left[\text{FE}^1_{\ell_1} + \beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right]\right\}.$$

This chooser fixed effect is more specific than the assembly plant location as the chooser in our estimation corresponds to a plant, firm, model and the power and chemistry of the battery being sourced.[15] The correspondence between estimated coefficients and structural variables is

$$\begin{aligned} w^1_{\ell_1} &= \exp\left[-\text{FE}^1_{\ell_1}/\theta_1\right] \\ \tau^1_{\ell_1\ell_2} &= \exp\left\{-\left[\beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right]/\theta_1\right\} \\ \Phi^1_{\ell_2}(m) &= \exp\left[-\text{FE}^1_{\ell_2}(m)\right] \end{aligned}$$

Since tariffs are direct cost shifters, and they have a unit pass-through elasticity, the coefficient β^t_1 estimates θ_1 (as in Head and Mayer, 2019).

An important component of the nesting approach to estimating production costs along the value chain is the integration of upstream costs into downstream decisions. The nested logit approach replaces the actual cost of cells (which includes the error term) with the expected cost of the optimal cell supplier. With extreme value distributions, this expected cell cost for assembly plant ℓ_2 is

$$\mathbb{E}c^1_{\cdot,\ell_2}(m) = \left(\Phi^1_{\ell_2}(m)\right)^{-1/\theta_1}. \tag{12}$$

We refer to this term as the *inclusive cost*; it depends on the sourcing potential of ℓ_2 and differs across car models due to various choice sets of the firm $\mathcal{L}_1(f)$ and the cell group g_1 used by m.

The second decision concerns the sourcing of EVs from assembly plants for each model-market

[15] These narrow definitions of choosers ensure single-sourcing in almost all cases, as we see in figure 4.

combination. Using (7), the probability that ℓ_2 is selected as the location of assembly is

$$\mathbb{P}^2_{\ell_2|n}(m) = \frac{\left[\left(w^2_{\ell_2}\right)^{\alpha_{22}}\left(\mathbb{E}c^1_{\cdot,\ell_2}\right)^{\alpha_{12}}\tau^2_{\ell_2 n}\right]^{-\theta_2}}{\Phi^2_n(m)}, \text{ with}$$
$$\Phi^2_n(m) = \sum_{g_2 \in G_2}\Gamma^2_{mg_2}\sum_{\ell_2 \in L_2} y_{fg_2\ell_2}\left[\left(w^2_{\ell_2}\right)^{\alpha_{22}}\left(\mathbb{E}c^1_{\cdot,\ell_2}(m)\right)^{\alpha_{12}}\tau^2_{\ell_2 n}\right]^{-\theta_2}. \tag{13}$$

The estimating equation for destination markets when choosing among potential vehicle assembly facilities is

$$\mathbb{P}^2_{\ell_2|n}(m) = \exp\left[\text{FE}^2_{\ell_2} + \text{FE}^2_n(m) + \beta^2_D \ln D_{\ell_2 n} + \beta^2_t \ln\left(1 + t^2_{\ell_2 n}\right) + \beta^2_\Phi \text{FE}^1_{\ell_2}(m)\right], \tag{14}$$

where $\text{FE}^1_{\ell_2}(m)$ was estimated in equation (11), and accounts for determinants of the upstream (cells) costs. The correspondence between the estimated coefficients and structural parameters is

$$\begin{aligned}
\left(w^2_{\ell_2}\right)^{\alpha_{22}} &= \exp\left(-\text{FE}^2_{\ell_2}/\theta_2\right)\\
\tau^2_{\ell_2 n} &= \exp\left\{-\left[\beta^2_D \ln D_{\ell_2 n} + \beta^2_t \ln\left(1 + t^2_{\ell_2 n}\right)\right]/\theta_2\right\}\\
\alpha_{12} &= -\theta_1\beta^2_\Phi/\theta_2\\
\Phi^2_n(m) &= \exp\left[-\text{FE}^2_n(m)\right]
\end{aligned}$$

Again, as tariffs shift costs and final prices proportionately, the coefficient β^t_2 provides a direct estimate of θ_2. The Antràs and De Gortari (2020) approach of a single independent path shock would imply $\theta_1 = \theta_2$ in the two probability equations.[16] As we estimate the two equations separately, we do not impose this coefficient restriction. We use the equivalence in terms of likelihood functions between Poisson (with fixed effects) and multinomial logit estimators (Guimaraes et al., 2003), which enables us to estimate equations (11) and (14) as high-dimensional fixed effects Poisson Pseudo Maximum Likelihood (PPML) regressions.

Combining the two production stages, we can express the log variable cost of the full path as a function of our current estimates and additional parameters:

$$\begin{aligned}
\ln c(\boldsymbol{\ell}_{mn}) = &-\kappa\Big\{\text{FE}^2_{\ell_2} + \beta^2_D \ln D_{\ell_2 n} + \beta^2_t \ln\left(1 + t^2_{\ell_2 n}\right)\\
&-\beta^2_\Phi\left[\text{FE}^1_{\ell_1} + \beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right]\Big\} + u(\boldsymbol{\ell}_{mn}).
\end{aligned} \tag{15}$$

The Gumbel scale parameter $\kappa \equiv 1/\theta_2$ governs the relative importance of observed determinants of marginal costs versus fixed costs. Assuming that ad valorem tariffs on vehicles $(1 + t_{\ell_2 n})$ proportionately shift path cost $c(\boldsymbol{\ell}_{mn})$, the parameter κ is identified from the tariff effect on vehicle sourcing decisions: $\kappa = -1/\beta^2_t$. The coefficient on the inclusive cost of cells provides the weight

[16] As noted in the appendix of Antràs and De Gortari (2020) the IID path shock also implies that the product of each stage's conditional probability collapses to a single conditional logit formula, referred to as the joint logit in Ben-Akiva and Lerman (1985).

on the cell stage relative to vehicle assembly. Our model implies $-\beta_{\Phi}^2 = \alpha_{12}\,(\theta_2/\theta_1)$. For reasons outlined in Ben-Akiva and Lerman (1985), our prior is that $\theta_2 \leq \theta_1$. Consequently, $-\beta_{\Phi}^2$ is a lower bound of the α_{12} cost share. In practice, our tariff estimates at both stages imply that $\theta_2 \approx \theta_1$, so that $-\beta_{\Phi}^2$ is close to α_{12}.

6.3 Filters Applied Before Estimation

Given the computational demands of this high-dimensional discrete-choice problem and the subsequent SMM estimation, we restrict the number of markets, firms, and models. Because our counterfactuals consider policy changes only in North America and Europe, we exclude 29 "China-only" firms that are dedicated to the Chinese market, which we define as firms selling fewer than 2,000 cars outside China.[17] The precise selection criteria for the filtered sample are:

1. Restrict world sales to the 24 countries that dominate either EV production (cell or assembly) or EV sales. These countries account for 99.9% of cells produced, 98.6% of EVs produced, and 97.1% of EV sales.[18]

2. Keep the 15 largest firms, ranked by world sales after excluding China-only firms. These top 15 firms, listed in table 3, account for 99.5% of world production excluding China-only firms.

3. Restrict to the 137 models offered by the 15 large firms that have price data available in 2022. Prices are needed in the SMM for two reasons: i) we use market shares in value to compute variable profits, ii) we allow fixed costs to vary with the implied quality of the model.[19]

None of the above filters is required for the sourcing estimation, since car prices are irrelevant for sourcing decisions, and PPML is computationally fast. We nevertheless apply these filters to ensure the sourcing estimates are aligned with the sample used later in the SMM.[20]

6.4 Sourcing Estimation Results

Table 4 reports the two-step nested discrete choice estimates of cost parameters in the EV industry.[21] They characterize variable costs of production and trade costs from cells to vehicle assembly, and from assembly to dealerships in the 24 markets.

[17] The "China-only" firms account for 23.4% of world sales, but only 0.3% of sales outside China.

[18] This definition of world comprises Austria, Belgium, Canada, China, Czechia, Denmark, Finland, France, Germany, Hungary, Italy, Japan, Mexico, Netherlands, Norway, Poland, Portugal, Slovakia, South Korea, Spain, Sweden, Switzerland, United Kingdom, United States.

[19] Appendix C details how we retrieve quality for each car model, combining quantity and price data.

[20] As we detail in section 8, the SMM combines plants into larger areas, but the sourcing regressions do not impose this aggregation because the plants' point locations give more precise distance and country information that is helpful for estimating the distance and border effects.

[21] Additional specifications that vary the control variables for each stage are shown in the online appendix tables O.A1 and O.A2. Those estimations also consider the whole sample available without the restriction imposed on major countries and firms.

Table 3: Top 15 firms in 2022

No.	Firm	# Markets	# Models	Sales	Cum. shr.
1	Tesla	23	4	1,296,470	22.8
2	BYD	10	14	903,239	38.7
3	SAIC	17	21	807,223	52.9
4	Volkswagen	24	23	600,159	63.4
5	Geely	24	15	394,257	70.3
6	Hyundai	23	15	351,553	76.5
7	Stellantis	18	19	297,034	81.7
8	BMW	24	8	237,472	85.9
9	Renault	19	6	168,535	88.9
10	Mercedes-Benz	24	9	164,960	91.8
11	Nissan-Mitsubishi	23	8	158,739	94.6
12	Ford	22	4	124,818	96.8
13	General Motors	7	7	83,013	98.2
14	Toyota	23	8	49,374	99.1
15	Rivian	3	3	24,334	99.5

Number of markets and sales are restricted to the 24 countries. Sample excludes firms that sell fewer than 2,000 cars outside China.

Table 4: Combined sourcing regression results

Stage:	(1) Cells	(2) Vehicle
Border	-0.953^a	-1.04^a
	(0.319)	(0.254)
log distance	-0.382^a	-0.112^c
	(0.021)	(0.062)
RTA	0.458	0.869^a
	(0.320)	(0.214)
Inc. cost of cells		-0.234^a
		(0.084)
log GDP per capita	0.213^c	0.206^b
	(0.118)	(0.087)
log(1+tariff)	-8.49^a	-8.56^a
	(2.34)	(1.74)
Observations	7,945	15,793
Squared Correlation	0.322	0.265

Note: The inclusive cost of cells are the expected costs of cells for assemblers, $\mathbb{E}c^1_{\cdot,\ell_2}$, defined in equation (12). Log GDP per capita is measured at the sub-national level.

Geographic proximity clearly matters for sourcing of both cells and vehicles. Border effects are large: the coefficients imply that domestic sourcing is 2.6–2.8 times as likely as importing, holding other things equal. Distance (measured using the great circle formula) affects cell sourcing but has a less significant effect on vehicle sourcing. Tariffs strongly influence sourcing decisions at both stages. The tariff elasticities are -8.49 and -8.56 for cells and vehicles, respectively, and are not significantly different from each other. The coefficient on vehicle sourcing is remarkably similar to the one obtained by Head and Mayer (2019). RTAs, which removes non-tariff barriers to trade among member countries, strongly affect EV sourcing but less so for battery cells.

As we detailed above, the coefficient on the inclusive cost of cells in the vehicle sourcing regression (Column 2) provides the weight to use on the cell stage in equation (15). We can use the coefficient to determine the implied cost share of cells: $\alpha_{12} = -\theta_1\beta_\Phi^2/\theta_2$. As $\beta_\Phi^2 = -0.234$ and estimated $\theta_1 \approx \theta_2$, the implied cost share of cells is around 23%. AlixPartners (2021) computes the share of batteries including packs and the battery management system as 8,000 EUR or 34% of the car they studied. Cells accounted for 75% of that cost, or 25.5% of the vehicle cost. This figure is remarkably close to our estimate, although the coefficient on inclusive cell cost varies across specifications (as seen in table O.A2 of the online appendix).

The regressions also include source country fixed effects, as well as sub-national (state/province) productivity variation measured using log GDP per capita.[22] However, we do not use these location-specific marginal cost and productivity estimates. The choice sets of open facilities for each firm are subject to selection bias: our framework predicts that firms open facilities based on a combination of low variable and fixed costs.[23] We therefore use the SMM estimation developed in the next sections to recover both cost terms jointly.

7 From Path Costs to Variable Profits

The nested logit parameters allow us to compute (log) cost *differences* between any two paths based on their observable characteristics. Next, we need to compute—as a function of the remaining parameters—the variable profit for any potential path in *levels*. This is because the firms' optimization problem shown in equation (1) subtracts fixed costs from the levels of variable profits. We recover variable profits in levels as a function of path cost differences, using Exact Hat Algebra (EHA), pioneered by Dekle et al. (2007).

[22] Details on country fixed effects and sub-national GDP per capita are provided in the online appendix O.A.2 and appendix B.6, respectively.

[23] A second concern of the fixed effects for some countries is that they are identified by the choices of one or two models, leading to over-fitting. Despite these two concerns, the fixed effects retain a useful purpose in the sourcing regressions. Conditioning on them allows us to recover consistent estimates of the dyadic frictions.

7.1 Demand and Market Structure

Use of EHA is facilitated by the constant elasticity of substitution (CES) form for demand.[24] One convenient feature of CES is that the average own-price elasticity is widely reported in the literature. Table B1 in the appendix provides 18 different estimates, taken mainly from the Industrial Organization literature. We set $\eta = 4.0$, the median of the negative of those estimates. Each country n spends a fixed amount $R_n = \sum_m p_{mn} q_{mn}$ on light vehicles of all fuel types, where p_{mn} are delivered prices and q_{mn} are quantities. The price of model m in country n incorporates tariffs and consumer subsidies.

The EV share R_n^{EV}/R_n under CES is given by $(P_n^{\text{EV}}/P_n)^{1-\eta}$, where the P_n^{EV} and P_n are the price index for EVs and all light vehicles. The market share of model m (in values, as a fraction of *total* expenditures) depends on the ratio of the model's appeal-adjusted price p_{mn}/ξ_{mn} to P_n:

$$s_{mn} \equiv \frac{p_{mn} q_{mn}}{R_n} = \left(\frac{p_{mn}}{\xi_{mn}}\right)^{1-\eta} (P_n)^{\eta-1}. \tag{16}$$

The all-vehicles price index, P_n, is decomposable as

$$(P_n)^{1-\eta} = \left(P_n^{ICE}\right)^{1-\eta} + \left(P_n^{EV}\right)^{1-\eta}, \qquad \text{with} \qquad P_n^{EV} = \left[\sum_m z_{mn} \left(\frac{p_{mn}}{\xi_{mn}}\right)^{1-\eta}\right]^{\frac{1}{1-\eta}}, \tag{17}$$

where z_{mn} is set to 1 for all models m sold in n (and 0 otherwise). The price index for internal combustion engine vehicles, P_n^{ICE}, is exogenous. To simplify computation, we approximate price/cost markups as $\eta/(\eta-1)$.[25] Prices are therefore given by $p_{mn\boldsymbol{\ell}} = [\eta/(\eta-1)] c_{mn\boldsymbol{\ell}}$, where delivered marginal costs, $c_{mn\boldsymbol{\ell}}$, have two pieces. The first is the path cost $c(\boldsymbol{\ell}_{mn})$, to which we return below. The second piece is a model-level cost shifter υ_m. It incorporates all causes of cross-model cost differences that do not depend on the path, such as the variable costs of producing a higher quality vehicle (more steel, more electronic sensors, etc.).

Combining these assumptions, variable profits can be expressed as a function of the path $\boldsymbol{\ell}_{mn}$:

$$\pi_{mn\boldsymbol{\ell}} = \frac{R_n s_{mn\boldsymbol{\ell}}}{\eta}, \qquad \text{with} \qquad s_{mn\boldsymbol{\ell}} = \left(\frac{\eta}{\eta-1} \frac{c(\boldsymbol{\ell}_{mn}) \upsilon_m}{\xi_{mn}} \frac{1}{P_n}\right)^{1-\eta}. \tag{18}$$

Any *change* in path $\boldsymbol{\ell}$ shifts variable profits of a model m sold in n through *changes* in $(c(\boldsymbol{\ell}_{mn})/P_n)^{1-\eta}$. The non-path-specific factors—R_n, υ_m, ξ_{mn}, $\eta/(\eta-1)$—all cancel. This completes the specification

[24] This assumption differs from most of the literature on the car industry which uses nested logit or mixed logit demands (Van Biesebroeck and Verboven, 2025). The main difference is that logit demand has price responsiveness parameter α that multiplies price, while CES has an elasticity η that multiplies log price. In the discrete choice derivation of logit demand, α corresponds to the marginal utility of income and $1/\alpha$ is the limiting value of the *additive* markup as market shares become small. There are two problems with this here: first, poorer countries should have higher marginal utilities of income; second, a markup of \$10,000 seems reasonable for a Tesla priced at \$50,000 but not for a Chinese model priced at \$15,000. These issues could be addressed with a household-specific α as is common in mixed logit models, but we would lose the tractability needed for our facility-location problem.

[25] Endogenizing markups would not matter much with this demand structure because individual EV firm market shares are small relative to all vehicles. We leave a thorough consideration of oligopoly pricing to future research.

of $\pi(c(\boldsymbol{\ell}_{mn}), A_n)$ used in the formulation of the firm's total profit shown in equation (1) (the MU-FLP objective function). Variable profits depend on the path and the aggregator ($A_n = P_n^{\eta-1}$), thus complying with Assumption 1. We now show how $\pi_{mn\boldsymbol{\ell}}$ for any potential path can be recovered from observed market shares, along with its cost change relative to the observed path.

7.2 Hat Algebra for Variable Profits

Following the notational standard established by Dekle et al. (2007), a variable with a "hat" on it, refers to the ratio of a model-predicted variable relative to its observed value. We use superscript $^\circ$ to denote the car industry configuration (comprising expenditure, paths, market shares, and the price index in each market) that we observe in the 2022 data. For each market n, let R_n° denote total expenditures on cars. This is associated with a reference (unobserved) price index P_n°. For each market n in our data, we can also measure $s_n^{\circ,\text{ICE}}$ as the aggregate market share of ICEs and s_{mn}° as the market share for EV model m in market n. These $^\circ$ values allow us to compute equilibrium outcomes for all the potential configurations relative to the observed configuration.

For any potential path $\boldsymbol{\ell}_{mn}$, the *changes* in path costs relative to the path *observed* in the data, denoted $\boldsymbol{\ell}_{mn}^\circ$, is given by

$$\hat{c}_{mn\boldsymbol{\ell}} \equiv \frac{c_{mn\boldsymbol{\ell}}}{c_{mn\boldsymbol{\ell}^\circ}} = \frac{c(\boldsymbol{\ell}_{mn})}{c(\boldsymbol{\ell}_{mn}^\circ)}. \tag{19}$$

Using (16) and assuming model-market appeal ξ_{mn} is invariant, we can express proportional changes in market shares as a function of changes induced by new paths as well as price index changes that aggregate the $\hat{c}_{mn\boldsymbol{\ell}}$ across all models:

$$\hat{s}_{mn\boldsymbol{\ell}} \equiv \frac{s_{mn\boldsymbol{\ell}}}{s_{mn}^\circ} = \left(\frac{\hat{c}_{mn\boldsymbol{\ell}}}{\hat{P}_n}\right)^{1-\eta}, \tag{20}$$

where $\hat{P}_n \equiv P_n/P_n^\circ$ is the (unobserved) change in the price index for market n. We describe later how to recover the implied endogenous change in market demand. Profit $\pi_{mn\boldsymbol{\ell}}$ from (18) can now be expressed as a function of cost *changes* for any path:

$$\pi_{mn\boldsymbol{\ell}} = \frac{1}{\eta} s_{mn}^\circ R_n^\circ \left(\frac{\hat{c}_{mn\boldsymbol{\ell}}}{\hat{P}_n}\right)^{1-\eta}. \tag{21}$$

Hat algebra is helpful in obtaining variable profits by path *in levels* because it allows the market shares in the data to absorb many parameters that would otherwise need to be calibrated. Specifically, the observed market shares s_{mn}° absorb the demand shifters ξ_{mn}, the cost shifters υ_m, the markup $\eta/(1-\eta)$, and the *level* of the price index, P_n°, that prevails in the observed equilibrium. A limitation of this approach is that we cannot obtain variable profits for a model in markets where it is not observed to be offered. Consequently, we cannot allow models to enter markets in the simulation that they do not enter in the data.

The last computation needed to obtain the level of variable profits in (21) is $\hat{P}_n$. Hat algebra

for $P_n^{1-\eta}$ yields:

$$\hat{P}_n^{1-\eta} = s_n^{\circ,\text{ICE}} + \sum_{m \in \mathcal{M}_n^\circ} s_{mn}^\circ \hat{c}_{mn}^{1-\eta}, \tag{22}$$

where $\mathcal{M}_n^\circ$ denotes the set of all EV models sold in market n in our data, and the observed market share of ICE models in n is

$$s_n^{\circ,\text{ICE}} = 1 - \sum_{m \in \mathcal{M}_n^\circ} s_{mn}^\circ. \tag{23}$$

In (22), $\hat{c}_{mn} = c(\boldsymbol{\ell}_{mn}^*)/c(\boldsymbol{\ell}_{mn}^\circ)$ (with the path $\boldsymbol{\ell}$ dropped in subscript) is defined in terms of the *optimal* chosen path $\boldsymbol{\ell}_{mn}^*$ in the MUFLP optimization. For models $m \in \mathcal{M}_n^\circ$ that are optimally dropped from market n, $\hat{c}_{mn}^{1-\eta} = 0$.

8 Simulated Method of Moments Estimation

In the previous section, we relied on hat algebra to reduce the number of parameters in the estimation (there are implicitly 3288 different ξ_{mn}). This method relies on observing initial market shares (s_{mn}°). We cannot use this parameterization approach to recover variable profits for markets that were not entered in the data.[26] Therefore, we treat the ϕ_{mn} as sunk costs: specifically, they are zero for markets entered in the data and infinite for those markets not entered. We therefore set the z_{mn} to zero for any market the model has not entered in the data. The firm chooses the optimal binary z_{mn} for all the remaining markets. The firm may choose to set $z_{mn} = 0$ for all markets n (even those with positive sales in the data) when the production of model m is unprofitable. This occurs when adding a model would entail a facility fixed cost (for that model's platform) that exceeds the variable profits that model generates. This "soft" z constraint will help us recover the fixed cost parameters by targeting the number of models sold in each market.

When estimating the SMM, we aggregate the individual plant locations (used in the sourcing estimation) into larger geographic units. Potential locations (the sets L_k) are the units within the 24 countries with observed production in the relevant stage in 2022. For locations that do not produce in the data, there is no $\boldsymbol{\ell}_{mn}^*$ that we can use to infer marginal cost changes. The aggregation yields a set of feasible locations of size $|L_1| = 12$ and $|L_2| = 15$, though the effective choice space is larger because firms choose both where to activate and which production lines to open at each location (e.g., NMC versus LFP cells).[27] To maintain parsimony in the SMM estimation, we estimate mean marginal costs at the continent-level; location-level *delivered* marginal costs vary due to borders, distances, and path cost shocks. We define three continents, $\mathcal{N} \in \{\text{Am}, \text{As}, \text{Eu}\}$.[28]

[26] The alternative approach taken by Castro-Vincenzi et al. (2025) is to separately estimate the demand system in product characteristic space. This allows them to incorporate model entry into new markets and recover bounds for the associated fixed costs.

[27] Geely, the firm with the maximum number of cell types (3) and platforms (8), has $2^{3\times12+8\times15} \approx 9.13 \times 10^{46}$ possible configurations.

[28] The geographic units we group as "Am" comprise four regions in the United States plus Mexico. "Eu" combines the UK and four subregions of the European Union (Germany+Austria, France+Belgium, Eastern Europe, and Italy+Spain+Portugal). "As" comprises Japan, South Korea, and three regions within China. The three L_2 locations not

8.1 Remaining Parameters: Fixed Effects and Fixed Costs

The remaining parameters we need to fully determine the firms' objective function in (1) are variable cost shifters $\text{FE}^k_{\mathcal{N}(\ell_k)}$ and fixed costs $\phi_{fg_k\ell_k}$. We parameterize the fixed cost $\phi_{fg_k\ell_k}$ as a draw from a lognormal distribution

$$\ln \phi_{fg_k\ell_k} \sim \text{Normal}\left(\ln\left[\mu_{fg_k\ell_k} \times R^{\text{EV}}_W\right], \sigma_k\right), \\ \text{with } \mu_{fg_k\ell_k} = \exp\left[\ln \rho^k_{\mathcal{N}(\ell_k)} + \rho^{\text{HQ},k} \ln D_{h(f)\ell_k} + \rho^{\xi,k} \ln \tilde{\xi}_{fg_k}\right], \tag{24}$$

where $\rho^k_{\mathcal{N}}$ denotes the expected fixed cost of a single production line in continent $\mathcal{N}$ for stage k as a share of *observed* world EV revenue, R^{EV}_W. The parameter $\rho^{\text{HQ},k}$ denotes the elasticities of fixed costs with respect to distance from firm headquarters, denoted $h(f)$, for each stage of production. Because fixed costs are likely higher for plants producing more complex and powerful varieties, we also allow them to vary with firm-level quality within each grouping of cells and car models. Specifically, $\rho^{\xi,k}$ is the elasticity of expected fixed costs with respect to quality, $\tilde{\xi}_{fg_k}$, which corresponds to the geometric mean of ξ_{mn} across the destinations for the relevant fg_k.[29] For the rest of the paper, we define $\boldsymbol{\rho} \equiv \{\rho^k_{\mathcal{N}}, \rho^{\text{HQ},k}, \rho^{\xi,k} \mid \mathcal{N} \in \{\text{Am}, \text{As}, \text{Eu}\}, k \in \{1,2\}\}$ as the vector of 10 parameters that determine expected log fixed costs. We use the same continental grouping by $\mathcal{N} \in \{\text{Am}, \text{As}, \text{Eu}\}$ for the location-specific marginal cost differences, reducing those fixed effects to a vector $\mathbf{FE} \equiv \{\text{FE}^k_{\mathcal{N}} \mid \mathcal{N} \in \{\text{As}, \text{Eu}\}, k \in \{1,2\}\}$ of 4 parameters, normalizing the parameters for $\mathcal{N} = \text{Am}$ to zero. Overall, there are thus 16 parameters $\{\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}\}$ to estimate through SMM.

8.2 Selection Correction for Observed Paths

We compute the cost difference $\hat{c}_{mn\boldsymbol{\ell}} = c(\boldsymbol{\ell}_{mn})/c(\boldsymbol{\ell}^\circ_{mn})$ for every potential path $\boldsymbol{\ell}_{mn}$ relative to the path $\boldsymbol{\ell}^\circ_{mn}$ observed in the data. This requires draws of the shocks $u(\boldsymbol{\ell}_{mn})$ from our sourcing regression in section 6. Those shocks are distributed $u(\boldsymbol{\ell}_{mn}) \sim -\text{Gumbel}(0, \kappa)$, where $\kappa = 1/8.56 = 0.12$ using the inverse of estimated tariff coefficient in Column 2 of table 4. However, the shock for the observed path $u(\boldsymbol{\ell}^\circ_{mn})$ is not a random draw from this distribution, but the realization that led the firm to choose $\boldsymbol{\ell}^\circ_{mn}$. For example, if the path observables for $\boldsymbol{\ell}^\circ_{mn}$ imply a high cost—equivalently, a low estimated probability $\mathbb{P}(\boldsymbol{\ell}^\circ_{mn})$ in equation (9)—then the actual $u(\boldsymbol{\ell}^\circ_{mn})$ draw must have been low in order to induce the firm to choose that path.

Following Hanemann (1984), we correct for this selection bias by exploiting the fact that the distribution of $u(\boldsymbol{\ell}_{mn})$—*conditional* on path $\boldsymbol{\ell}^\circ_{mn}$ being chosen—is also Gumbel, but with a location parameter shifted by $-\kappa \ln \mathbb{P}(\boldsymbol{\ell}^\circ_{mn})$. Therefore, we calculate $c(\boldsymbol{\ell}^\circ_{mn})$ using a draw from the following

included in L_1 because they do not produce cells in 2022 are France+, Italy+, and Mexico. When measuring distances between locations in the SMM, we average the latitude and longitude of the plants within each geographic unit.

[29] Appendix C shows how we calibrate each ξ_{mn} and aggregate.

distribution:[30]

$$u(\boldsymbol{\ell}_{mn}) \mid_{\boldsymbol{\ell}_{mn}=\boldsymbol{\ell}^{\circ}_{mn}} \sim -\text{Gumbel}\left(-\kappa \ln \mathbb{P}(\boldsymbol{\ell}^{\circ}_{mn}), \kappa\right). \tag{25}$$

8.3 Industry Equilibrium Constraints

A simulation j is characterized by a set of lognormal draws for the fixed costs $\phi_{fg_k\ell_k,j}$ and Gumbel draws for the cost differences $\hat{c}_{mn\boldsymbol{\ell},j}$. Given these draws, the firm's realized variable profit for any potential path follows from (21), where $\pi_{mn\boldsymbol{\ell},j}$ depends on the simulation-specific $\hat{c}_{mn\boldsymbol{\ell},j}/\hat{P}_{n,j}$. We assume that firms do not observe the cost draws of their competitors, and therefore face uncertainty regarding the realized price index $\hat{P}_{n,j}$. Firms thus optimize based on their anticipated profits:

$$\mathbb{E}\left[\pi_{mn\boldsymbol{\ell},j}\right] = \frac{1}{\eta} s^{\circ}_{mn} R^{\circ}_{n} \left(\hat{c}_{mn\boldsymbol{\ell},j}\right)^{1-\eta} \hat{\mathbb{A}}_{n}, \tag{26}$$

where $\hat{\mathbb{A}}_{n} \equiv \mathbb{E}\left[\left(\hat{P}_{n,j}\right)^{\eta-1}\right]$ captures the anticipated market demand adjustment resulting from uncertainty over the realized price index change $\hat{P}_{n,j}$. The term $\hat{\mathbb{A}}_{n}$ is an endogenous variable that links the individual firm optimizations together across markets. The estimation sets the $\hat{\mathbb{A}}_{n}$ equal to their sample moments across simulations j using the expression for $\hat{P}_{n,j}$ from (22):

$$\hat{\mathbb{A}}_{n} = \frac{1}{J}\sum_{j=1}^{J}\left(\hat{P}_{n,j}\right)^{\eta-1} = \frac{1}{J}\sum_{j=1}^{J}\left(s^{\circ,\text{ICE}}_{n} + \sum_{m\in\mathcal{M}^{\circ}_{n}} s^{\circ}_{mn}\hat{c}^{1-\eta}_{mn,j}\right)^{-1}. \tag{27}$$

Similar incomplete information approaches to market equilibrium have been used by Seim (2006), Sabal (2025), and Dingel and Tintelnot (2026). Whereas the first two papers model the industry as an oligopoly, we retain the monopolistic competition in constructing the anticipated demand adjustment $\hat{\mathbb{A}}_{n}$. In this respect, our approach resembles the "continuum-case rational expectations" beliefs in Dingel and Tintelnot (2026). Since equation (27) imposes an equilibrium constraint across simulations, the incomplete information approach offers the additional benefit of allowing us to exploit the tools of mathematical program with equilibrium constraints developed by Su and Judd (2012).[31]

8.4 Moments and Estimation

We simulate $J = 100$ times. Given a guess of parameters $\{\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}\}$, for each set of draws and market demand adjustments $\hat{\mathbb{A}}_{n}$, firms solve their facility location problem, formulated as the mixed-integer linear program in section 5.2. The SMM targets 47 moments, matching simulated

[30] For potential paths that coincide with the observed path $\boldsymbol{\ell}^{\circ}_{mn}$, the numerator of $\hat{c}$ uses the actual u draw as the denominator. That is, the draw from the conditional distribution. Hence, $\hat{c}_{mn\boldsymbol{\ell}^{\circ}} = 1$ for the observed paths of all mn combinations.

[31] The complete information counterpart would require each firm to observe the cost draws of all its rivals and simultaneously solve all the location decisions and corresponding market demand vector. Our approach avoids cycling in the interconnected location decisions and reduces the computational burden.

means to their observed counterparts. These moments are organized into five blocks, with predictions at each location aggregated to the same three continents $\mathcal{N} \in \{\text{Am}, \text{As}, \text{Eu}\}$ corresponding to the cost parameters.

The first block matches dyadic flows across continents by targeting expenditure shares between and within continents at each stage.[32] The second block targets the number of car models sold in each continent expressed as a ratio relative to the data. The third block targets the number of production lines at each continent and stage, as a ratio relative to the number of production lines in the data. A production line is an activated facility $y_{fg_k\ell_k} = 1$ for group g_k (a platform for assembly, and battery category for cells). The simulated number of production lines sums the $y_{fg_k\ell_k}$ for each stage and continent $\mathcal{N}$. The fourth block captures how those production lines vary with respect to the location of the firm's headquarters. It targets the share of production lines in a continent based on the continental headquarters of its owner (for both cells and assembly). Lastly, we target the standard deviations of realized log fixed costs from the news articles compiled in table 2, namely 1.50 for cell plants and 1.62 for assembly plants.

Moments in the first two blocks use bilateral flows and model entry contribute to the identification of the relative variable cost differences across continents and the level of variable costs relative to fixed costs. The remaining blocks help to identify the mean fixed costs by continent, how they vary with distance to firm headquarters and with quality, and their variance.

Formally, the vector of moments function, $v(\cdot)$, specifies the differences between the observed outcomes and those predicted by the model. At the true parameter value, $\mathbb{E}\big[v\left(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}; \hat{\mathbb{A}}\right)\big] = 0$ and equation (27) holds. We estimate both market demand adjustments $\hat{\mathbb{A}}$ and structural parameters $\{\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}\}$ such that the SMM objective function is minimized subject to the market demand equations as constraints:

$$\min_{\mathbf{FE},\boldsymbol{\rho},\boldsymbol{\sigma},\hat{\mathbb{A}}} \check{v}\left(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}; \hat{\mathbb{A}}\right)' W \check{v}\left(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}; \hat{\mathbb{A}}\right), \quad \text{subject to} \quad \hat{\mathbb{A}} = f(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}; \hat{\mathbb{A}}), \tag{28}$$

where $\check{v}(\cdot)$ is the simulated estimate of the moment function, $f(\cdot)$ is the right-hand side of equation (27), and W is the SMM weighting matrix. We use an identity matrix W, with the weights on monadic moments scaled up by a factor of three (the number of continents) so that monadic and dyadic moment blocks receive equal weight.

One difficulty in estimating this type of problem is that there are frequently multiple local optima. To search for a global optimum, we implement a multi-start global optimization algorithm, TikTak, explained and benchmarked by Arnoud et al. (2019). The algorithm starts with a uniform exploration of the parameter space and uses the information it accumulates to increasingly focus the search on the most promising region.[33]

[32] For cells, the flows from continent o to d are evaluated using the vehicle values assembled in d with cells sourced from o.

[33] We implement TikTak with 1000 random starts. The algorithm retains the 25 best starts and augments this set with a "prepend" parameter vector (based on priors or a vector of previous estimates). We then use a local optimizer that imposes the endogenous market demand constraint, along with a tightened convergence criteria.

Table 5: SMM parameter estimates

Param.	$\mathcal{N}$	Description	Estimate	90% C.I.
FE^1	As	Variable cost advantage of cell plant	4.91	[2.14, 5.78]
FE^1	Eu	(by region, Am = 0)	5.94	[4.08, 7.97]
FE^2	As	Variable cost advantage of assembly plant	-0.21	[-1.36, 0.82]
FE^2	Eu	(by region, Am = 0)	-0.25	[-0.89, 0.97]
ρ^1	Am	Fixed cost of cell plant (by region)	0.30	[0.22, 0.90]
ρ^1	As		0.12	[0.04, 0.21]
ρ^1	Eu		0.54	[0.25, 1.07]
ρ^2	Am	Fixed cost of assembly plant (by region)	0.16	[0.11, 0.64]
ρ^2	As		0.03	[0.03, 0.25]
ρ^2	Eu		0.18	[0.16, 1.38]
$\rho^{\text{HQ, dist,1}}$		Fixed cost HQ distance elas. (cells)	0.47	[0.40, 0.83]
$\rho^{\text{HQ, dist,2}}$		Fixed cost HQ distance elas. (assembly)	0.78	[0.64, 1.75]
$\rho^{\xi_{fg_1}}$		Fixed cost quality elas. (battery type)	3.40	[2.47, 3.96]
$\rho^{\xi_{fg_2}}$		Fixed cost quality elas. (platform)	2.76	[2.31, 4.06]
σ_1		Fixed cost dispersion (cells)	2.12	[1.54, 2.74]
σ_2		Fixed cost dispersion (assembly)	1.99	[1.29, 2.33]

Confidence intervals given by the second lowest and second highest of 40 bootstrap samples drawn with replacement. Sampling within headquarters continents ensures that the bootstrap sample moments include activity on each continent.

Table 5 shows the results of the estimation. The $\text{FE}^k_{\mathcal{N}}$ estimates determine differences in log wages (variable costs) between $\mathcal{N}$ and North America. The $\boldsymbol{\rho}$ parameters determine the mean of the log fixed cost draws. Fixed costs are systematically higher for cell and EV factories that are farther from the carmaker's headquarters. Quality (the model-level demand shifter) raises fixed costs more than proportionally, with elasticities of 3.40 for cells and 2.76 for vehicles. The estimated $\sigma_1 = 2.12$ and $\sigma_2 = 1.99$ of the fixed cost distribution imply standard deviations for realized log fixed costs of 1.95 for cell plants and 2.03 for assembly plants, which are slightly larger than the targeted values of 1.50 and 1.62, respectively.

Variable costs relative to America are given by $\exp(-\kappa \text{FE}^k_{\mathcal{N}})$. Relative mean fixed costs in stage k between two continents are given by the ratio of their respective $\rho^k_{\mathcal{N}}$ estimates. [34] Since the variable and fixed cost estimates can point in different directions, a calculation of relative average costs provides a useful composite measure of production-cost competitiveness. The ratio of average costs is a weighted average of the ratios of fixed and variable costs with the weights given by the respective shares of total costs, so that we can back out relative average costs with the formula

$$\frac{\text{AC}^k_{\mathcal{N}}}{\text{AC}^k_{\text{Am}}} = \frac{\text{FC}}{\text{TC}} \times \frac{\rho^k_{\mathcal{N}}}{\rho^k_{\text{Am}}} + \left(1 - \frac{\text{FC}}{\text{TC}}\right) \times \exp(-\kappa \text{FE}^k_{\mathcal{N}}).$$

[34] See equations (15) and (24) respectively.

Our estimates imply the ratio of fixed costs to world sales is 7% in the EV sector.[35] Under constant markup, the ratio of variable costs to revenues is $(\eta - 1)/\eta = 0.75$, which in turn implies FC/TC = 0.085.

Table 6 uses these relationships to re-express the variable and fixed cost parameters in a way that is much easier to interpret than the raw SMM estimates shown in the first 10 rows of table 5. Columns (3) and (6) compute average production costs for cells and assembly in Europe and Asia, both compared to the North America.

Table 6: Variable, fixed and average cost comparisons

Region	Cells			Vehicles		
Ratio:	var.	fixed	avg.	var.	fixed	avg.
Europe	-50.06	83.30	-38.67	2.93	12.54	3.75
Asia	-43.64	-57.80	-44.85	2.47	-78.79	-4.47

% difference in variable, fixed and average costs with respect to North America.

We estimate fixed costs to be highest in Europe and lowest in Asia. This is true for both cells and vehicles, but Europe's disadvantage in cells is much larger. North America has a substantial cost disadvantage in average costs of cells relative to the other two continents. The gaps are much smaller for vehicles, and North America has a slight average cost advantage over Europe.

The confidence intervals shown in table 5 are obtained via bootstrap sampling of firms, as in Antràs et al. (2017) and Bernard et al. (2022). We hold fixed the number of firms drawn by headquarters continent: six for Asia, four for North America, and five for Europe. For each bootstrap sample, we construct the empirical moments used in the structural estimation, simulate 100 sets of cost shocks, and *globally* estimate parameters of the model for that sample.[36]

Sampling firms ensures that our key interdependencies, including dependencies across locations, models, and paths, are preserved within a firm. Across firms, the dependencies are channeled through the endogenous price index. However, we estimate $\hat{\mathbb{A}}$ close to one across all countries in the baseline, which implies that price indices are nearly fixed and dependencies across firms are weak. We further verify that the market demand adjustments remain close to one in bootstrap estimates.

The confidence intervals reported in table 5 are based on the lower and upper 5% quantiles of the bootstrap parameters. The wide range exhibited for some parameters arises in part because the procedure samples over a small number (15) of highly asymmetric firms. Thus, a particular

[35]Predicted total realized fixed costs equal 18.85 billion and world (anticipated) EV sales are 267.36 billion, and therefore $18.85/267.36 \approx 0.07$.

[36]Depending on which firms are drawn in the bootstrap sample, EV revenue in each market, R_n^{EV}, is computed accordingly. Keeping ICE vehicle revenue in each market the same as in the full sample and summing it with bootstrapped EV revenue, we obtain the total vehicle revenue, R_n° equivalent, in each bootstrap sample. Market share in the bootstrap sample, s_{mn}° equivalent, is calculated by dividing the EV revenue of the drawn firms to the bootstrapped total vehicle revenue.

bootstrap sample might replace a giant firm like Tesla with several instances of the much smaller Rivian. Since the variation in the bootstrap moments is so large, we view these as conservative confidence intervals. What matters in the end is whether the imprecision in the estimates generates volatile counterfactual policy predictions. In the next section we show prediction ranges for counterfactuals. The reassuring finding is that policy effects are very stable across bootstrap parameter sets.

8.5 Targeted Fit

The plausibility of our policy counterfactuals depends to a large extent on whether the estimated model can replicate the key features of intercontinental battery and EV production and trade. We therefore examine the fit of the targeted moments. Figure 8 shows the count of production lines on each continent in the data (horizontal axis) and in our estimated model (vertical axis). It captures whether the model accurately predicts the $y_{fg_k\ell_k}$ in the firm's problem, which in turn depends on correctly estimating fixed costs of activating plants. All six points fit the 45-degree line closely, with Asia's dominance of both stages clearly replicated by the model. Table 7 shows the fit of model entry, z_{mn} in the optimization. Because of the "soft z" procedure, the baseline prediction must lie below the data. However, predicted entry in all three continents is close to actual entry. Furthermore, the EV models that the simulation predicts to be profitable account for over 92% of actual revenues.

Figure 8: Production lines by continent

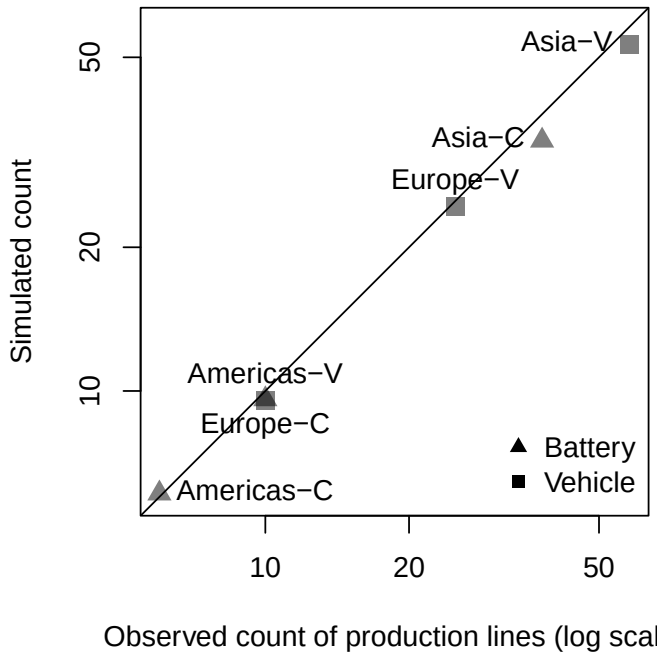


Table 7: Model market entry

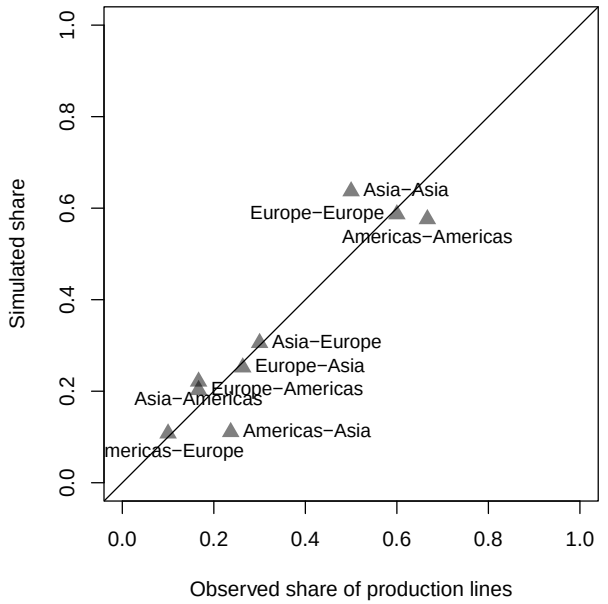
	# car models		Rev.
	data	baseline	(%)
Americas	49	43.0	92.1
Asia	111	99.9	92.5
Europe	83	77.2	94.2

Note: Baseline model predictions are averaged across all simulations. Column (1) is the targeted data moment. Column (3) is the % of observed sales by car models we keep in the prediction, averaged across simulations.

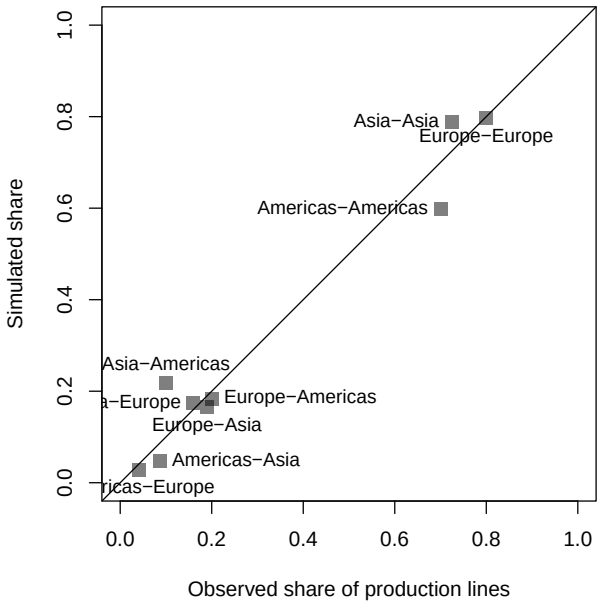
Figure 9: Fit of the targeted dyadic moments

Production lines by firm origin–production location

(a) battery cells

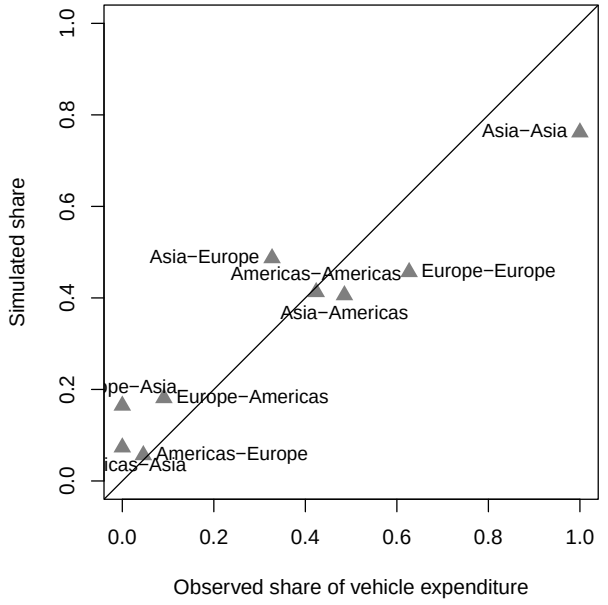


(b) vehicle assembly

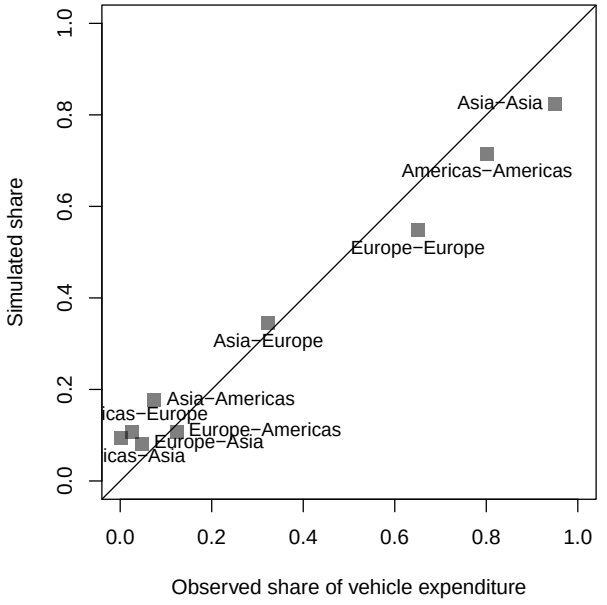


Expenditure shares by production location–destination

(c) battery cells



(d) vehicle assembly



Note: The SMM targets these moments.

Figure 9 displays the continent-to-continent fits for two more important moments. Panels (a) and (b) show the number of production lines by headquarters-production dyads. For example, in panel (a), European vehicle makers account for about 20% of the cell production lines in North America, which is only slightly above the data. The fit for vehicles is also strong, with the simulation replicating pronounced intra-continental bias, although the model overpredicts the share of Asian makers in North America. Panels (c) and (d) show the share each production origin accounts for in *expenditure*. For example, the simulations predict that vehicles with Asia-made cells make up about half of all vehicles assembled in Europe, compared with roughly 35% in the data. Panel (d) again shows home bias in all three continents, but in this case it is based on location of assembly, rather than location of headquarters.[37]

9 Counterfactual Policy Simulations

In recent years governments in North America, Europe and Asia have put in place (and sometimes removed) policies to stimulate purchases of EVs such as consumer subsidies. Over time, contingencies have been added to those subsidies. We will focus on widely used requirements regarding the location of production along the supply chain, in our case, the location of assembly and cell production.

More specifically, we model the following scenarios:

1. **Subsidy to EV buyers:** The government subsidizes EV purchases regardless of the locations of assembly and battery manufacture.

2. **Subsidy conditional on domestic assembly:** Only cars assembled in the same continent as the buyer are eligible for the subsidy.

3. **Subsidy conditional on domestic cell-vehicle supply chain:** To qualify, both cell and vehicle must be produced in the same continent as the buyer.

Policy 1 resembles the common practice in Europe and Canada (and the US before the 2022 Inflation Reduction Act) where subsidies were not contingent on production locations. Policy 2 adds the "buy domestic" requirement at the assembly level. We impose this at the continental level, which is similar to the IRA requirement that included Canada and Mexico. Policy 3 mirrors the IRA policies regulating the input sources of EVs to be eligible for the subsidy. The IRA phased in mineral and components requirements for obtaining a full $7,500 subsidy. Had the policy not been eliminated in 2025, 100% of components would be required from the US or Free Trade Agreement (FTA) by 2029. Allcott et al. (2024) evaluate the effects of the IRA, taking into account its detailed requirements, as well as the exemption for leased cars. Their framework takes production locations as given, whereas the simulations here focus on how policies induce changes in facility locations.

[37] Appendix E shows the fit of the estimated model with respect to several *untargeted* moments.

Upstream requirements for subsidies are not unique to the IRA. For example, France introduced in 2024 an "Eco score" cutoff for the CO2 emissions of the overall production process in order for an EV to be eligible for the demand subsidy. Without specifically naming China, the policy *de facto* excludes cars made with Chinese batteries.[38] In 2025 the French government added a €1,000 supplement for EVs with European batteries. In our simulation, the rule is that only cells produced inside the continent are eligible for the consumer subsidy.

Production subsidies and tariffs are additional policies that feature prominently in the auto industry. Our framework easily accommodates those policies, but we relegate their analysis combined with consumer subsidies to appendix F due to space considerations. Another salient policy we do not explicitly model is an emission tax on ICEs. This is because our unconditional 20% subsidy for EVs (Policy 1) is equivalent to a uniform 25% tax on ICEs in terms of all the market outcomes we describe.[39]

9.1 Policy Simulation Results

Up to now, we have used $\hat{c}_{mn,j}$ to denote the full change in model m's delivered marginal cost to n, inclusive of existing tariffs and subsidies that our counterfactual simulations hold constant across scenarios. We now introduce $\hat{t}_{mn,j}$ to denote the *additional* subsidy enacted in our counterfactual scenarios, which we compare to a *baseline* case where $\hat{t}_{mn,j} = 1, \ \forall m, n, j$. Thus, while $\hat{c}_{mn,j}$ denotes the cost change in the baseline relative to data, $\hat{c}_{mn,j}\hat{t}_{mn,j}$ gives the cost change under a counterfactual scenario.

In order to make the scenarios comparable, we set a common counterfactual subsidy rate of 20% at the continental level: $\hat{t}_{mn,j} = .8, \ \forall n \in \mathcal{N}$—whenever model m is eligible for the policy. Since eligibility may also depend on the location of production along the value chain (except in Policy 1), $\hat{t}_{mn,j}$ varies across simulations j for a given model m and destination n based on the chosen path $\boldsymbol{\ell}_{mn}$.

For each counterfactual, we report three sets of outcomes, aggregating from individual countries n to their respective continents $\mathcal{N}$, and averaging across simulations $j \in J$:

1. Changes in the number of production lines
2. Changes in EV expenditure
3. Decomposition of expenditure changes into:
   (a) the direct effects of policy,
   (b) the changes in delivered marginal costs due to path changes,
   (c) changes in model entry in terms of the number of varieties offered in each market,
   (d) the associated change in market demand.

[38] Malgouyres et al. (2025) show that failure to attain the Eco-score minimum entails a 60% reduction in sales.

[39] See section A.1 in the appendix for a derivation of this equivalence. The equivalence does not extend to the fiscal consequences for the government's budget balance.

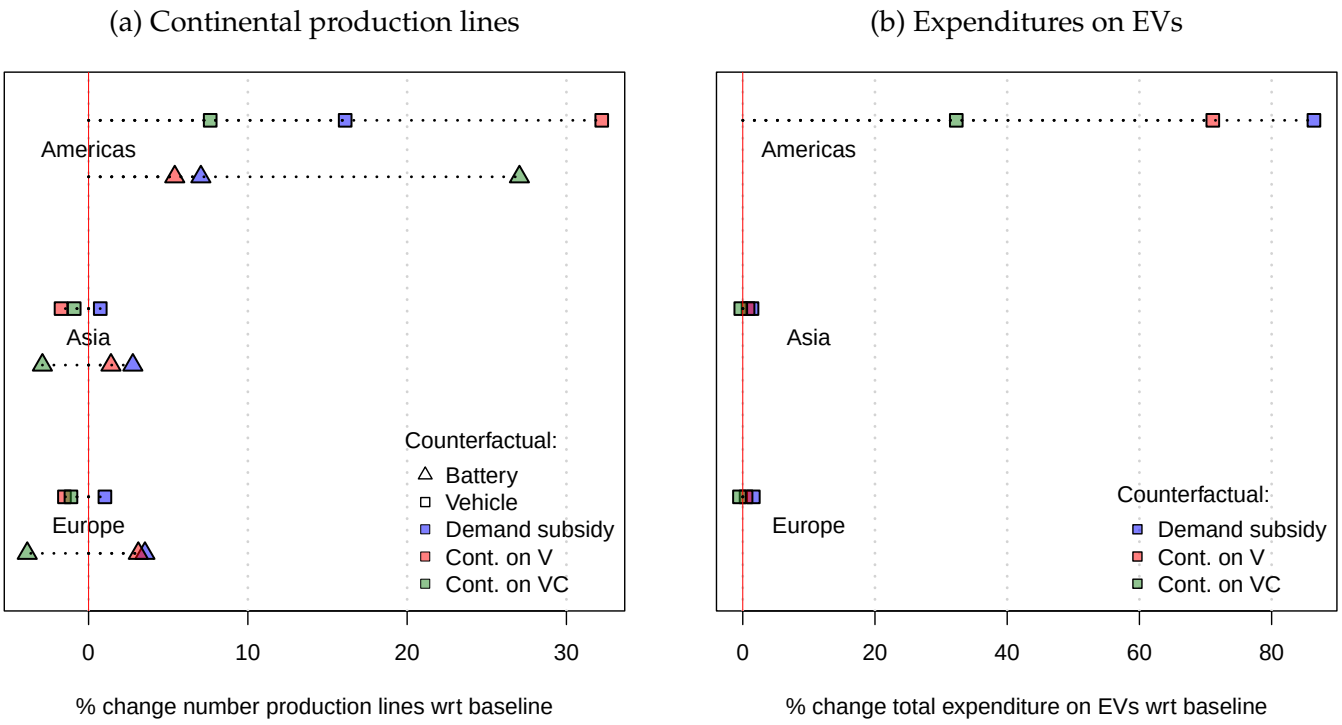
We report all of those outcomes for each scenario relative to the baseline where $\hat{t}_{mn,j} = 1$. Production lines capture production and employment, while EV expenditure proxies for emissions reductions. Since we hold total expenditure on vehicles fixed, the change in EV expenditure corresponds to the reduction in ICE expenditure.[40] Throughout the counterfactual exercises, we re-solve for the new equilibrium price index via fixed-point iteration such that the market equilibrium condition specified in (27) holds after the policy change.

The change in EV expenditure is computed as follows. In the data, the total EV expenditure in market n is $R_n^{\circ,\text{EV}} \equiv R_n \sum_{m \in \mathcal{M}_n^\circ} s_{mn}^\circ$. Let $R_{\mathcal{N}}^{\circ,\text{EV}} \equiv \sum_{n \in \mathcal{N}} R_n^{\circ,\text{EV}}$ denote the EV expenditures at the continent level. In each simulation j, an EV model's expenditure change (relative to the expenditure in the data) is $\hat{c}_{mn,j}^{1-\eta} \hat{t}_{mn,j}^{1-\eta} \hat{P}_{n,j}^{\eta-1}$. The continental change in EV expenditure is an average of those changes over continents (with weight $R_n^{\circ,\text{EV}} / R_{\mathcal{N}}^{\circ,\text{EV}}$), over simulations (with weight $1/J$), and over models (with weight s_{mn}°):

$$\bar{\hat{R}}_{\mathcal{N}}^{\text{EV}} = \sum_{n \in \mathcal{N}} \frac{R_n^{\circ,\text{EV}}}{R_{\mathcal{N}}^{\circ,\text{EV}}} \sum_j \frac{1}{J} \sum_{m \in \mathcal{M}_{n,j}} s_{mn}^\circ (\hat{c}_{mn,j}^{1-\eta} \hat{t}_{mn,j}^{1-\eta} \hat{P}_{n,j}^{\eta-1}) \tag{29}$$

9.1.1 North America

Figure 10: Counterfactual results with 20% subsidy in North America



[40] We do not model explicitly how CO2 emissions from internal combustion engines change in response to policies. However, figure ES.1 in Bieker (2021) shows that battery EVs emit 60–69% less CO2 over their lifetime than gas cars in the US and Europe.

Figure 10(a) provides the continental summary of our plant activation indicators y in response to a 20% subsidy for EV purchases in North America (Policies 1–3). Squares correspond to the change in assembly production lines and triangles to cell production lines for each continent. These changes are measured relative to the baseline prediction of the model, which as shown in figure 8(a) closely matches the data. The unconditional subsidy (Policy 1 in blue) raises the number of production lines in the Americas by 16% for assembly and 7% for cells. This indicates a significant impact of the subsidy on production capacity, highlighting the effectiveness of financial incentives in stimulating EV manufacturing. The complementarity between assembly and cell production in the subsidizing region also appears clearly in these results.

Figure 10(b) provides the percent change in expenditures on EVs implied by the policy changes (recalling that the total expenditure on cars is constant, this implies that the policies displace ICE sales). The unconditional subsidy increases expenditure on EVs by 86%. This increase is so large that the rise in the number of vehicles assembled in the Americas does not come at the expense of Europe or Asia. The number of assembly lines on those continents rises by 1% and 0.7%, respectively, and the collateral benefit for cells is even slightly larger for both regions. Although the subsidy is ostensibly neutral in its application, not being contingent on production locations, it nonetheless reallocates production shares in favor of the subsidizing region. In a constant returns model without fixed costs, production in all three continents would increase by the same proportion.

Imposing domestic final assembly as a condition for the subsidy (Policy 2 in red: "Buy American" generalized to include Canada and Mexico) raises production lines by nearly a third compared to the baseline. This hurts Asia by about 2% of its assembly lines relative to the baseline. European assembly is similarly adversely affected by the buy-American subsidy. Online appendix O.B displays dyadic results in figure O.B3 and table O.B4, in terms of the count of car models sourced from each continent. They show that the assembly-contingent subsidy promotes cell exports from Asia to North America (for 5 extra models), while lowering vehicle exports (4 fewer models).

While successful in terms of domestic protection, figure 10(b) shows that Policy 2 lowers expenditures on EVs and thus is not appealing as an environmental policy relative to the unconditional subsidy in Policy 1.[41] The policy is also unsuccessful at promoting cell production in America. Hence, we turn to the natural alternative where the subsidy is contingent on *both* local assembly and cell production.

Policy 3 does indeed increase cell production lines by 27% (four times more than Policy 1 and five times more than Policy 2). However, since North America is the costliest place to produce cells, the relocation raises costs and hurts vehicle assembly, whose increase is less than a quarter of that under Policy 2. Importantly, the weaker outcome on production is accompanied by a sharp reduction in the EV expenditure relative to the other two policies. Unless the objective is strongly biased toward local battery production, Policy 3 is dominated by the alternative subsidy

[41] Tables F6 and F7 in appendix F present the numbers underlying the visualizations in figures 10 and 11.

schemes.[42] Policy 3 has the added drawback of alienating trade partners: vehicle and cell production in the other two continents unambiguously suffer from this extreme local content restriction on the North American demand subsidy.

Table 8: Variation in counterfactual production lines (North American policy)

Policy	EV prod. lines (#)				Cell prod. lines (#)				EV expenditure (%)			
	Sim		Boot		Sim		Boot		Sim		Boot	
	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3
Americas												
1: Unconditional	1	2	1.2	1.6	0	1	0.3	0.4	71.4	98.3	85.9	91.8
2: Continental V	2	4	2.3	3.3	0	1	0.3	0.5	56.5	83.3	69.8	75.7
3: Continental V+C	0	1	0.6	0.9	1	2	1.2	1.7	19.8	46.0	34.8	40.8
Asia												
1: Unconditional	0	1	0.4	0.5	0	2	0.9	1.1	0.0	1.9	1.2	1.7
2: Continental V	-1	0	-0.8	-0.5	0	1	0.4	0.6	-0.1	1.2	0.6	1.3
3: Continental V+C	-1	0	-0.4	-0.2	-2	0	-1.0	-0.8	-0.5	0.1	-1.1	-0.4
Europe												
1: Unconditional	0	1	0.4	0.5	0	1	0.4	0.5	0.3	2.7	1.6	2.0
2: Continental V	-1	0	-0.4	-0.2	0	1	0.2	0.3	-0.5	1.4	0.2	0.6
3: Continental V+C	0	0	-0.2	-0.1	-1	0	-0.4	-0.3	-1.2	0.1	-0.7	-0.3

Sim: Figures obtained from the variation over simulation draws of counterfactual outcomes with table 5 parameter estimates. This is always an integer for production lines.
Boot: Figures obtained from the variation over bootstrapped estimates for *mean* outcomes over simulations.

All results in our main counterfactual figure 10 depend on the point estimates from the SMM. Since these estimates differ from true values, we would like to assess how policy outcomes vary across parameter values drawn from the confidence interval. Moreover, counterfactual changes shown in this figure are constructed from a set of 100 simulations, and are thus subject to simulation uncertainty. Table 8 investigates these two sources of variation in the three sets of predicted outcomes (EV and cell production lines and EV expenditure), reporting the first and third quartiles over simulations and bootstrap replications.

The most salient feature of table 8 is that the bootstrap interquartile ranges are almost always substantially narrower than the those across simulations, indicating that simulation variation matters more than parameter uncertainty. The narrow bootstrap variation implies that the policy rankings described above remain robust. Policy 1 is most effective in expanding EV purchases, Policy 2 expands EV production lines the most and Policy 3 is best at expanding cell production lines. Cross-continent effects are smaller than within-continent, but always positive for the unconditional subsidy. The V+C contingency in North America is systematically bad for production and EV adoption in Europe and Asia.

[42] In appendix table F6 we see that augmenting policy 3 with a cell production subsidy would add vehicle production lines relative to policy 3 alone. Policy 3+4 maximizes the number of cell production lines in North America and may be justified on non-environmental grounds if this outcome is sufficiently important.

The three rows of table 9 unpack the components of the change in North American EV expenditure caused by the North American subsidies.[43] The last column reproduces the expenditure change graphically shown in the top line of figure 10(b).[44] The first column reports the share of paths (mn, j combinations) eligible for the subsidy. The second column reports the corresponding share of revenue. Revenue shares are larger, indicating that firms are more likely to adjust the paths of models with higher market shares to attain eligibility. The next two columns report an average of the policy ($\hat{t}_{mn,j}$) and delivered cost ($\hat{c}_{mn,j}$) changes across continents, simulations, and models, aggregated using the CES cost index (denoted with a ˜). For the cost changes $\hat{c}_{mn,j}$, it takes the form:

$$\tilde{\hat{c}}_{\mathcal{N}}^{1-\eta} = \sum_{n\in\mathcal{N}} \frac{R_n^{\circ,\mathrm{EV}}}{R_{\mathcal{N}}^{\circ,\mathrm{EV}}} \sum_j \frac{1}{J} \sum_{m\in\mathcal{M}_{n,j}} \frac{s_{mn}^{\circ}}{\bar{s}_n^{\mathrm{EV}}} \hat{c}_{mn,j}^{1-\eta}, \tag{30}$$

where $\bar{s}_n^{\mathrm{EV}} \equiv (1/J)\sum_j \sum_{m\in\mathcal{M}_{n,j}} s_{mn}^{\circ}$ is the average share of EVs across simulations for market n, accounting for models in $\mathcal{M}_n^{\circ}$ that drop out of $\mathcal{M}_{n,j}$ (the soft z for that model in simulation j is set to zero). Thus, all the weights in the CES index—across countries n, simulations j, and models m—sum to 1.

Table 9: Cost and expenditure changes in North America

Policy	Eligible share		Cost index change		Shifters		EV Exp.
	path	revenue	subsidy	costs	variety	demand	change
			$\tilde{\hat{t}}_{\mathcal{N}}$	$\tilde{\hat{c}}_{\mathcal{N}}$	$\bar{s}_{\mathcal{N}}^{\mathrm{EV}}$	$\hat{\mathbb{A}}_{\mathcal{N}}$	$\bar{\hat{R}}_{\mathcal{N}}^{\mathrm{EV}}$
1: Unconditional	100.0	100.0	-20.0	-4.5	2.5	-16.5	86.4
2: Continental V	43.4	90.5	-17.4	-3.4	1.9	-13.6	71.1
3: Continental V+C	21.5	68.3	-14.7	2.2	0.8	-6.2	32.3

Policies defined at the beginning of section 9. Policies are applied in the Americas, and outcomes are for North America. For the variable in each column, defined in text, the number reported is the percentage difference of that variable between the policy scenario and baseline.

The CES cost index has the usual interpretation that it induces the same change in aggregate expenditures as the underlying set of individual cost changes averaged in the index, given our monopolistic assumption with a constant markup. The term $\tilde{\hat{t}}_{\mathcal{N}}^{1-\eta}$ is defined analogously as the CES index of the policy cost changes $\hat{t}_{mn,j}^{1-\eta}$, using the same weights. Thus, for the unconditional subsidy (Policy 1), $\tilde{\hat{t}}_{\mathcal{N}}$ is exactly 20% below baseline, while the reduced take-up shown in the first two columns lowers that number in the subsequent rows.

The next two columns aggregate the demand shifters for EV variety changes and market demand—$\bar{s}_{\mathcal{N}}^{\mathrm{EV}} \equiv \sum_{n\in\mathcal{N}}(R_n^{\circ,\mathrm{EV}}/R_{\mathcal{N}}^{\circ,\mathrm{EV}})\bar{s}_n^{\mathrm{EV}}$ and $\hat{\mathbb{A}}_{\mathcal{N}} \equiv \sum_{n\in\mathcal{N}}(R_n^{\circ}/R_{\mathcal{N}}^{\circ})\hat{\mathbb{A}}_n$—from markets n to continent $\mathcal{N}$. These are weighted averages of the market-level changes. For the EV varieties, the

[43] Section A.3 of the appendix derives a full decomposition of the average EV expenditure change $\bar{\hat{R}}_{\mathcal{N}}^{\mathrm{EV}}$ into the components $\tilde{\hat{t}}_{\mathcal{N}}, \tilde{\hat{c}}_{\mathcal{N}}, \bar{s}_{\mathcal{N}}^{\mathrm{EV}}, \hat{\mathbb{A}}_{\mathcal{N}}$. That decomposition also includes covariance terms between those components.

[44] Table O.B5 in the online appendix reports impacts of the North American policy for the other two continents.

weights are the shares of EV expenditures, whereas the weights for the overall market demand change are the shares of all light vehicle expenditures.

The EV variety shifter $\bar{s}_{\mathcal{N}}^{\text{EV}}$ captures the change in model entry between the baseline and a counterfactual scenario. It is positive for the subsidy scenarios, indicating that additional models are offered relative to the baseline. The demand shifter $\hat{\mathbb{A}}_{\mathcal{N}}$ captures the endogenous change in market demand induced by a policy. It averages the CES price index changes $\hat{P}_{n,j}$ across markets and simulations (see equation 27). This shifter represents the average expenditure change for a model at a constant price. Given constant markups, this applies to any model whose cost does not change ($\hat{c}_{mn,j} = 1$), including all ICE models. A negative market demand shifter thus reflects an increase in competition.

The unconditional subsidy (Policy 1) allows all paths leading to sales in North America to obtain the 20% subsidy, which results in a 2.5% increase in $\bar{s}_{\mathcal{N}}^{\text{EV}}$ compared to the baseline, an increased variety effect. Firms also rearrange their value chains such that the delivered marginal cost $\tilde{\hat{c}}_{\mathcal{N}}$ falls by 4.5% in addition to the 20% subsidy. Partly compensating these cost reductions is an increase in competition equivalent to a 16.5% expenditure loss at constant prices relative to the baseline. Overall, the net effect on EV expenditure is an impressive 86% increase.

Policy 2 imposes an additional constraint that drastically lowers the share of paths receiving the subsidy. However, the revenue share of eligible paths remains very high, emphasizing that the models associated with larger sales tend to self-select into eligible paths. All effects are logically dampened by the local assembly constraint, and the objective of favoring local assembly harms the gain in EV adoption by 15 percentage points (86−71).

Policy 3 is even worse from an environmental perspective, as few EVs sold in North America end up qualifying for the subsidy when it is contingent on local cell sourcing. Only 21.5% of the realized paths are eligible. The small number that do draw sufficiently low fixed costs in North America to source cells within the region are still the best-selling vehicles since they represent 68% of sales. On top of the attenuated subsidy, variety, and market demand effects relative to Policies 1 and 2, contingency on the entire supply chain raises average delivered marginal cost by 2.2% due to path changes. Policy 3 is worse than non-intervention from the point of EV production in both Asia and Europe (as shown in figure 10(a)), while increasing EV production in North America *less* than the other two policies. This is an example where the policy is successful at attracting production of the major input (cells), but at the cost of reduced production of vehicles that use those local cells. The gain in EV adoption falls by a further 39 percentage points (71−32), leaving few arguments in favor of this policy.

9.1.2 Europe

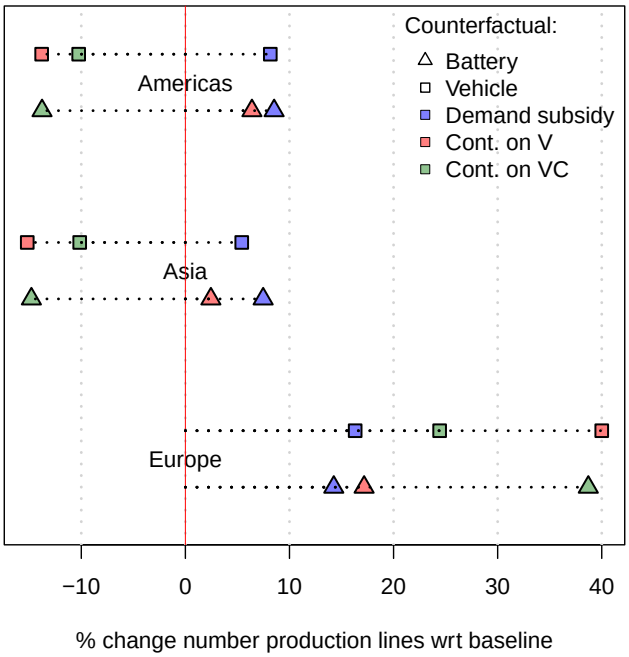
Figure 11 provides our main policy simulation results for Europe. While there are many shared aspects with North America, the predicted results for Europe differ in important dimensions. Panel (a) shows larger changes in production lines. The biggest impact is that Policy 2 generates a 40% increase in vehicle production lines in Europe (10 extra lines above the baseline level). The

majority of the added lines comes at the expense of EV facilities in the other two regions (around −14% in each). In contrast to the North American version of this policy, there is little stimulus to cell production lines in Asia. On the expenditure side, the European version of Policy 1 is less successful, with EV expenditures rising by 63% as compared to 86% in North America. In online appendix table O.B3, we show that these EU findings are approximately as robust across different sets of parameter estimates as the results for North America shown in table 8.

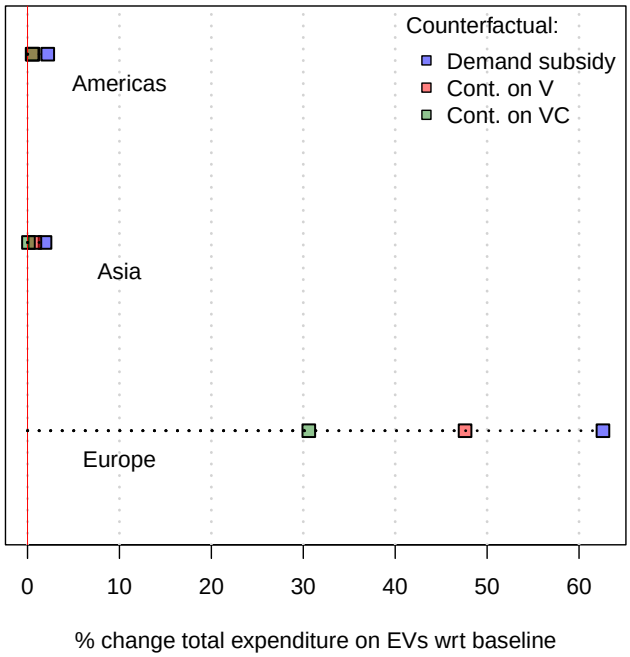
As was the case in North America, the Policy 3 requirement of domestic cells lowers the number of assembly plants: the green square is to the left of the red square. The trade-off is less pronounced in Europe, where the V+C contingency yields more assembly lines than the unconditional subsidy (the blue square is to the left of the green square). We can see the explanation for this difference in the cost comparisons shown in table 6. Europe is much more competitive in cell manufacturing than North America: Europe's average costs are 39% lower than North America and only about 6% higher than Asia.

Figure 11: Counterfactual results with 20% subsidy in Europe

(a) Continental production lines



(b) Expenditures on EVs



As noted earlier, the expenditure boost from unconditional subsidies in Europe is 24 percentage points lower than the increase in North America. Table 10 helps us understand why. First, the reduction in variable costs is lower (3.0% vs 4.5%) and this difference is magnified by the demand elasticity of 4. Second, the variety effect is smaller (1.9% vs 2.5%). Third, the demand contraction is larger (-20.1% vs -16.5%) because EV shares were larger in Europe in 2022. Together, these components account for an 18 percentage point gap in expenditure increases between North America and Europe. The remainder comes from a mixture of covariance and aggregation effects explained

in the complete decomposition shown in appendix A.3.

The other decomposition results shown in table 10 are broadly similar to those we saw for North America. Policy 3 reduces subsidy take-up to 47.3% of the potential paths (much higher than the 21.5% take-up of Policy 3 in North America). This leads to a lower effective subsidy of just 12.1%, resulting in substantial reductions in EV adoption compared to the other two policies. The cost competitiveness of cell manufacturing in Europe mitigates the adverse cost effects of Policy 3 (column 4), which now rise by 0.2% instead of 2.2% in North America.

Table 10: Cost and expenditure changes in Europe

Policy	Eligible share		Cost index change		Shifters		EV Exp.
	path	revenue	subsidy	costs	variety	demand	change
			$\tilde{\hat{t}}_{\mathcal{N}}$	$\tilde{\hat{c}}_{\mathcal{N}}$	$\bar{s}^{\text{EV}}_{\mathcal{N}}$	$\hat{\mathbb{A}}_{\mathcal{N}}$	$\bar{\hat{R}}^{\text{EV}}_{\mathcal{N}}$
1: Unconditional	100.0	100.0	-20.0	-3.0	1.9	-20.1	62.6
2: Continental V	68.4	85.3	-15.8	-1.6	1.3	-15.3	47.6
3: Continental V+C	47.3	66.8	-12.1	0.2	0.6	-9.9	30.6

Policies defined at the beginning of section 9. Policies are applied in Europe, and outcomes are for Europe. For the variable in each column, defined in text, the number reported is the percentage difference in that variable between the policy scenario and baseline.

9.1.3 Production Subsidies and Tariffs

In the main text we have focused on just three versions of the demand subsidy that allow us to consider the core trade-offs. Our framework can easily handle a range of other policies, and we now briefly summarize the simulated impacts of production subsidies and tariffs.[45] As seen in table F7, cell production subsidies and tariffs are much less effective at increasing EV purchases than the consumer subsidies. The former increase EV expenditures by 5.1% and 6.3% in North America and Europe, respectively. Tariffs on vehicles *lower* EV expenditure by about 10% on both continents, and this is worsened when those tariffs apply to cells as well. The Biden IRA policy of combining contingent consumer subsidies with cell production subsidies ("3+4") does indeed do more for EV adoption than Policy 3 in isolation, but it still performs worse in this dimension than Policies 1 and 2 on both continents, albeit with a smaller gap in Europe. From the point of view of opening new battery factories, table F6 shows that cell production subsidies (Policy 4) are more effective than clean EV consumer subsidies (Policy 1). The combined Policy "3+4" maximizes the number of cell factories in whichever continent it is implemented.

Tables O.B5 and O.B6 in the online appendix dig into the details of the consequences of EV and battery tariffs. In both cases, we see rises in marginal costs and shrinking numbers of car models offered in the countries imposing the tariffs. Another notable feature is that despite the fact that the majority of paths will be subject to tariffs, revenue shares of those paths are much

[45] The complete set of policy counterfactuals are reported in appendix F.

lower (column 2). The effective tariff in column 3 resulting from a statutory 20% tariff is under 5% in North America (reflecting widespread tariff jumping by models with high market shares) and under 10% in Europe.

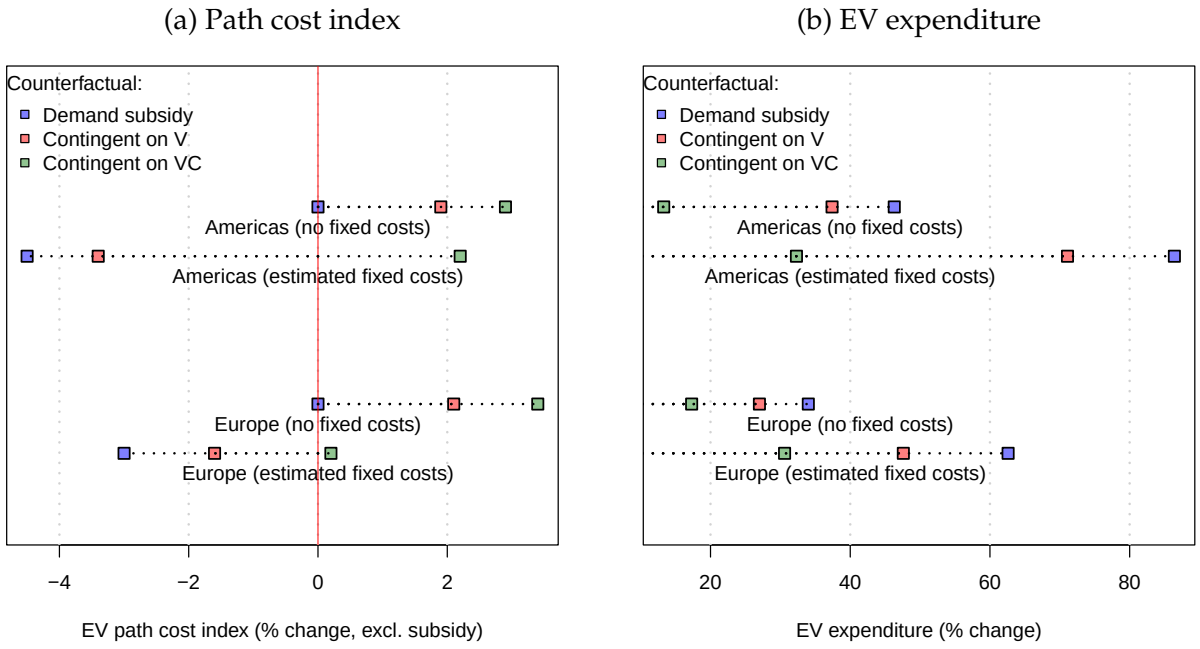
We have not computed welfare effects of the tariffs and subsidies. With large social costs of carbon, welfare will tend to rise or fall depending on changes in demand for the emission-reducing product. In the Gerarden et al. (2025) analysis of tariffs on solar panels, the environmental component of welfare moves in the same direction as consumer surplus (down for tariffs, up for production subsidies), but is about twenty times larger.

9.2 Quantifying the Role of Fixed Costs

We conclude with a different type of computational scenario that highlights the relevance of our framework—capturing the trade-offs between fixed and variable production costs with endogenous facility choice—for the modeling of policy counterfactuals. To highlight those trade-offs, we set the facility fixed costs to zero and then rerun the subsidy policy counterfactuals (Policy 1–3).[46]

We focus on the consequences for the induced change in the delivered marginal cost $\tilde{\hat{c}}_{\mathcal{N}}$ and associated EV expenditure share $\tilde{\hat{R}}^{\text{EV}}_{\mathcal{N}}$ in the continent that implements the policy (for both North America and Europe). Those outcomes are shown in panels (a) and (b) of figure 12, respectively. For comparison, we replicate the results from tables 9 and 10 that we previously reported with our estimated facility fixed costs. Recall that the cost change $\tilde{\hat{c}}_{\mathcal{N}}$ captures the average change in cost $\hat{c}_{mn,j}$ due to changes in the optimal chosen paths, which *do not include the subsidy*.

Figure 12: Comparison with zero fixed costs



[46]For this exercise, we keep the variable cost parameters at their estimated values.

With no facility fixed costs, the MUFLP optimization is trivial: The lowest delivered path cost $\hat{c}_{mn,j}$ is always chosen among the set of paths through all the potential facilities in both stages. There are no interconnections between chosen paths based on the set of open facilities (any facility can be opened at zero cost). Thus, the unconditional subsidy (Policy 1) will have no effect on the chosen paths and the resulting delivered marginal cost: the blue squares for this case are at zero for Policy 1. EV prices will all be reduced by the constant subsidy $\tilde{\hat{t}}_{\mathcal{N}} = .8$. The expenditure shares in panel (b) reflect that constant price change along with the endogenous change in market demand $\hat{\mathbb{A}}_{\mathcal{N}}$ such that expenditures on all cars (including ICEs whose prices do not change) remain constant. This is why the expenditure share increases are more muted for Europe: the share of EVs is higher than in the Americas, and therefore the decrease in market demand $\hat{\mathbb{A}}_{\mathcal{N}}$ is steeper.

In this case with no fixed costs, the contingent subsidies (Policies 2 and 3) must weakly increase the delivered marginal cost for all models. Those policies may be successful in their protectionist goal of inducing a path change from a "Foreign" one in the baseline to a domestic one that is eligible for the subsidy. But such a change will result in higher delivered marginal cost above its minimum attained under Policy 1.[47]

As the restrictions on the subsidy eligibility increase, the average delivered marginal cost will initially increase, but then decrease beyond a threshold. In the limit when the restrictions are so onerous that they do not induce any path changes, the resulting policy will not affect the average delivered marginal cost.[48] Thus, the cost increases associated with the more restrictive Policy 3 could be *lower* than for Policy 2 if it is close enough to that limit. However, we do not find that to be the case empirically: the tougher restrictions on the subsidy in Policy 3 induce *higher* costs than for Policy 2.

As we have previously reported for our framework with fixed costs, all three subsidy policies induce changes in the delivered costs $\tilde{\hat{c}}_{\mathcal{N}}$. Under both the baseline and unconditional subsidy, the lowest delivered cost path is often not chosen because fixed costs constrain which facilities are open. The unconditional subsidy may induce a switch to a path with lower variable cost $\hat{c}_{mn,j}$. The larger market scale driven by the subsidy makes it worthwhile to incur higher fixed facility costs (including opening a new facility on the new path) in exchange for the lower variable cost. However, the subsidy cannot increase the delivered cost to the subsidized region.[49] Thus, in our framework with fixed costs, the unconditional subsidy must decrease—at least weakly—the average delivered marginal cost.

This is not the case for the contingent subsidies: In those scenarios, a firm may be induced to choose a higher path cost $\hat{c}_{mn,j}$ that is eligible for the subsidy, so long as the higher cost is outweighed by the subsidy. Thus, our structure does not restrict the sign of the cost change $\tilde{\hat{c}}_{\mathcal{N}}$ under Policies 2 and 3. Furthermore, the cost savings under those contingent policies could be

[47] If the restricted policy does not induce a change in path, either because the original path is eligible or because the higher cost of the eligible paths outweigh the benefit of the subsidy—then the delivered marginal cost does not change.

[48] This initial increase and subsequent decrease in delivered cost is closely related to the "Laffer Curve" effect for rules of origin studied by Head et al. (2024).

[49] If the new path with the higher cost is optimal with the subsidy, it must be optimal without the subsidy.

even greater than under the unrestricted subsidy. In appendix A.2, we construct a simple example highlighting this case and show that there is always a range of fixed costs such that a contingent subsidy induces a lower $\hat{c}_{mn,j}$ than the unconditional subsidy. Thus, the existence of substantial fixed costs introduces the possibility of a "win-win" outcome for industrial policy: adding trade restrictions (contingencies) can amplify price reductions as well as deliver a protectionist outcome (facility relocation).[50] However, we find that this does not hold empirically for the EV industry. There are substantial cost savings associated with the higher scale of production induced by an unconditional subsidy. But adding contingencies is not win-win. The assembly contingency achieves the protectionist goal but shrinks the cost savings. And the full local content requirement is lose-lose: it fails to deliver assembly relocation and wipes out the cost savings.

10 Conclusion

Combining subsidies for consumers to buy EVs with incentives for EV makers to source their batteries locally has been a persistent theme of industrial policy in this sector. As described in Barwick et al. (2025), the Chinese government restricted subsidy eligibility to EVs whose batteries came from a "White List" of (Chinese) suppliers. The Biden administration's IRA continued in this vein with a set of escalating upstream requirements to receive the full $7,500 consumer credit for EV purchases. While both of those policies have been phased out, the *Financial Times* reported[51] that the European Commission plans to introduce "Made in Europe" requirements of up to 70 percent for cars and other products. The report generated complaints by automakers and praise by battery makers.

Analyzing the effects of subsidies with and without local content restrictions is not straightforward because the success of such policies depends in large part on the willingness of industries to relocate production. The battery and electric vehicle industry is an important example of a value chain where optimal location conditions are determined in a complex way, balancing trade costs, fixed costs, and a mix of substitution and complementarity effects. Just solving the multi-stage problem computationally has been a challenge, leading most researchers to leave out fixed costs. However, fixed costs are important because they lead to increasing returns, which themselves make policy consequences much harder to predict.

We find that these increasing returns are empirically important: A 20% unconditional consumer subsidy to EVs decreases consumer prices by 25% (for North America) and 23% (for Europe). Those additional price decreases stem from the changes in production paths that would be irrelevant under constant returns to scale. Theoretically, those price decreases could be magnified by introducing local content requirements, which would also satisfy protectionist objectives (the win-win outcome we previously described). Empirically, we find that those production contingencies *shrink* the price savings relative to the unconditional subsidy, leading to substantially

[50] In addition, for a green industry, lower prices both increase consumer surplus and reduce emission externalities.
[51] December 2nd, 2025

lower EV adoption. The combined cell and assembly contingency halves the EV adoption increase relative to the unconditional subsidy.

The unconditional subsidy works both as an environmental and industrial policy, delivering higher EV adoption and additional production lines—both in the subsidizing region and in the other regions. Adding contingencies can deliver some protectionist outcomes, but with different magnitudes and signs. Relative to the unconditional subsidy, the assembly contingency relocates vehicle plants to the subsidizing region, increasing the number of plants by 14% (North America) and 20% (Europe). But it fails to induce the relocation of cell plants.

Imposing the combined cell and assembly contingency induces mixed protectionist outcomes relative to the unconditional subsidy. For North America, there is a relocation of cell production, with a 20% increase in facilities. But this is offset by a 7% decrease in assembly facilities. On the other hand, the same policy for Europe achieves an unambiguous protectionist goal of raising both cell and vehicle production (21% for cells and 7% for vehicles).

As we have just described, the relative appeal of different industrial policies involves a complex set of trade-offs that arise from the endogenous plant relocations across multiple stages of production in an industry equilibrium. Policymakers face similarly difficult evaluations of alternatives for other industries that involve high fixed costs spread across multiple stages of production. For example, the US IRA also specifically targeted the renewable power and semiconductor industries. Those sectors, among others, are common targets for industrial policy—including trade restrictions—by governments around the world. Our framework provides a way of computationally evaluating the effectiveness of a broad range of policy scenarios intended to promote those sectors.

References

AlixPartners (2021). Accompagner la filière automobile dans la transition énergétique, connectée et partagée. Slide Deck.

Allcott, H., R. Kane, M. S. Maydanchik, J. S. Shapiro, and F. Tintelnot (2024). The effects of Buy American: Electric vehicles and the Inflation Reduction Act. Working Paper 33032, National Bureau of Economic Research.

Allen, T. and C. Arkolakis (2022). Welfare effects of transportation infrastructure improvements. *Review of Economic Studies 89*(6), 2911–2957.

Anderson, S., A. De Palma, and J. Thisse (1992). *Discrete choice theory of product differentiation*. MIT Press.

Antràs, P. and A. De Gortari (2020). On the geography of global value chains. *Econometrica 88*(4), 1553–1598.

Antràs, P., T. C. Fort, and F. Tintelnot (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review 107*(9), 2514–64.

Antràs, P. and D. Chor (2022). Global value chains. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics: International Trade, Volume 5*, Volume 5 of *Handbook of International Economics*, pp. 297–376. Elsevier.

Antràs, P., E. Fadeev, T. C. Fort, and F. Tintelnot (2024). Exporting, global sourcing, and multinational activity: Theory and evidence from the United States. *The Review of Economics and Statistics 1*, 1–48.

Arkolakis, C., F. Eckert, and R. Shi (2023). Combinatorial discrete choice: A quantitative model of multinational location decisions. Working Paper 31877, National Bureau of Economic Research.

Arnoud, A., F. Guvenen, and T. Kleineberg (2019). Benchmarking global optimizers. Working Paper 26340, National Bureau of Economic Research.

Barwick, P. J., H.-S. Kwon, S. Li, Y. Wang, and N. B. Zahur (2024). Industrial policies and innovation: Evidence from the global automobile industry. Working Paper 33138, National Bureau of Economic Research.

Barwick, P. J., H.-S. Kwon, S. Li, and N. B. Zahur (2025). Drive down the cost: Learning by doing and government policies in the global EV battery industry. Working Paper 33378, National Bureau of Economic Research.

Ben-Akiva, M. E. and S. R. Lerman (1985). *Discrete choice analysis: theory and application to travel demand*, Volume 9. MIT press.

Bernard, A. B., E. Dhyne, G. Magerman, K. Manova, and A. Moxnes (2022). The origins of firm heterogeneity: A production network approach. *Journal of Political Economy 130*(7), 1765–1804.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Bieker, G. (2021). Global comparison of the life-cycle greenhouse gas emissions of passenger cars. White Paper.

Castro-Vincenzi, J. (2024). Climate hazards and resilience in the global car industry. Mimeo.

Castro-Vincenzi, J., E. Menaguale, E. Morales, and A. Sabal (2025). Market entry and plant location in multi-product firms. Mimeo.

de Gortari, A. (2020). Global value chains and increasing returns. Online note.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced trade. *American Economic Review 97*(2), 351–355.

Dingel, J. and F. Tintelnot (2026). Spatial economics for granular settings. *Econometrica*. Forthcoming.

Fajgelbaum, P. D., P. K. Goldberg, P. J. Kennedy, and A. K. Khandelwal (2020). The return to protectionism. *The Quarterly Journal of Economics 135*(1), 1–55.

Gerarden, T. D., B. K. Bollinger, K. Gillingham, and D. Xu (2025). Strategic avoidance and the welfare impacts of U.S. solar panel tariffs. Working Paper 34401, National Bureau of Economic Research.

Guimaraes, P., O. Figueirdo, and D. Woodward (2003). A tractable approach to the firm location decision problem. *The Review of Economics and Statistics 85*(1), 201–204.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica 52*(3), 541–561.

Head, K. and T. Mayer (2019). Brands in motion: How frictions shape multinational production. *American Economic Review 109*(9), 3073–3124.

Head, K., T. Mayer, and M. Melitz (2024). The Laffer curve for rules of origin. *Journal of International Economics 150*, 103911.

Head, K., T. Mayer, M. Melitz, and C. Yang (2026). The Integer Programming solution to facility location problems. Mimeo.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics 131*(3), 1291–1364.

Jia, P. (2008). What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

Johnson, R. C. and A. Moxnes (2023). GVCs and trade elasticities with multistage production. *Journal of International Economics 145*, 103796.

Khandelwal, A. (2010). The long and short (of) quality ladders. *The Review of Economic Studies 77*(4), 1450–1476.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013). Trade liberalization and embedded institutional reform: evidence from chinese exporters. *American Economic Review 103*(6), 2169–95.

Kummu, M., M. Taka, and J. H. Guillaume (2018). Gridded global datasets for gross domestic product and human development index over 1990–2015. *Scientific data 5*(1), 1–15.

Malgouyres, C., T. Mayer, and L. Nolden (2025). EV incentives and the trade-off between decarbonization and reshoring: Evidence from the French environmental score reform. Mimeo.

McFadden, D. (1978). Modeling the choice of residential location. In A. Karlqvist (Ed.), *Spatial Interaction Theory and Planning Models*, pp. 75–96. Elsevier.

Oberfield, E., E. Rossi-Hansberg, P.-D. Sarte, and N. Trachter (2024). Plants in space. *Journal of Political Economy 132*(3), 867–909.

Ortiz-Astorquiza, C., I. Contreras, and G. Laporte (2018). Multi-level facility location problems. *European Journal of Operational Research 267*(3), 791–805.

Sabal, A. (2025). Product entry in the global automobile industry. Mimeo.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

Su, C.-L. and K. L. Judd (2012). Constrained optimization approaches to estimation of structural models. *Econometrica 80*(5), 2213–2230.

Tintelnot, F. (2017). Global production with export platforms. *The Quarterly Journal of Economics 132*(1), 157–209.

Tyazhelnikov, V. (2022). Production clustering and offshoring. *American Economic Journal: Microeconomics 14*(3), 700–732.

Van Biesebroeck, J. and F. Verboven (2025). Demand, competition and public policy in the automobile industry. Mimeo.

Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics 60*(2), 293–314.

Yi, K.-M. (2003). Can vertical specialization explain the growth of world trade? *Journal of Political Economy 111*(1), 52–102.

Yi, K.-M. (2010). Can multistage production explain the home bias in trade? *The American Economic Review 100*(1), 364–393.

Appendix

A Additional Theoretical Results

A.1 Equivalence between ICE Taxes and Unconditional Subsidies

In this section, we show that a common subsidy $\hat{t}_{mn,j} = \hat{t}_n < 1$ for all EV models m sold in market n is equivalent to an ICE tax $\hat{t}'_n = (\hat{t}_n)^{-1} > 1$ on vehicles sold in n. Under the EV subsidy, the CES price index in n is given by:

$$\hat{P}_{n,j}^{1-\eta} = s_n^{\circ,\text{ICE}} + \hat{\mathbb{C}}_{n,j}^{1-\eta} = s_n^{\circ,\text{ICE}} + \hat{t}_n \hat{C}_{n,j}^{1-\eta}.$$

Under an ICE tax, the CES price index P'_n in n is given by:

$$\hat{P'}_{n,j}^{1-\eta} = \hat{t'}_n^{1-\eta} s_n^{\circ,\text{ICE}} + \hat{C'}_{n,j}^{1-\eta},$$

where $\hat{C'}_{n,j}$ is the new cost index in simulation j under that tax.

Let $\mathbb{A}'_n$ denote the associated anticipated market demand under the tax. If $\hat{\mathbb{A}}'_n = \hat{t}_n^{1-\eta}\hat{\mathbb{A}}_n$, then the anticipated profits for all models m in any simulation j will be identical in both the subsidy and tax scenarios. In that case, all firms will make the same location choices leading to the same cost index in the two scenarios in all simulations j: $\hat{C'}_{n,j} = \hat{C}_{n,j}, \forall j$.

This will hold whenever $\left(\hat{t}_n^{1-\eta}\right)^{-1} \hat{P}_{n,j}^{1-\eta} = \hat{P'}_{n,j}^{1-\eta}, \forall j$, or equivalently whenever:

$$s_n^{\circ,\text{ICE}} \left(\hat{t}_n^{1-\eta}\right)^{-1} + \hat{C}_{n,j}^{1-\eta} = \hat{t'}_n^{1-\eta} s_n^{\circ,\text{ICE}} + \hat{C'}_{n,j}^{1-\eta}, \quad \forall j.$$

Lastly, note that this will be satisfied for all j so long as $\hat{t'}_n = (\hat{t}_n)^{-1}$.

A.2 Delivered Marginal Cost Reductions from Subsidies

With increasing returns to scale (fixed costs and constant marginal costs), it is straightforward to show that a subsidy can reduce the delivered marginal cost. Less intuitively, we also show that adding an upstream restriction on a subsidy can achieve a further reduction in delivered marginal costs relative to an unconditional one. We demonstrate these propositions with a simple two-country example. Recall that $\hat{c}_{mn,j}$ is the change (relative to the data) in delivered marginal cost—exclusive of the subsidy itself—for a given m, sold in market n in simulation j. Let $\hat{c}'_{mn,j}$ denote the cost change under the unconditional subsidy (Policy 1) and $\hat{c}''_{mn,j}$ the cost change under the conditional subsidy (it may require assembly as in Policy 2 or cells and assembly as in Policy 3). Variables without primes correspond to the baseline. We show that there will always be a combination of fixed costs and variable cost differences such that $\hat{c}''_{mn,j} < \hat{c}'_{mn,j} < \hat{c}_{mn,j}$.

Let there be two countries, East (E) and West (W). The maintained assumption is that the consumer in E will be served by a plant in E, no matter what decision is made for the W consumer.

Hence, we can focus on $n = W$ and take a representative m and j. There are two paths $\ell = E$ and $\ell = W$, which correspond to the production location being in the East or West for customers served in West. Under the unconditional policy, both paths are eligible. Let $\hat{c}_W < \hat{c}_E$, which means that the path $\ell = W$ featuring local production has lower delivered marginal cost. The policies are defined such that $\hat{t}'_E = \hat{t}'_W = 1 - s$ and $\hat{t}''_E = 1$ and $\hat{t}''_W = 1 - s$, where the subscript indicates the path to $n = W$ and s is the *ad valorem* subsidy for eligible paths. We assume that model m has a small enough market share that we can ignore changes in the price index in response to the chosen path. We can therefore set $\hat{P}_W = \hat{P}'_W = \hat{P}''_W = 1$. Letting k denote the determinants of variable profit that do not depend on the path, we can then specify variable profits from sales in W under both policies and paths:

$$\begin{aligned} \pi_W &= k\hat{c}_W^{1-\eta}, & \pi_E &= k\hat{c}_E^{1-\eta}, \\ \pi'_W &= \pi_W(1-s)^{1-\eta}, & \pi'_E &= \pi_E(1-s)^{1-\eta}, \\ \pi''_W &= \pi'_W, & \pi''_E &= \pi_E. \end{aligned}$$

For any $s > 0$, $\pi''_E < \pi'_E$, since the contingency reduces the variable profit of serving W remotely.

Given $\hat{c}_W < \hat{c}_E$, the difference in variable profit between the choice of W versus E is lowest under the baseline, higher under Policy 1, and highest under Policy 2:

$$\pi_W - \pi_E < \pi'_W - \pi'_E < \pi''_W - \pi''_E.$$

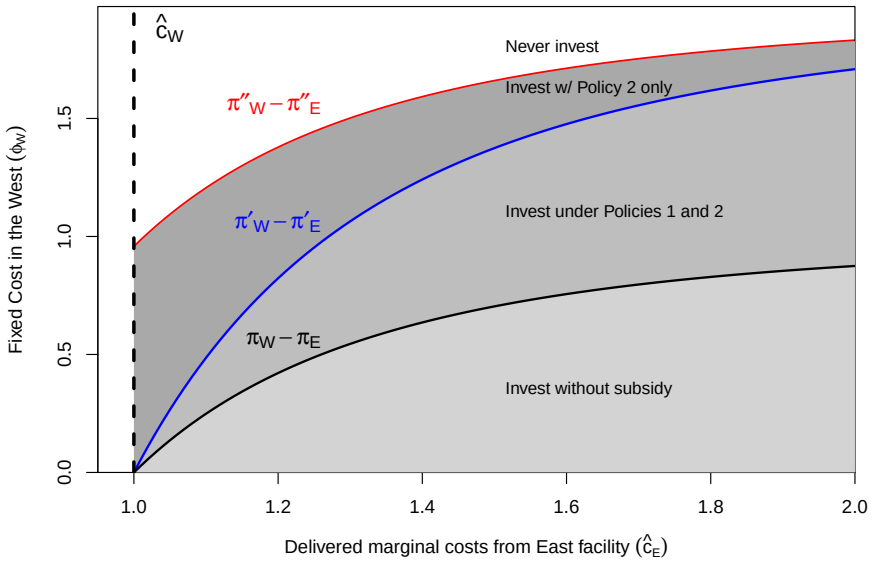
Let ϕ_W denote the fixed cost of activating W ($y_W = 1$), thereby permitting path $\ell = W$. Those profit differences above define four possible ranges for this fixed cost. When $\phi_W < \pi_W - \pi_E$, the firm always chooses W over E. When $\pi_W - \pi_E < \phi_W < \pi'_W - \pi'_E$, the firm no longer chooses W under the baseline. In this case, both subsidies (Policies 1 and 2) decrease the delivered marginal cost to W. When $\pi'_W - \pi'_E < \phi_W < \pi''_W - \pi''_E$, the firm no longer chooses W under the unconditional subsidy. This is the less intuitive case where the local content requirements decrease the delivered marginal cost. And lastly for completeness, the firm will never invest in W when ϕ_W exceeds $\pi''_W - \pi''_E$.

We show those four ranges for the fixed cost ϕ_W in Figure A1 with $k = 1$ (choice of numeraire for profits), $\hat{c}_W = 1$ (horizontal shifter for the whole plot), and $s = .2$. The variable profit differences between W and E are increasing for all scenarios, reflecting the well-known "proximity versus concentration" tradeoff.

A.3 Decomposition of Continental EV Expenditure Change

In Tables 9 and 10 of the main text, we provided a breakdown of the continental EV expenditure changes $\tilde{\hat{R}}^{\text{EV}}_{\mathcal{N}}$ into CES cost index changes $\tilde{\hat{c}}_{\mathcal{N}}$ and $\tilde{\hat{t}}_{\mathcal{N}}$ that average the cost and policy changes at the model-market-simulation level ($\hat{c}_{mn,j}$ and $\hat{t}_{mn,j}$) and the aggregate shifters for EV variety changes $\bar{s}^{\text{EV}}_{\mathcal{N}}$ and market demand $\hat{\mathbb{A}}_{\mathcal{N}}$. In this appendix, we derive a full decomposition of that EV

Figure A1: Delivered marginal cost reductions



expenditure changes $\bar{\hat{R}}^{\mathrm{EV}}_{\mathcal{N}}$ into those components, along with their covariances.

As shown in (29), $\bar{\hat{R}}^{\mathrm{EV}}_{\mathcal{N}}$ aggregates the market-level average EV expenditures $\bar{\hat{R}}^{\mathrm{EV}}_{n}$:

$$\bar{\hat{R}}^{\mathrm{EV}}_{\mathcal{N}} = \sum_{n \in \mathcal{N}} \frac{R^{\circ,\mathrm{EV}}_{n}}{R^{\circ,\mathrm{EV}}_{\mathcal{N}}} \bar{\hat{R}}^{\mathrm{EV}}_{n}, \qquad \bar{\hat{R}}^{\mathrm{EV}}_{n} = \sum_{j} \frac{1}{J} \sum_{m \in \mathcal{M}_{n,j}} s^{\circ}_{mn} (\hat{c}^{1-\eta}_{mn,j} \hat{t}^{1-\eta}_{mn,j} \hat{P}^{\eta-1}_{n,j}).$$

This market-level expenditure can then be decomposed as follows:

$$\bar{\hat{R}}^{\mathrm{EV}}_{n} = \hat{A}_n \frac{\bar{s}^{\mathrm{EV}}_{n}}{s^{\circ,\mathrm{EV}}_{n}} \left[\tilde{\hat{c}}^{1-\eta}_{n} \tilde{\hat{t}}^{1-\eta}_{n} + \mathrm{Cov}(\hat{c}^{1-\eta}_{mn,j}, \hat{t}^{1-\eta}_{mn,j}) \right] + \mathrm{Cov}\left[\left(\hat{P}^{\mathrm{EV}}_{n,j} \right)^{1-\eta}, \hat{P}^{\eta-1}_{n,j} \right], \tag{A.1}$$

where the market-level CES cost index change for $\tilde{\hat{c}}_n$ is defined analogously to $\tilde{\hat{c}}_{\mathcal{N}}$ for a continent $\mathcal{N}$ with a single market n; the market-level CES cost index change for $\tilde{\hat{t}}_n$ is defined in the same way based on $\tilde{\hat{t}}_{\mathcal{N}}$; and $\hat{P}^{\mathrm{EV}}_{n,j}$ is the price index change for EVs only.[52] The covariance terms are

[52] Using (17) and (22), this price index change is given by:

$$\left(\hat{P}^{\mathrm{EV}}_{n,j} \right)^{1-\eta} = \frac{1}{s^{\circ,\mathrm{EV}}_{n}} \sum_{m \in \mathcal{M}^{\circ}_{n}} s^{\circ}_{mn} (\hat{c}_{mn,j} \hat{t}_{mn,j})^{1-\eta} = \frac{1}{s^{\circ,\mathrm{EV}}_{n}} (\hat{P}^{1-\eta}_{n,j} - s^{\circ,\mathrm{ICE}}_{n})$$

given by:

$$\text{Cov}(\hat{c}_{mn,j}^{1-\eta}, \hat{t}_{mn,j}^{1-\eta}) \equiv \frac{1}{J\bar{s}_n^{\text{EV}}} \sum_j \sum_{m \in \mathcal{M}_{n,j}} s_{mn}^{\circ} \left(\hat{c}_{mn,j}^{1-\eta} - \tilde{\hat{c}}_n^{1-\eta}\right) \left(\hat{t}_{mn,j}^{1-\eta} - \tilde{\hat{t}}_n^{1-\eta}\right),$$

$$\text{Cov}\left[\left(\hat{P}_{n,j}^{\text{EV}}\right)^{1-\eta}, \hat{P}_{n,j}^{\eta-1}\right] \equiv \frac{1}{J} \sum_j \left[\left(\hat{P}_{n,j}^{\text{EV}}\right)^{1-\eta} - \left(\bar{\hat{P}}_n^{\text{EV}}\right)^{1-\eta}\right] \left(\hat{P}_{n,j}^{\eta-1} - \hat{\mathbb{A}}_n\right),$$

where $\bar{\hat{P}}_n^{\text{EV}}$ is the CES price index that averages the $\hat{P}_{n,j}^{\text{EV}}$ across simulations such that $\left(\bar{\hat{P}}_n^{\text{EV}}\right)^{1-\eta} = \sum_j \left[\left(\hat{P}_{n,j}^{\text{EV}}\right)^{1-\eta}\right]/J$. The second covariance derivation also uses the definition of $\hat{\mathbb{A}}_n$ in (27) as the average across simulations of $\hat{P}_{n,j}^{\eta-1}$.

Tables 9 and 10 decompose the average EV sales into the continental average of the four components $\hat{A}_n$, $\frac{\bar{s}_n^{\text{EV}}}{s_n^{\circ,\text{EV}}}$, $\tilde{\hat{c}}_n^{1-\eta}$, $\tilde{\hat{t}}_n^{1-\eta}$ in (A.1). This decomposition is exact at the country level when the two covariance terms are zero. The covariance $\text{Cov}(\hat{c}_{mn,j}^{1-\eta}, \hat{t}_{mn,j}^{1-\eta})$ reflects the fact that the CES index of a product is not exactly equal to the product of the CES indexes – except when one of the product components is constant, as in the unconditional subsidy. In the other cases, it is second order (it would not show up in a log-linearization of the CES index); and empirically those covariances are small. The second covariance $\text{Cov}\left[\left(\hat{P}_{n,j}^{\text{EV}}\right)^{1-\eta}, \hat{P}_{n,j}^{\eta-1}\right]$ reflects the capacity for market-level EV sales to increase by displacing ICE sales as P_n^{EV} decreases and P_n^{ICE} remains constant (recall that total vehicle sales R_n are held constant). When the EV share is low, this covariance is negligible. As the EV share rises, the covariance term becomes increasingly negative, offsetting the cost reduction terms.

B Data Appendix

The first three data sets are all proprietary data sets sold to us by IHS Markit. The other data used in the paper are publicly available.

B.1 IHS HVBD

The High Voltage Battery Data (HVBD) module provides value chain data by light vehicle model. We use only Battery Electric Vehicles (BEVs) in our analysis, excluding hybrids, defined here as all vehicles with combustion engines. The data show configurations of brand models, showing the supplier firm and plant location for each cell, module, pack, and battery management system used in a car or light truck. The location and owner of the vehicle assembly plant is also shown, but not the destination market where the vehicle is sold. Quantities for each configuration are provided from 2015 to 2023. However, the last year is currently a projection.

B.2 IHS Sales

The Sales module provides plant of production and destination market for each model sold. The sales quantities are available from 2015 to 2023. This data set lacks detail on the battery. Furthermore, it does not distinguish BEVs, hybrid, and internal combustion versions with the same nameplate (model name). For that we must combine information from the New Registrations data set.

B.3 IHS New Registrations (NRQ, NRP)

The data we refer to as "newreg" comes to us in two pieces: The NRQ file has quantities and no prices. The NRP file has prices and quantities, but not for all models. NRP is also not available for 5 of our markets of interest: Czechia, Slovakia, Mexico, Hungary and Korea. NRQ does not have quantities for Japan. We explain how we deal with those different missing data in section C.

We can obtain ξ_{mn} only for models with prices (the NRP data). The raw data from IHS often express model names slightly differently from the HVBD and Sales data. After some effort to match the model names across IHS modules, we have the merged data has 99% overlap in sales for the set of 15 big firms. Another problem is that some models in the NRQ have Model recorded as the brand name with "Unspecified" as the model name. There are also some Brand names (Make) that are unspecified. Collectively, these cases account for very small shares of the total sales in NRQ, under 2% in all markets in 2022, and 0.5% in all markets except China and Germany. These observations are dropped.

B.4 Tariff data

The primary source of tariffs is WITS, a database maintained by the World Bank. The rates reported are ad-valorem, and at the HS6 digits level of detail. For battery cells, we take the rates reported for HS 850790. For vehicles, we apply the rates of 870380 for passenger cars, and 870490 for commercial vehicles. Note that 870380 is the HS for electric-only vehicle since the major revision of 2017. We use 870390 until the country declares 870380. For both stages we use the data provided by Fajgelbaum et al. (2020) to adjust the rates following the wave of Trump tariffs against China. Regarding the frequent missing values in WITS raw data, we follow the same method as in Head and Mayer (2019): we fill the holes with linear interpolation, and when the missing data is for the latest years, we replace it with the latest available for this pair of countries and this HS.

B.5 Geography

Distances are calculated from plant to plant for cells to assembly, using latitude and longitude data and exploiting the R **geosphere** package for great circle distances. For distance from the assembly plant to consumer market, we use the **world.cities** data set contained in the R **maps** package. It contains latitude and longitude of close to 1000 cities for France, Japan, Germany, the

USA, and Italy, over 500 cities for most countries, with Sweden having the fewest cities at 104. We average the **geosphere**::distGeo distances for each plant-city pair to obtain plant-market distances. To allow discontinuities associated with intercontinental trade, we define continents based on by the R **countrycode** package, which also provides iso codes for each of the countries named in the IHS data sets.

B.6 Sub-national per capita GDP

We construct measures of sub-national level of per capita Gross Domestic Product (GDP) using data from Kummu et al. (2018).[53] They provide a global harmonized and gap-filled per capita GDP dataset at national and sub-national (provincial, municipality) administrative units from 1990 to 2015. We then map EV battery cell and assembly plants' coordinates to their sub-national boundaries by the shortest distance. Once the link between plant and administrative unit is established, we use the per capita GDP measure in 2015 given in constant 2011 international US dollars.

B.7 Demand elasticity estimates

Table B1 provides the preferred (absolute) value of the average own price elasticities estimated in 18 studies that estimate demand for cars (ICE and EVs). The left column gives figures for papers focusing on EVs, while the right column does not distinguish between EVs and traditional gas-powered vehicles, and therefore mostly covers ICEs. In Table B1, the median over all 18 studies is 4. A recent paper by Van Biesebroeck and Verboven (2025) also provides average own-price elasticities for 14 recent papers (mostly for ICEs), with a median value of 5, very close to our right column. We set η to the overall median value: 4.

Table B1: Price responsiveness of car demand from recent literature

EV-only	**elas.**	**Mainly ICEV**	**elas.**
Barwick et al. (2024)	4.2	Beresteanu and Li (2011)	8.4
Kwon (2023)	4.4	Castro-Vincenzi (2024)	4.3
Li et al. (2022)	3.7	Colon and Gortmaker (2020)	3.9
Li (2023)	3.7	Coşar et al. (2018)	14.9
Li et al. (2017)	1.3	Goldberg (1995)	3.3
Linn (2022)	5.3	Goldberg and Verboven (2001)	5.2
Muehlegger and Rapson (2022)	2.1	Grieco et al. (2024)	5.4
Springel (2021)	1.8	Head and Mayer (2019)	3.9
Xing et al. (2021)	2.8	Li (2018)	9.5
Median	3.7	**Median**	5.2
Overall Median	4.0		

[53] The dataset can be downloaded from `https://datadryad.org/dataset/doi:10.5061/dryad.dk1j0`.

C Calibration of demand parameters

As shown in section 8.1, expected fixed cost draws rise with the quality of the cars associated with each factory, which we denote as $\tilde{\xi}_{fg_k}$. The implicit assumption is that manufacturing cells or vehicle assembly for higher-quality cars requires more complex and expensive equipment. The primitive underlying the firm-group quality index is ξ_{mn}, the model-market demand shifters we refer to as "appeal" following Hottman et al. (2016). The distinction we draw between appeal and quality is that the former varies across countries, depending on consumer perceptions, whereas the latter consists of the universally valued elements of the vehicle. We obtain our quality index as a geometric mean across markets where the model is sold and the models in a given group:[54]

$$\tilde{\xi}_{fg_k} = \exp\left(\sum_n \sum_{m \in \mathcal{M}_f} z_{mn} \Gamma^k_{mg_k} \ln \xi_{mn}\right)$$

The calibration of ξ_{mn} demand shifters uses IHS model-level car prices in each market, as well as the sales of EV and ICE cars. Demand is given by

$$q_{mn} = R_n P_n^{\eta-1} \xi_{mn}^{\eta-1} p_{mn}^{-\eta}, \tag{C.2}$$

Inverting (C.2), we recover ξ_{mn} as a ratio of appeal relative to the geometric mean of EV models in market n.[55] Letting $\tilde{p}_n$ and $\tilde{q}_n$ denote the geometric means of EV prices and quantities in market n, the inversion yields

$$\xi_{mn} = \left(\frac{p_{mn}}{\tilde{p}_n}\right)^{\frac{\eta}{\eta-1}} \left(\frac{q_{mn}}{\tilde{q}_n}\right)^{\frac{1}{\eta-1}}, \tag{C.3}$$

where we have adopted the Hottman et al. (2016) normalization that the geometric mean of ξ_{mn} is one in every market n (only relative appeal matters). Using our estimate of $\eta = 4$, we compute the relative appeal, ξ_{mn} of each EV model in each market.

The next step is to attribute groups for each model m (that is to fill in the $\Gamma^k_{mg_k}$ matrix). The **newreg** data does not provide information on characteristics of the cell or the platform on which the vehicle was built. However, it provides "trims", which are destination specific names ("Long-range" or "Standard" for instance), associated with the weight of the vehicle and its price, and also the power of the electric motor. For vehicle platforms, things are straightforward since all our models of interest are built on a single platform in 2022, independently of their trim, plant of assembly and destination. For cell groups, a little more work is needed. As shown in Table D3, most brand-model combinations come with only one cell group (NC-Pouch for instance), independently of the trim. However, 11 car models have two cell groups in substantial proportions

[54] Recall from equation (4) that $\Gamma^k_{mg_k} = 1$ for m in group g_k and zero otherwise.

[55] Our approach here is closest to Hottman et al. (2016). It is an intellectual descendant of the Berry (1994) inversion but here we used quantities relative to an inside good. Khandelwal (2010) also backs out appeal in a logit demand, but he uses a regression approach with instruments for prices. Khandelwal et al. (2013) imposes a price responsiveness parameter as we do, but then estimates appeal across markets as a residual from a fixed-effects regression.

(those are high sales vehicles such as the Tesla 3 and Y, the VW ID.3 and ID.4, the Kia Niro and the MG 4 for instance). For those, we have to map trims to cell groups. Manual online search shows that more powerful battery packs are associated with heavier and pricier vehicles and more powerful electric motors. We therefore compute median price, weight, and horsepower, and classify each trim as above or below those medians. Taking the average of those classifications for a given trim-destination, we then attribute the most powerful cell group if the mean classification is ≤ 1.5.

Then we need to attribute a ξ_{mn} to model-destination combinations without price data in NRP. This is done with a regression-based approach on the sample with price data in NRP:

$$\log \xi_{mn} = -0.35 \times \text{Border}_{h(m)n} + 0.049 \times \text{RTA}_{h(m)n} + \text{FE}_m + \text{FE}_n + \epsilon_{mn},$$

where $h(m)$ is the headquarters country of brand-model m, adding distaste for non-domestic and non-RTA models to model and destination fixed effects (FE_m explains nearly three quarters of the quality variance in our sample). We then reconstruct missing ξ_{mn} as $\exp(\text{FE}_m - 0.35 \times \text{Border}_{h(m)n} + 0.049 \times \text{RTA}_{h(m)n})$.

Table C2 shows the top 10 and bottom 10 models in terms of appeal (for the models serving at least 10 markets, and not accounting for observations reconstructed with the method above). Topping the list of appeal are models by Porsche, BMW, Audi and all four Tesla models. The four top models are between 2.5 and 3 times more attractive than the average offer over the set of countries where they are sold. It is revealing that the destination in which those are particularly appealing is China, a market in which many inexpensive models are available. Our computations reveal that, given its high price, the Mercedes-Benz EQS must be 6 times more attractive than the average model in China to achieve its actual level of sales there, against less than 2 times the average model sold in Canada. The bottom appeal list features smaller cars made by a group of firms that has almost no intersection with the set of firms making the top 10 models. The models in that bottom group are revealed to have an appeal between 35 and 70% of the average car. Another notable pattern is that there seems to be some home brand advantage: The Renault Zoe has most appeal in France, and the two MG models have most appeal in the UK, where the brand was born (even though it is now owned and operated by SAIC).

D Additional empirical results

D.1 Evidence on multi-sourcing and capacity increases

Figures D2 and D3 are examples of multi-sourcing for vehicles found in our data. In figure D2, we focus on sales of Tesla models in Germany. The Model S and Model X are always sourced in Fremont, those are single sourcing cases all along. The Model 3 (in blue) comes from both Fremont and Shanghai in 2020. But this year is also the opening of the Chinese factory. Sales from Shanghai continue to increase thereafter, while the Californian-made Model 3 are not exported to Germany as soon as 2021. The story is very similar with the Model Y (in red). It comes in 2021

Table C2: Quality and market-level appeal of the top 10 and bottom 10 models

brand model	N mkts	Quality ($\tilde{\xi}_m$)	best ξ_{mn} country	best ξ_{mn} value	worst ξ_{mn} country	worst ξ_{mn} value
Highest quality models						
Porsche Taycan	19	3.02	CHN	5.82	CAN	1.91
Tesla Model Y (N)	18	2.86	CHN	4.24	JPN	1.96
Mercedes-Benz EQS	19	2.76	CHN	6.09	CAN	1.76
BMW iX	18	2.55	CHN	4.12	FRA	1.55
Audi e-tron GT	18	2.06	AUT	3.38	PRT	1.41
Tesla Model 3 (N)	19	1.87	CAN	3.26	NOR	1.23
Audi e-tron	18	1.85	NOR	2.84	POL	1.10
Tesla Model S	15	1.85	USA	4.07	PRT	0.87
Tesla Model X	14	1.74	USA	4.59	PRT	1.08
BMW i4	19	1.68	PRT	2.31	CAN	0.92
Lowest quality models						
Citroen C4	14	0.70	ESP	1.14	POL	0.34
Toyota bZ4X	15	0.68	CHN	1.03	USA	0.41
Renault Zoe	15	0.65	FRA	1.24	NOR	0.37
Skoda Enyaq Coupe	11	0.64	CHE	1.07	BEL	0.44
Kia Niro (N)	10	0.55	GBR	1.06	USA	0.25
Opel Corsa	15	0.55	DEU	0.98	FIN	0.32
MG ZS (N)	13	0.54	GBR	0.92	CHN	0.17
BMW iX1	10	0.53	NOR	0.80	GBR	0.20
Kia Soul	12	0.51	PRT	0.76	ITA	0.21
Dacia Spring	11	0.38	ITA	0.68	DNK	0.22
MG 4 (N)	11	0.36	GBR	0.57	AUT	0.19

Note: Models ranked by their quality, $\tilde{\xi}_m$, measured as the model m geometric mean of ξ_{mn} over all markets n in 2022. Only models with data in 10 or more market included.

from Shanghai, and there is multi-sourcing in 2022, with the Berlin factory opening and gradually ramping up production. The made-in-China Model Y volumes fall in 2023 and go to zero in 2024. Both cases of multi-sourcing appear to be transitions. The story is the same for the Volkswagen ID.4 sold in North America as shown in figure D3(a). The model is launched in 2021 in the USA and Canada, and sourced in the German plant of Mosel. The US-based Chattanooga plant starts producing this car in 2022, with the flows from Germany going to zero in 2023. Mexican imports are negligible, but come from the US factory.

Figure D2: Transitory multi-sourcing by Tesla in Germany

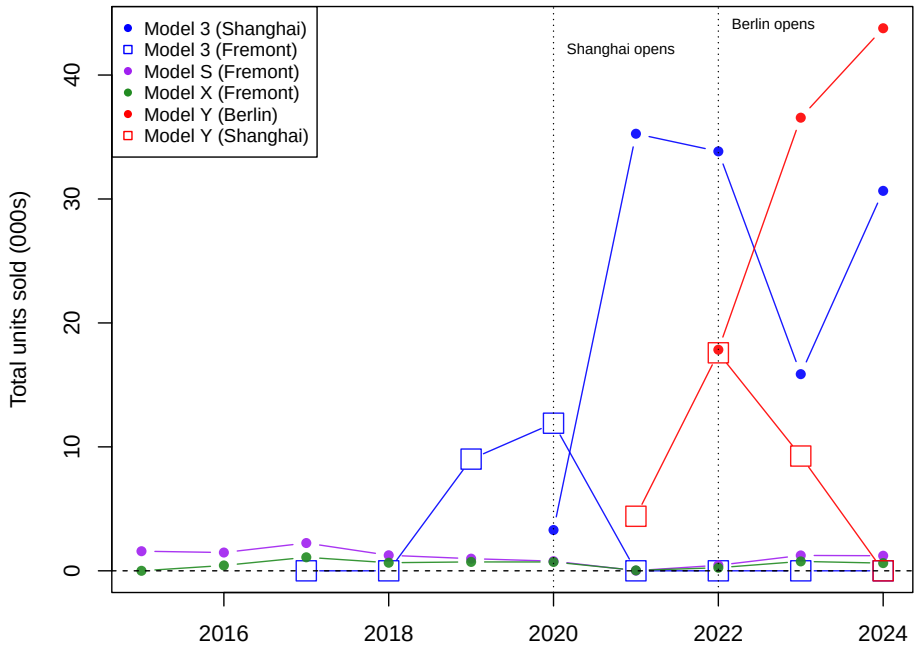
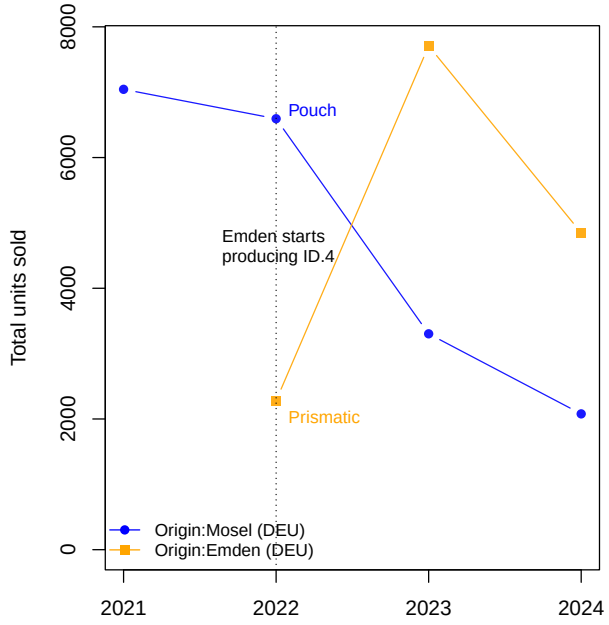


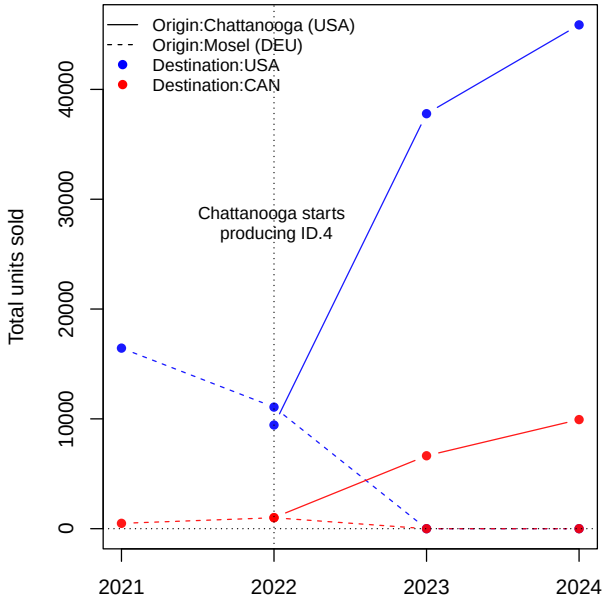
Figure D3(b) shows a different case where the same VW ID.4 is imported by Sweden from two German plants starting in 2022, and continues that way onward. It turns out that those two factories are producing different shapes of the NMC battery which powers this car. Although we do not observe it directly in our sales data, those ID.4s coming from Mosel and Emden are likely to be different trims with different battery specifications. We will treat those as such in our regressions.

Figure D4 shows capacity increases in two factories owned by Tesla: the first one in Fremont, California, and the Shanghai factory opened in 2020. Total output in 000s of cars produced is depicted in black, and specific models are in other colors. In both cases, we see that the production goes from relatively small amounts (around 50 thousands in 2015 for Fremont, and less than 200 thousands for Shanghai in 2020) to several multiples of the initial number very fast. The Fremont factory output is multiplied by around 10 in 8 years, and the factor is more than 4 in 3 years for Shanghai. Capacity constraints do not seem a dominant feature in this industry.

Figure D3: Transitory vs trim-based multi-sourcing

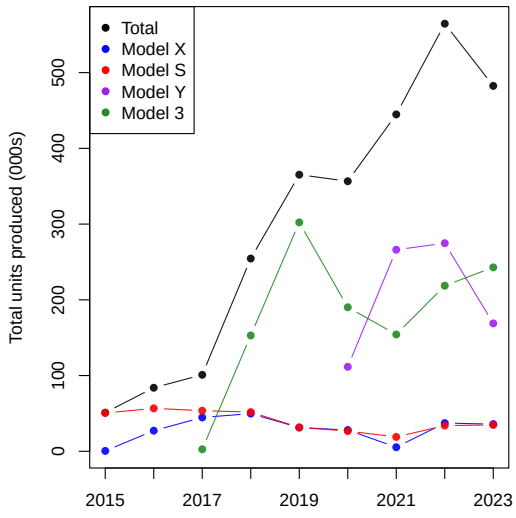


(a) VW ID.4 sold in Sweden

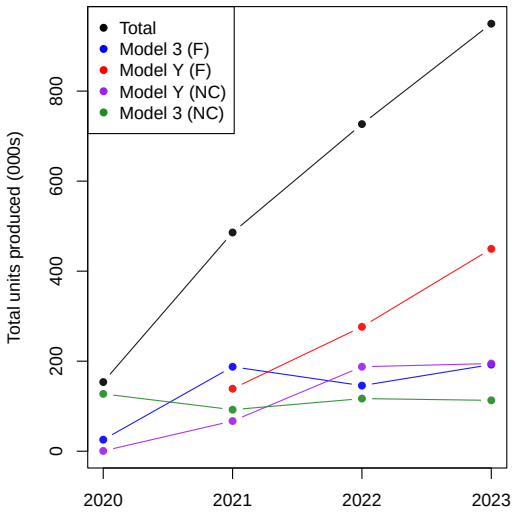


(b) VW ID.4 sold in North America

Figure D4: Tesla's production increases in Fremont and Shanghai



(a) Tesla-Fremont



(b) Tesla-Shanghai

D.2 Cell and vehicle groupings

Our model assumes that a distinct fixed cost is paid for each group g_k installed in cell and assembly plants. The matrix Γ^1 with rows m and columns g_1 is composed of indicators equal to 1 if model m uses group g_1 for cells and zero otherwise. The groups for cells distinguish combinations of material category (NMC or LFP) and shape (cylinder, pouch or prismatic) of the cell. For vehicles, the indicators in the Γ^2 matrix specify the *platform* g_2 for each model m, e.g. "GEN III" for Tesla Model 3 and Model Y, "MEB" for VWs such as ID.4, ID.3, Audi Q4, and Skoda Enyaq. Platforms establish the architecture of core vehicle systems (e.g. steering, chassis) and dictate the placement of their components. They require a specific organization of the assembly line and its machine tools.

Table D3: Cell variants per model

# of matcat-shapes per model	count
Single group	102
Two groups	24
— one dominant by 75%	13
— none dominant	11

Note: The table counts material-shape groups for cells used in a model in 2022 and separates models into single-group or two-group categories (material consists of two categories, NMC or LFP).

Table D3 reports the number of cell groups (material category and shape) that can be found for each of the car models we consider in the counterfactual analysis. Out of the 128 models, there are 24 which are built with two different cell groups. In more than half of the cases (13 out of 24), one group is largely dominant (more than three quarters of the output), such that we attribute the dominant group to this car model. For the remaining 11 car models with no dominant group, we consider both versions as separate car models, which brings the overall number to 139.

Table D4 describes how many variants of the cells are produced in a typical cell factory. It turns out that the vast majority of plants (85%) produce only one shape or one material category (70%) of cells. Even for both characteristics constituting g_1, the most frequent case (63%) is that a plant produces only one combination. Even expressed in shares of total output (measured in GWh), a little more than a quarter of world cell production takes places in plants producing more than two groups.

E Untargeted fit

This section provides evidence of the model's fit to untargeted moments. Figure E5 compares the continent-to-continent shares of EV models sourced in each stage, with the observed shares in 2022 on the x-axis, and the predictions on the y-axis. Figure E6 draws similar bilateral shares, but

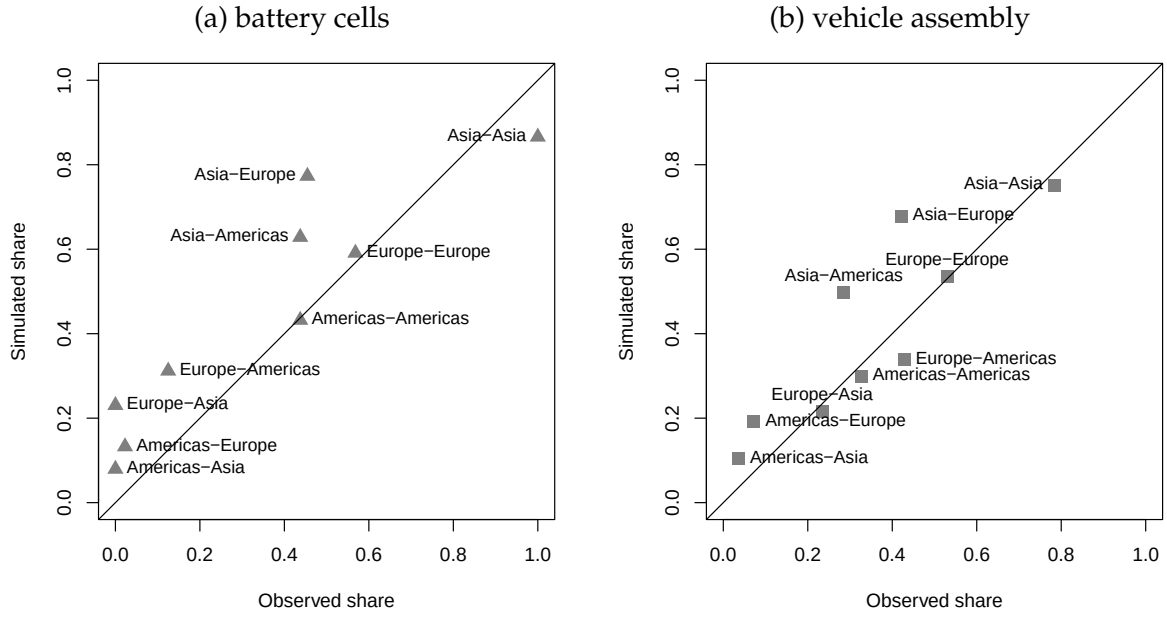
Table D4: Cell plants are specialized in shape and material produced

	Shape			Mat. Cat.			Shape & Mat. Cat.		
# variants:	count	%	% GWh	count	%	% GWh	count	%	% GWh
1	73	84.9	66.0	60	69.8	45.5	54	62.8	37.8
2	12	14.0	32.0	24	27.9	33.9	26	30.2	35.7
3	1	1.2	2.1	2	2.3	20.6	4	4.7	4.1
4	0	0.0	0.0	0	0.0	0.0	2	2.3	22.4

Note: The table gives counts and shares of shapes and material category produced in cell plants. Material category consists of two categories (NMC or LFP). Counts are numbers of cell plants in 2022.

in terms of counts of sourcing assignments $m\ell_k\ell_{k+1}$. For the two figures, the fit is overall good with the most salient deviation being the lower realized shares of Asia to Europe and America shares in both stages of production.

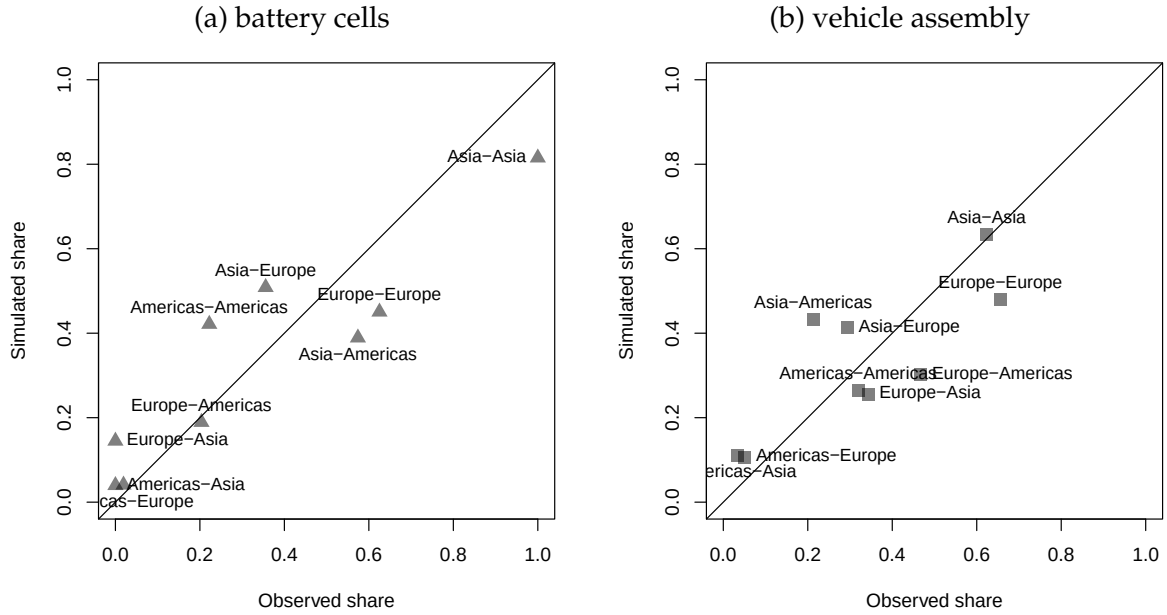
Figure E5: Untargeted fit of the estimated model: origin-destination model sourcing



Notes: These are the share of BEV models each continent sources from the others, for both cells and vehicles.

Table E5 lists a number of monadic untargeted moments for both stages and the 3 continents. The number of locations entered, the count of plants, the number of car models per plant, and the number of model-plant combinations active. Our model does reasonably well in terms of locations entered and plants opened, but overestimates the counts of models per plant.

Figure E6: Untargeted fit of the estimated model: origin-destination sourcing arcs



Notes: These are the share of assignments $m\ell_k\ell_{k+1}$ over all assignments received at destination aggregated to continental level, for both cells and vehicles.

Table E5: Untargeted fit in other extensive margins

Continent	Stage	Potential locations	Data				Model Prediction			
			# loc. entered	# plants	# models per plant	# model-plant combination	# loc. entered	# plants	# models per plant	# model-plant combination
Americas	C	4	4	6	1.3	8	3.2	5.4	3.5	18.8
Asia	C	5	4	29	4.1	118	5	28.4	5.2	148.3
Europe	C	3	2	9	3	27	3.0	8.9	5.3	47.7
Americas	V	5	5	9	2	18	4.3	8.4	2.8	23.2
Asia	V	5	5	24	4	96	5	31.9	3.7	118.7
Europe	V	5	5	15	2.9	44	5.0	18.9	3.7	70.4

Predicted values are averages across simulations.

F Additional Counterfactual Results

F.1 Policy Results

Here, we provide results for the full set of counterfactual scenarios. We analyzed five different policy changes (with only the first three reported in the main text):

1. **Subsidy to EV buyers:** The government subsidizes EV purchases regardless of the locations of assembly and battery manufacture.
2. **Subsidy conditional on domestic assembly:** Only cars assembled in the same continent as the buyer are eligible for the subsidy.
3. **Subsidy conditional on domestic cell-vehicle supply chain:** To qualify, both cell and vehicle must be produced in the same continent as the buyer.
4. **Production subsidy for domestically produced cells:** We denote the combination of production subsidies with the contingent consumer subsidy policy 3 as "3+4".
5. **Tariffs on imported cells and vehicles:** Tariff on both imported cells and vehicles are policy 5. A variant ("5V") restricts the tariff to vehicles.

In order to make the scenarios comparable, we set a common counterfactual tariff or subsidy rate of 20% at the continental level: $\hat{t}_{mn,j} = .8,\ \forall n \in \mathcal{N}$ for the subsidies and $\hat{t}_{mn,j} = 1.2,\ \forall n \in \mathcal{N}$ for the tariff—whenever model m is eligible under the counterfactual scenario. Cell tariffs are eligible and imposed if cells are produced outside and exported to the policy region. Vehicle tariffs are eligible and imposed if EVs are assembled outside and exported to the policy markets.

Section 9 in the main text presents the results graphically for a subset of the policies we investigated. Here we present the complete set of results in tabular form, for production lines in Table F6 and for EV expenditures in Table F7. Those are provided for all our counterfactual policies.

Table F6: Production lines† by policy

Receiving	Policy*							
Production	0	1	2	3	4	3+4	5V	5
North American policies								
				Vehicles				
Americas	9.6	11.1	12.6	10.3	9.9	10.9	10.6	9.1
Asia	53.2	53.6	52.3	52.7	53.2	52.7	52.2	52.5
Europe	24.4	24.6	24.0	24.1	24.3	24.0	23.9	24.2
				Cells				
Americas	6.1	6.5	6.4	7.8	8.1	9.7	6.0	6.3
Asia	33.5	34.4	33.9	32.5	32.8	31.8	33.1	32.7
Europe	9.6	9.9	9.9	9.2	9.4	9.0	9.4	9.3
European policies								
				Vehicles				
Americas	9.6	10.3	8.2	8.6	9.7	8.5	8.2	8.3
Asia	53.2	56.1	45.1	47.8	53.0	46.7	45.6	46.5
Europe	24.4	28.4	34.2	30.4	25.4	31.9	28.5	27.1
				Cells				
Americas	6.1	6.6	6.5	5.3	5.7	4.8	5.9	5.7
Asia	33.5	36.0	34.3	28.5	31.6	26.3	32.1	31.0
Europe	9.6	11.0	11.3	13.3	11.7	14.9	9.8	10.3

† Production lines are defined as model-plant combinations.
* Policies defined at the beginning Appendix F. Policy 0 is the baseline with all existing 2022 policies absorbed in the estimated path costs.

Table F7: Expenditures on EVs by policy*

Receiving	Policy						
Production	1	2	3	4	3+4	5V	5
				North-American policies			
Americas	86.4	71.1	32.3	5.1	46.2	-9.8	-16.1
Asia	1.4	0.8	-0.3	1.9	1.1	-0.3	-1.1
Europe	1.6	0.5	-0.5	1.2	0.7	-0.8	-1.7
				European policies			
Americas	2.2	0.6	0.5	6.2	5.9	-1.2	-1.8
Asia	1.9	0.8	0.1	4.3	4.7	-0.8	-1.3
Europe	62.6	47.6	30.6	6.3	43.2	-10.6	-14.7

* Policies defined at the beginning Appendix F.

F.2 Zero vs Baseline Fixed Costs

Table F8: Comparison with zero fixed costs

Policy	Path cost index change		EV expenditure change	
	No FC	Baseline FC	No FC	Baseline FC
		Americas		
1: Unconditional	0.00	-4.50	46.30	86.40
2: Continental V	1.90	-3.40	37.40	71.10
3: Continental V+C	2.90	2.20	13.30	32.30
4: Subsidy C	0.10	0.50	2.20	5.10
3+4	4.70	3.20	19.70	46.20
5V: Tariff V	1.40	1.10	-7.20	-9.80
5: Tariff V+C	1.10	1.70	-12.50	-16.10
		Europe		
1: Unconditional	0.00	-3.00	34.00	62.60
2: Continental V	2.10	-1.60	27.00	47.60
3: Continental V+C	3.40	0.20	17.30	30.60
4: Subsidy C	0.20	-0.40	3.80	6.30
3+4	4.70	0.30	23.90	43.20
5V: Tariff V	1.60	0.90	-6.30	-10.60
5: Tariff V+C	1.40	1.00	-9.10	-14.70