Markup Heterogeneity, Openness and the Pro-Competitive Gains from Trade with Persistent Trade Relationships

Preliminary and incomplete - Please do not circulate

Paul Ko* María Pía Olivero† Wei Qian‡

May, 2026

Abstract

We study the pro-competitive gains from trade arising from a nested-CES demand structure and the implied non-constant markups in a dynamic model with trade relationships that persist over time. To do so we extend the framework in Firooz, Heins and Mathur (2025) to introduce dynamics and allow for "deep habits" á la Ravn, Schmitt-Grohé and Uribe (2006) in trade relationships. Specifically, we want to model the case of relationships between importers and exporters that are costly to build, and how this feature will impact how exporters set prices when facing changing trade costs in foreign markets. Will the introduction of this dynamic element make pro-competitive gains more likely, even when head-to-head *competition across countries is not that prevalent?*

Markups vary over time due to the differences between within- and across-sector elasticities of demand (just as in previous work) as well as because of a "deep habits" feature that makes markups decreasing in the amount of past trade. The latter effect is along the lines of models with switching costs in the IO literature.

In the empirical section, we work with detailed establishment-level data for India between 2009 and 2013 from the Annual Survey of Industries of the Ministry of Statistics and Programme Implementation. While previous work has focused on China and Taiwan, in India the structural reforms and opening up of the economy were different and took place much later.

Our (still in progress) analysis highlights the importance of cross-country differences in demand elasticities and the degree of persistence in trade relationships in shaping the welfare gains from trade. We also explore the role of sector-specific tariff structures in this framework with dynamic elements.

*Swarthmore College
†Swarthmore College
‡Haverford College

1 Introduction

We study the pro-competitive gains from trade arising from a nested-CES demand structure and the implied non-constant markups in a dynamic model with trade relationships that persist over time. In contrast to the classic "small gains" result emphasized by the recent literature, we expect to find more sizable gains arising from the fact that it is costly to break importer-exporter relationships and, as a result, these become persist over time. Our analysis (still in progress) highlights the importance of cross-country differences in demand elasticities and the degree of persistence in trade relationships in shaping the welfare gains from trade. We also explore the role of sector-specific tariff structures in this framework.

We extend Firooz, Heins and Mathur (2025, hereafter FHM) and Edmond, Midrigan and Xu (2015, hereafter EMX) to allow for dynamics in trade relationships. Specifically, we want to model the trade relationships between importers in country n and exporters from country i. Since building these relationships is costly, the more they import in period $(t-1)$, the more they will find it optimal to import from that particular country in the following period t. Importing from others would require importers and exporters to incur the costs of building a new relationship, which is non-optimal. Thus, these importer-exporter links are persistent over time.

We ask how this feature will impact how exporters set prices when facing a change in trade costs in foreign markets and thus, how these dynamics in trade links will shape the pro-competitive gains from trade.

These relationships are modelled through the well-known "deep habits" feature á la Ravn, Schmitt-Grohé and Uribe et al (2006, hereafter RUSG). The producers of final goods will develop "deep habits" in the demand for intermediate inputs. This makes markups inversely related to the level of trade, a behavior consistent with the empirical evidence on the countercyclicality of markups (see Domowitz et al. (1986), Lebow (1992), Chevalier and Scharfstein (1995 and 1996), Galeotti and Schiantarelli (1998), Bloch and Olive (2001) and Campello (2003), among others).

We believe it makes sense to think that over time firms develop habits for particular inputs from a particular exporter in country i. These inputs are already imperfectly substitutable with each other, but over time this can be increasingly so as the more an importer in country n buys

from a particular producer in country i in $t-1$, they have to adjust their production processes and plant structure to this type of input, so the more they will need to buy from that same producer i in t.

Our intuition is that with "deep habits" (or for what matters any other friction that generates markups that are endogenously related to the volume of trade and not just trade shares), the pro-competitive gains from trade are more likely to be positive. One of the main results in EMX is that a weak pattern of comparative advantage in individual sectors is a requirement for large gains from trade. We conjecture that this might not be such a strong requirement once we account for dynamics of trade relationships. Below we explain why.

EMX find that trade can significantly reduce markup distortions and therefore, generate welfare gains if two conditions are satisfied: (i) there must be large inefficiencies associated with markups, i.e., extensive misallocation, and (ii) trade must in fact expose producers to greater competitive pressure. The second condition is satisfied if there is a high degree of cross-country correlation in the productivity with which producers within a given sector operate ($\tau(\rho)$ is high).

Since the elasticity of demand and hence the markup will be a function of aggregate trade X, and X is a function of sectoral price independently from the sector shares, even when $\tau(\rho) = \left(1 - \frac{1}{\rho}\right)$ is high and the sectoral share in the domestic market is not very sensitive to changes in trade costs, sectoral prices should still fall in response to opening to trade (reduction in the iceberg trade cost τ) so that the elasticity of demand should still rise and markups should still fall. So, this extension might relax the requirements (in terms of the pattern of comparative advantage) for obtaining large gains from trade.

While the current empirical evidence on pro-competitive gains is mostly from Edmond, Midrigan and Xu (2015) for the case of Taiwan and Hsu, Lu and Wu (2020) for the case of China, we study the economy of India. We see this as an added contribution of our work. The opening of Taiwan began in the late 1950's and was an export-oriented industrialization strategy for an economy deeply integrated into global manufacturing supply chains.[1] Taiwan's strong export

[1]In Taiwan the transition away from import-substitution policies began in the late 1950's, with most scholars identifying the period 1958–1962 as the key turning point toward export-oriented industrialization. During this period, Taiwan reduced foreign-exchange controls, simplified import and export procedures, introduced export incentives and tax rebates, devalued the currency to support competitiveness, and shifted the emphasis of their

promotion policies allowed for a gradual integration into global markets. Very differently, India's opening process represented a much later and more abrupt shift away from a highly regulated and inward-looking economic regime. Beginning in 1991, India implemented a broad set of structural reforms in response to a balance-of-payments crisis, including substantial reductions in tariffs and import licensing requirements, deregulation of domestic industries, liberalization of foreign direct investment, and a progressive opening of financial and trade markets. Also, whereas Taiwan's earlier liberalization occurred in the context of a relatively small economy, India entered the liberalization process as a much larger and more domestically oriented economy with widespread government intervention and high trade barriers. As a result, the implications of trade opening in India potentially differ in important ways from those documented in previous work.

Our results are still in progress, and they will be available for presentation at the conference. We will show the model's response to a unilateral trade liberalization, as well as to a global tariff war in which all countries raise their tariffs retaliating with each other. We also aim to study the response of the model to tariffs that vary at the sector-level. We measure welfare gains from trade and we define the "pro-competitive" gains as the difference in welfare measures between our model and one with constant markups (in which trade does not impact firms' markups and profit structure). We will also solve the FHM version of the model without deep habits (which is embedded in our model when setting a parameter ψ=0). This will allow us to quantify the contribution (to gains from trade) of dynamics versus that of changing trade shares. Last, we will analyze the response of wages and the implications of persistent trade relationships for labor markets. Does the fact that importers are to a degree "locked-in" their relationships with existing exporters matter for their labor demand behavior and therefore, for their willingness to pay their employees?

policies toward manufacturing exports.

2 Literature Review

Our paper fits broadly in the international trade literature that studies endogenous markups with country- and sector-specific demand features. This literature underscores the importance of understanding the distribution and dispersion (across both sectors within a country and countries) of markups and the pro-competitive gains (anti-competitive losses) from trade. The aggregate gains (losses) do depend in fundamentally interesting ways on the features of that distribution.

Edmond, Midrigan and Xu (2015) (hereafter EMX) study gains from trade that are efficiency-driven, and related to pro-competitive and misallocation-reducing forces. They do so in a multi-sector model with oligopolistic competition and variable markups. The model features a nested CES demand system á la Atkeson and Burstein (2008) which implies a linear relationship between inverse markups and market shares. The authors discipline the model using Taiwanese manufacturing data at the firm-level to calibrate sectoral shares and elasticities of substitution, fixed operating and export costs, and the distribution of productivity including the correlation structure of sector-specific productivity shocks across countries.

Quantitatively, the paper shows that trade can substantially reduce markup dispersion and hence misallocation when dominant firms in the domestic country compete with foreign firms that produce with close enough productivity levels, i.e. firms that engage in head-to-head international competition. In the benchmark study of the Taiwanese cases, moving from autarky to observed trade levels raises aggregate productivity by about 12%. A drawback of EMX is that their results are very sensitive to the distribution of productivity and market shares across countries. When productivity is assumed to be not highly correlated across countries, trade actually results in losses.

Despite the presence of variable markups and oligopolistic competition, in Edmond, Midrigan and Xu (2015) total gains from trade still remain well approximated by the Arkolakis–Costinot–Rodríguez-Clare (ACR) formula according to which gains from trade can be approximated using only the elasticity of trade flows with respect to trade costs and the domestic expenditure share (the share of aggregate spending on home goods).

Firooz, Heins, and Mathur (2025) (hereafter FHM) develop a quantitative multi-country, multi-sector trade model with variable markups to study how pro-competitive gains from trade differ across countries and sectors. The key innovation is the integration of rich cross-country heterogeneity in demand elasticities, firm size distributions, productivity, trade costs, and tariffs within a nested CES framework where markups respond endogenously to market shares. Using trade data for 29 countries and estimating roughly 3,000 sector- and country-specific import demand elasticities, the authors show that richer countries tend to trade goods with lower demand elasticities (higher markups), while poorer countries specialize more in high-elasticity (lower markup) sectors. Quantitatively, the paper challenges the notion that trade universally delivers sizable pro-competitive benefits. Consistent with the "elusive gains" idea from previous work, FHM obtain anti-competitive losses from trade for a significant share of the countries in their sample. This is because a trade-off arises in the model after a country opens up to trade. On the one hand, foreign exporters gain increased market power which raises the degree of misallocation and is welfare-reducing. On the other, domestic markups fall which reduces the misallocation friction. For several countries that comprise the sample of FHM, the former effect outweighs the latter. Larger and less open countries are the ones that tend to gain more from trade. The effects depend critically on trade shares and the responsiveness of markups to trade shares which in turn depends on the difference between across- and within-sector demand elasticities. An important conclusion from FHM is that it is not just the level of ad-valorem tariffs that matters for how much countries gain (or lose) from trade, but the sectoral structure/distribution of tariffs. This is a crucial result because it implies that trade policy should focus not just on the level of tariffs but on the dispersion across sectors.

In a similar modeling framework, Firooz (2024) studies labor market distortions as a transmission channel through which trade-induced competition can reshape both aggregate welfare and distributional outcomes. He does so by extending the static framework in FHM in two ways: through dynamic considerations and through frictions in labor markets. In the model, firms operate under oligopolistic competition á la Atkeson–Burstein (2008), so markups depend on firm market shares and therefore respond to trade liberalization through changes in demand

elasticities. On the labor side, firms face search and matching frictions, vacancy posting costs and firing costs. Workers and firms engage in a wage bargaining process.

A key mechanism is that lower trade costs and increased foreign competition raise the elasticity of labor demand and make it from responsive to idiosyncratic shocks. This amplifies job reallocation, increases dispersion in firm size, and shifts employment toward larger, higher-wage firms, which in turn affects unemployment and residual wage inequality. The model is disciplined using Colombian plant-level data, and reduced-form evidence shows that firms with higher labor elasticities of revenue exhibit higher job turnover and that trade liberalization raised both elasticities and job reallocation.

Relative to a constant-markup version of the model, trade liberalization leads to substantially larger increases in unemployment and residual wage inequality—nearly three times as large in the steady state. The rise in unemployment reduces the mass of operating varieties and raises the price index, offsetting standard pro-competitive channels. As a result, the pro-competitive gains from trade become negative in the presence of labor market frictions.

Feenstra (2018) complements EMX by providing evidence that the magnitude of pro-competitive gains is highly model-dependent and sensitive to structural assumptions on the distribution of sectoral productivities. Feenstra (2018) reopens the question of whether monopolistic-competition models can generate meaningful variety and pro-competitive gains from trade. He argues that standard Pareto productivity assumptions can mechanically eliminate these channels, and shows that using a bounded Pareto distribution restores both product-variety and markup-reduction effects. Using data for the United States, Feenstra finds that product variety and lower markups account for a large share of welfare gains from trade expansion. Rather than relying on oligopolistic head-to-head competition as EMX do, this paper shows that pro-competitive forces can also arise within a heterogeneous-firm monopolistic-competition framework once the productivity distribution is modified.

Similar to us, Hsu, Lu, and Wu (2020) move the analysis away from Taiwan towards the much larger economy of China. They extend the EMX framework to China, using Chinese firm-level census data to quantify gains from trade during a period of rapid trade-cost reductions. They

find total gains from improved openness of 7.1 percent, with pro-competitive effects accounting for about 20 percent of the total gains. Similar to previous work in this literature, they show that head-to-head competition and endogenous markups can meaningfully increase welfare gains relative to standard ACR-type formulas.

Asturias, García-Santana, and Ramos (2019) also study the case of India, although they focus on gains from improvements in transportation infrastructure (the Golden Quadrilateral highway project), not trade shocks. While theirs is a model of internal trade across Indian states, they also exploit variable markups as the source of improvements in allocative efficiency. They calibrate the model to Indian manufacturing and estimate real income gains of 2.7%, with allocative-efficiency gains accounting for 7.4% of total gains and up to 18% in some states. This paper also highlights the importance of studying an economy very different from Taiwan, one in which the policies working to open up the economy were not limited to increased exposure to foreign competition, but included significant reductions in the frictions associated to domestic trade across regions.

Following EMX, Holmes, Hsu, and Lee (2014) develop an index of allocative efficiency that depends on the distribution of markups and separates this channel from standard Ricardian gains from trade, and they decompose how changes in trade frictions affect welfare through cost-change and price-change channels.

Impullitti, Licandro, and Rendahl (2022) study pro-competitive gains from trade in the context of endogenous firm innovation. In a model with heterogeneous firms that compete oligopolistically, a reduction in trade costs, raises market concentration, lowers the equilibrium number of firms and stimulates innovation by increasing the scale of the surviving firms. In their model and differently from EMX, reductions in markup dispersion play only a negligible role. It is concentration and innovation that are the central drivers of gains from trade.

Heid and Stähler (2024) also study the pro-competitive gains from trade but this time in the spirit of the empirical gravity literature. They also conclude that welfare gains from trade liberalization are larger once competition effects are incorporated. They conclude that abolishing the European Single Market would raise markups for EU countries by a range of 2-6%.

3 The Model

The two-tier CES framework for sectors and sub-sectors aggregation (absent persistent trade relationships) in FHM closely resembles those in Atkeson and Burstein (2008) and Edmond et al (2015).

Introducing deep habits implies introducing dynamic aspects into the model, but other than that, our framework closely follows that in FHM. It is a model with N countries, K sectors and J subsectors. We denote the importing (home) country with n and exporting countries with i. In each sector indexed with k there are a total of $J(k)$ subsectors indexed by j.

In each period t, consumers in country n buy a final consumption final good Q_n produced by firms operating in a perfectly competitive industry according to the Cobb-Douglas production function:

$$Q_{n,t} = \prod_{k=1}^{K} Q_{n,t}^{k\alpha_n^k} \tag{1}$$

where Qn^k denotes a composite good in sector k and α_n^k is the expenditure share of sector k in country n.

In each sector k, this composite good is produced using a constant elasticity of substitution (CES) aggregate over subsectors j,

$$Q_{n,t}^k = \left(\sum_{j=1}^{J(k)} \left(q_{n,t}^{j(k)} \right)^{\frac{(\sigma_n^k - 1)}{\sigma_n^k}} \right)^{\frac{\sigma_n^k}{(\sigma_n^k - 1)}} \tag{2}$$

where $q^j(k)_n$ is a composite good in subsector j belonging to sector k in country n, and parameter σ_n^k measures the elasticity of substitution across the sub-sectors of sector k in country n.

From eq. (2), we can obtain the period t's demand for the composite good $q_{n,t}^{j(k)}$ as:

$$q_{n,t}^{j(k)} = \left(\frac{P_{n,t}^{j(k)}}{\mathcal{P}_{n,t}^k} \right)^{-\sigma_n^k} Q_{n,t}^k \tag{3}$$

where $P_n^{j(k)}$ is the price index for subsector $j(k)$ and P_n^k is the CES price index of sector k in the

importing country n and satisfies $\mathcal{P}_{n,t}^{k} = \left[\sum_{j=1}^{J(k)} \left(P_{n}^{j(k)}\right)^{1-\sigma_{n}^{k}}\right]^{\frac{1}{1-\sigma_{n}^{k}}}$.

The aggregate price index in country n is given by:

$$\mathcal{P}_{n,t} = \prod_{k=1}^{K} \left(\frac{P_{n,t}^{k}}{\alpha_{n}^{k}}\right)^{\alpha_{n}^{k}} \tag{4}$$

In turn, $q_{n,t}^{j(k)}$ is a CES aggregator over a finite number of varieties from across the world produced by firm f in each of the exporting countries i. Using the "deep habits" framework of Ravn et al (2006), in which consumers develop habits over individual varieties, in our case, over imports from a particular country in the past. Thus, the more country n imports from country i in period $t-1$, the more it will find it optimal to keep importing from the same partner in period t.

It is important to highlight that the "deep habits" framework is consistent with several other models of markup variation. The advantage of deep habits is that it provides a tractable way of dealing with endogenously varying markups in dynamic, general equilibrium.[2]

The parameter ψ in equation (**??**) governs the degree of "deep habits" in the demand for each variety of intermediates. The benchmark case in EMX and FMH with standard static preferences is embedded in this one for the case of $\psi = 0$. The parameter $\sigma_{n}^{j(k)} > 0$ denotes the intratemporal elasticity of substitution across subsectors in sector j.

The terms $q_{n,t-1}^{j(k)} = \sum_{i=1}^{N} \sum_{f=1}^{F_{in,t}^{j(k)}} q_{inf,t-1}^{j(k)}$ and $y_{t-1}^{F}(s) = \sum_{i=1}^{n(s)} y_{i,t-1}^{F}(s)$ denote the cross-sectional average level of consumption of variety i in period $t-1$, which the firm takes as exogenously given.

$$q_{n,t}^{j(k)} = \left[\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right)^{\frac{\sigma_{n}^{j(k)}-1}{\sigma_{n}^{j(k)}}}\right]^{\frac{\sigma_{n}^{j(k)}}{\sigma_{n}^{j(k)}-1}} \tag{5}$$

As explained in Ravn et al (2006), "deep habits" introduce an intertemporal effect in the

[2]Among others these models are the switching cost IO model of Klemperer (1995), the implicit collusion model of Rotemberg and Saloner (1986) and Rotemberg and Woodford (1992), the composition of aggregate demand model of Gali (1994), and the sluggish product-price adjustment of Yun (1996).

determination of equilibrium markups. This is because firms internalize the fact that, through habits, current pricing affects not only demand in period t, but future demand conditions as well.

Therefore, when the expected present discounted value of lifetime profits is high, firms have an incentive to invest in building a larger customer base today, which will also build the future stock of habits. They do so by lowering current markups and inducing higher current sales. This mechanism is similar to that of partial equilibrium models of customer-market pricing of Phelps and Winter (1970) or switching costs models of Klemperer (1995) in the Industrial Organization literature. In the deep habits model there is gradual substitution between differentiated goods rather than discrete switches among suppliers typical of IO papers. This renders the DH model much more analytically tractable since one does not face an aggregation problem. In equilibrium, buyers can distribute their purchases identically, and still suppliers face a gradual loss of customers if they raise their relative prices. The deep-habit-formation model can therefore be viewed as a natural vehicle for incorporating switching-cost/customer-market models into a dynamic general equilibrium framework (Ravn et al (2006)).

In our case, the "deep habits" framework allows us to study the link between heterogeneity in markups and the pro-competitive gains from trade in a multi-country, multi-sector model.

3.1 Deriving sectorial demand from individual producers at home and abroad

From the optimization problem of country n trying to minimize the cost of importing from a set of firms indexed by f in a set of countries indexed by i, we can obtain country n's demand for each of the varieties from each of the exporting countries in the following way:

$$\min \quad \sum_{i=1}^{N} \sum_{f=1}^{F_{in,t}^{j(k)}} p_{inf,t}^{j(k)} q_{inf,t}^{j(k)} \tag{6}$$

$$s.t.$$

$$q_{n,t}^{j(k)} = \left[\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)} \right)^{\frac{\sigma_n^{j(k)}-1}{\sigma_n^{j(k)}}} \right]^{\frac{\sigma_n^{j(k)}}{\sigma_n^{j(k)}-1}}$$

$$q_{inf,t}^{j(k)} = \left(\frac{p_{inf,t}^{j(k)}}{P_{n,t}^{k}} \right)^{-\sigma_n^{j(k)}} q_{n,t}^{j(k)} + \psi q_{n,t-1}^{j(k)} \tag{7}$$

Rearrange eqs.(7) and (3) for the relative prices, respectively to obtain:

$$\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{k}} = \left(\frac{q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}}{q_{n,t}^{j(k)}} \right)^{-\frac{1}{\sigma_n^{j(k)}}} \tag{8}$$

$$\frac{\mathcal{P}_{n,t}^{k}}{\mathcal{P}_{n,t}^{j(k)}} = \left(\frac{q_{n,t}^{j(k)}}{Q_{n,t}^{k}} \right)^{-\frac{1}{\sigma_n^{k}}} \tag{9}$$

Combining eqs. (8) and (4), we get:

$$p_{inf,t}^{j(k)} = \frac{\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)} \right)^{-\frac{1}{\sigma_n^{j(k)}}}}{q_{n,t}^{j(k)}} \frac{q_{n,t}^{j(k)}{}^{-\frac{1}{\sigma_n^{k}}}}{Q_{n,t}^{k}} \mathcal{P}_{n,t}^{k} \tag{10}$$

From eq. (**??**), we can obtain the price elasticity of demand as:

$$\epsilon_{inf,t}^{j(k)} \equiv - \left(\frac{\partial p_{inf,t}^{j(k)}}{\partial q_{inf,t}^{j(k)}} \frac{q_{inf,t}^{j(k)}}{p_{inf,t}^{j(k)}} \right)^{-1} \tag{11}$$

This derivative is comprised of two terms, which we work on separately below.

The first term is:

$$\begin{aligned}
&= \frac{1}{\sigma_n^{j(k)}} \left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right)^{-\frac{1}{\sigma_n^{j(k)}}-1} P_{n,t} \left(\frac{1}{Q_{n,t}^k}\right)^{-\frac{1}{\sigma_n^k}} \left(q_{n,t}^{j(k)}\right)^{\frac{1}{\sigma_n^{j(k)}}-\frac{1}{\sigma_n^k}} \frac{q_{inf,t}^{j(k)}}{p_{inf,t}^{j(k)}} \\
&= \frac{1}{\sigma_n^{j(k)}} \frac{q_{inf,t}^{j(k)}}{\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right)} \\
&= \frac{1}{\sigma_n^{j(k)}} \frac{q_{inf,t}^{j(k)}}{q_{n,t}^{j(k)}} \left(\frac{p_{inf,t}^{j(k)}}{P_{n,t}^{j(k)}}\right)^{\sigma_n^{j(k)}}
\end{aligned}$$

In the last step, we used $\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right) = \left(\frac{p_{inf,t}^{j(k)}}{P_{n,t}^{j(k)}}\right)^{-\sigma_n^{j(k)}} q_{n,t}^{j(k)}$.

Now note that

$$s_{inf,t}^{j(k)} = \frac{p_{inf,t}^{j(k)} q_{inf,t}^{j(k)}}{p_{n,t}^{j(k)} q_{n,t}^{j(k)}} \tag{12}$$

is the share of country n's imports coming from country i. Then, the first term of the inverse elasticity of demand is

$$\begin{aligned}
&= \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \frac{P_{n,t}^{j(k)}}{p_{inf,t}^{j(k)}} \left(\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}}\right)^{\sigma_n^{j(k)}} \\
&= \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \left(\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}}\right)^{\sigma_n^{j(k)}-1}
\end{aligned}$$

The second term is:

$$\begin{aligned}
&= -\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right)^{-\frac{1}{\sigma_n^{j(k)}}} P_{n,t} \left(\frac{1}{Q_{n,t}^k}\right)^{-\frac{1}{\sigma_n^k}} \left(\frac{1}{\sigma^j(k)_n} - \frac{1}{\sigma_n^k}\right) \left(q_{n,t}^{j(k)}\right)^{\frac{1}{\sigma^j(k)_n}-\frac{1}{\sigma_n^k}-1} \frac{\partial q^j(k)_{n,t}}{\partial q_{inf,t}^{j(k)}} \frac{q_{inf,t}^{j(k)}}{p_{inf,t}^{j(k)}} \\
&= -\left(\frac{1}{\sigma_n^{j(k)}} - \frac{1}{\sigma_n^k}\right) \frac{p_{inf,t}^{j(k)} q_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)} q_{n,t}^{j(k)}} \\
&= -\left(\frac{1}{\sigma_n^{j(k)}} - \frac{1}{\sigma_n^k}\right) s_{inf,t}^{j(k)}
\end{aligned}$$

In the second-to-last step, we used $\frac{\partial q_{n,t}^{j(k)}}{\partial q_{inf,t}^{j(k)}} = \left[\frac{\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)}\right)}{q_{n,t}^{j(k)}}\right]^{-\frac{1}{\sigma_n^{j(k)}}}$.

Therefore, combining both terms, the final expression for the price elasticity of demand is:

$$\begin{aligned}
\epsilon_{inf,t}^{j(k)} &= \left\{ \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \left(\frac{p_{inf,t}^{j(k)}}{P_{n,t}^{j(k)}} \right)^{\sigma_n^{j(k)}-1} - \left(\frac{1}{\sigma^j(k)_n} - \frac{1}{\sigma_n^k} \right) s_{inf,t}^{j(k)} \right\}^{-1} \\
&= \left\{ \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \left[\left(\frac{p_{inf,t}^{j(k)}}{P_{n,t}^{j(k)}} \right)^{\sigma_n^{j(k)}-1} - 1 \right] + \frac{1}{\sigma_n^k} s_{inf,t}^{j(k)} \right\}^{-1} \qquad (13)
\end{aligned}$$

It can be shown that in the absence of persistent trade relationships, i.e. without deep habits in the model and ψ=0, eq. (13) reduces to

$$\epsilon_{inf,t}^{j(k)} = \left\{ \frac{1}{\sigma_n^{j(k)}} (1 - s_{inf,t}^{j(k)}) + \frac{1}{\sigma_n^k} s_{inf,t}^{j(k)} \right\}^{-1} \qquad (14)$$

which is the expression in FHM (see Appendix A.1 for this derivation).

3.2 Production and Trade Frictions

Firm f in country i in sub-sector $j(k)$ produces using only labor according to a constant returns-to-scale (CRS) technology

$$q_{if}^{j(k)} = A_{if}^{j(k)} l_{i,f} \qquad (15)$$

in which the sub-sector, firm, and country-specific productivity is denoted with $A_{i,f}^{j(k)}$ and it is assumed to be constant over time.

Producers from country i need to pay an ad-valorem tariff $t_{in,t}^{j(k)}$ to be able to enter country n. Notice that the tariff is sector-specific, and will change over time when we introduce the trade policies that we want to study. There is also an iceberg trade cost to deliver a unit of good $j(k)$ from country i to country n that we denote with $d_{in}^{j(k)}$ which can also be sector-specific but is constant over time. As standard in the trade literature, these costs are intended to capture the well-known gravity variables, such as geographic distance, the presence of common borders and

language, etc.

Then, total trade frictions in each period t can be summarized as:

$$\tau_{in,t}^{j(k)} = d_{in}^{j(k)} \left(1 + t_{in,t}^{j(k)}\right) \tag{16}$$

where $\tau_{ii}^{j(k)} = 1 \quad \forall t$

Marginal costs for these producers are given by $\frac{w_{i,t}\tau_{in,t}^{j(k)}}{A_{i,f}^{j(k)}}$ so that the pricing choice by Cournot competitors is given by:

$$\max_{p_{inf}^{j(k)}, q_{inf}^{j(k)}} p_{inf}^{j(k)} q_{inf}^{j(k)} - q_{inf}^{j(k)} \left[\frac{w_{i,t}\tau_{in,t}^{j(k)}}{A_{i,f}^{j(k)}}\right] \tag{17}$$

subject to the demand equation (7).

The solution is

$$p_{inf,t}^{j(k)} = \mu_{inf,t}^{j(k)} \frac{w_{i,t}\tau_{in,t}^{j(k)}}{A_{i,f}^{j(k)}} \tag{18}$$

where $\mu_{inf,t}^{j(k)}$ is the markup charged by firm f of country i when selling to sub-sector $j(k)$ in country n at time t and satisfies the following condition with respect to the price elasticity of demand in equation (11).

$$\mu_{inf,t}^{j(k)} = \frac{\epsilon_{inf,t}^{j(k)}}{\epsilon_{inf,t}^{j(k)} - 1} \tag{19}$$

3.3 Aggregate Expenditure

Consumers in country n spend a total amount $I_{n,t}$ in each period t. Of that total, a fraction α_n^k of their total expenditure on sector k, and their aggregate expenditure on sub-sector $j(k)$ is

$$X_{n,t}^{j(k)} = \alpha_n^k I_{n,t} \left(\frac{P_{n,t}^{j(k)}}{\mathcal{P}_{n,t}^k}\right)^{1-\sigma_n^k} \tag{20}$$

Total expenditure $I_{n,t}$ is given by the sum of workers' labor income ($w_{n,t}L_{n,t}$), firms' profits ($Y_{n,t}$), tariff revenue for the government ($R_{n,t}$) and trade deficits ($D_{n,t}$)

$$I_{n,t} = w_{n,t}L_{n,t} + Y_{n,t} + R_{n,t} + D_{n,t} \tag{21}$$

Firms' profits can be written using the optimal pricing equation as:

$$Y_{n,t} = \sum_{k=1}^{K}\sum_{j=1}^{J}(k)\sum_{i=1}^{N}\sum_{f=1}^{F_{ni,t}^{j(k)}} \frac{1}{\epsilon_{nif,t}^{j(k)}} \frac{s_{nif,t}^{j(k)} X_{i,t}^{j(k)}}{(1+t_{ni,t}^{j(k)})} \tag{22}$$

Total tariff revenue earned by the government is

$$R_{n,t} = \sum_{k=1}^{K}\sum_{j=1}^{J}(k)\sum_{i=1}^{N}\sum_{f=1}^{F_{in,t}^{j(k)}} t_{inf,t}^{j(k)} \frac{s_{inf,t}^{j(k)} X_{n,t}^{j(k)}}{(1+t_{in,t}^{j(k)})} \tag{23}$$

Trade deficits defined as total imports minus total exports are give by

$$D_{n,t} = \sum_{k=1}^{K}\sum_{j=1}^{J(k)}\sum_{i=1}^{N}\sum_{f=1}^{F_{in,t}^{j(k)}} \frac{s_{inf,t}^{j(k)} X_{n,t}^{j(k)}}{(1+t_{in,t}^{j(k)})} - \sum_{k=1}^{K}\sum_{j=1}^{J(k)}\sum_{i=1}^{N}\sum_{f=1}^{F_{ni,t}^{j(k)}} \frac{s_{nif,t}^{j(k)} X_{i,t}^{j(k)}}{(1+t_{ni,t}^{j(k)})} \tag{24}$$

From these equations, labor market clearing implies that total revenue of country n's firms across the world net of profits should equal total labor costs by these firms, so that

$$w_{n,t}L_{n,t} = \sum_{k=1}^{K}\sum_{j=1}^{J(k)}\sum_{i=1}^{N}\sum_{f=1}^{F_{ni,t}^{j(k)}} \frac{(\epsilon_{nif,t}^{j(k)}-1)}{\epsilon_{nif,t}^{j(k)}} \frac{s_{nif,t}^{j(k)} X_{i,t}^{j(k)}}{(1+t_{ni,t}^{j(k)})} \tag{25}$$

Exports are given by:

$$X^{j(k)} = \sum_{k=1}^{K}\sum_{j=1}^{J(k)}\sum_{i=1}^{N}\sum_{f=1}^{F_{ni,t}^{j(k)}} \frac{s_{nif,t}^{j(k)} X_{i,t}^{j(k)}}{(1+t_{ni,t}^{j(k)})} \tag{26}$$

Imports are given by:

$$IM^{j(k)} = \sum_{k=1}^{K}\sum_{j=1}^{J(k)}\sum_{i=1}^{N}\sum_{f=1}^{F_{in,t}^{j(k)}} \frac{s_{inf,t}^{j(k)} X_{n,t}^{j(k)}}{(1+t_{in,t}^{j(k)})} \tag{27}$$

Then, the sector-level share of spending on domestic goods $\lambda^{j(k)}$ is $\lambda^{j(k)} = \frac{q^{j(k)}-X^{j(k)}}{q^{j(k)}-X^{j(k)}+IM^{j(k)}}$.

Then, the aggregate share of spending on domestic goods is $\lambda = \int_0^1 \lambda(s)ds$ and the index of import share dispersion is $\sigma(\lambda) = \frac{Var(\lambda)}{\lambda(1-\lambda)}$

3.4 Welfare

Aggregate real expenditure in the home country n is

$$Q_{n,t} = \frac{I_{n,t}}{\mathcal{P}_{n,t}} \tag{28}$$

which is the consumers' utility in period t, with aggregate welfare W_n given by

$$W_n = \sum_{t=1}^{\infty} \beta^{t-1} Q_{n,t} \tag{29}$$

3.5 Aggregate Productivity and Markups

Defining aggregate productivity in country i as A_i which satisfies

$$Q_{i,t} = A_{i,t} L_i \tag{30}$$

and it can be shown that A_i is the quantity-weighted harmonic average of firm-level productivities as in

$$A_{i,t} = \left[\sum_{k=1}^{K} \sum_{j=1}^{J(k)} \left(\sum_{f=1}^{F_{ii,t}^{j(k)}} \frac{1}{A_{i,f}^{j(k)}} \frac{q_{iif,t}^{j(k)}}{Q_{i,t}} + \sum_{n \neq i} d_{in}^{j(k)} \sum_{f=1}^{F_{in,t}^{j(k)}} \frac{1}{A_{i,f}^{j(k)}} \frac{q_{inf,t}^{j(k)}}{Q_{i,t}} \right) \right]^{-1} \tag{31}$$

Notice that even though idiosyncratic firm-sector-level productivities are constant over time, aggregate productivity can vary over time due to the weights/trade shares being time-dependent because they are policy dependent.

Similarly, the aggregate markup $\mu_{i,t}$ in country i and period t is defined as

$$\mu_{i,t} = \frac{\mathcal{P}_{i,t}}{\left(\frac{w_{i,t}}{A_{i,t}}\right)} \tag{32}$$

It can be calculated as a harmonic average

$$\mu_{i,t} = \left[\sum_{k=1}^{K} \sum_{j=1}^{J(k)} \left(\sum_{f=1}^{F_{ii,t}^{j(k)}} \frac{1}{\mu_{iif,t}^{j(k)}} \frac{p_{iif,t}^{j(k)} q_{iif,t}^{j(k)}}{\mathcal{P}_{i,t} Q_{i,t}} + \sum_{n \neq i} \frac{d_{in}^{j(k)}}{\tau_{in,t}^{j(k)}} \sum_{f=1}^{F_{in,t}^{j(k)}} \frac{1}{\mu_{inf,t}^{j(k)}} \frac{p_{inf,t}^{j(k)} q_{inf,t}^{j(k)}}{\mathcal{P}_{i,t} Q_{i,t}} \right) \right]^{-1} \tag{33}$$

3.6 Misallocation

As in FHM, we define misallocation as the percentage difference between aggregate productivity $A_{i,t}$ and efficient productivity $A_{i,t}^{eff}$ which is in turn defined as the aggregate productivity associated with marginal-cost pricing[3].

So, misallocation is calculated as

$$\left(1 - \frac{A_{i,t}}{A_{i,t}^{eff}} \right) \tag{34}$$

Another measure of misallocation, as in EMX, is

$$\frac{Var(\lambda)}{\lambda(1-\lambda)} \tag{35}$$

where is the cross-sectional average of domestic market shares across sectors $j(k)$.

4 Solution Strategy

To solve the model, we follow the strategy in FHM. In this section we summarize this method.

Given productivities $A_{if}^{j(k)}$ (estimated from the data on India), elasticities of substitution within and across sectors $\sigma_n^{j(k)}$ and σ_n^k (calibrated), Cobb-Douglas shares α_n^k (calibrated), labor endowments L_n, trade deficits D_n, iceberg trade costs $d_{in}^{j(k)}$ (estimated from the data on India's main trading partners), number of firms $F_{ni}^{j(k)}$, and ad-valorem tariffs $t_{in}^{j(k)}$, an equilibrium is characterized by a vector of wages that satisfy the equilibrium conditions (4), (8), and (11)-(19).

To solve for the equilibrium we use the same strategy as FHM in the following steps:

[3]This would prevail if all firms across the world set prices equal to the marginal cost of serving each destination market according to $p_{inf,t}^{j(k)} = \frac{w_{i,t} \tau_{in,t}^{j(k)}}{A_{i,f}^{j(k)}}$

1. We start with a guess on the vector of wages for all countries $\{w_n\}_{n=1}^N$.

2. Using our guess on wages and information on the number of firms in each market, iceberg trade costs d, tariffs, and productivities, we calculate the marginal cost of firm f from country i serving country n in sub-sector $j(k)$.

3. We solve the system of equations (9) and (8), (19), (13) and (12) and compute firms' market shares, prices, markups, and demand elasticities.

4. From the prices computed above, we use equation (**??**) to calculate sectoral price indices $\mathcal{P}_n^k$.

5. Using equations (20), (21), (22) and (23), we can compute aggregate firm profits, tariff revenues, sectoral expenditures, and total expenditures.

6. We next check the trade deficit equation (24) and update our guess for wages until this equation is satisfied.

7. Both before and after the introduction of trade policy, We focus on the steady-state of the model so that $q_{n,t}^{j(k)} = q_{n,t-1}^{j(k)}$.

8. To evaluate the gains from trade and the effects of raising tariffs, we repeat the above steps for a counterfactual set of tariffs. We then calculate and report percentage changes in profits, expenditures, prices, and welfare.

5 The Data and Estimation

We use establishment-level data from the Indian Annual Survey of Industries (ASI). The ASI is the principal source of industrial statistics in India and is conducted by the Ministry of Statistics and Programme Implementation (MoSPI). The dataset covers nearly the entire universe of establishments in India's organized manufacturing sector. Specifically, the ASI includes all establishments employing 10 or more workers with power and those employing 20 or more workers without power.

Our sample spans the years 2009-2013. In each year, the survey collects data from approximately 40,000 establishments, which can be linked across survey waves with a unique identifier. For each establishment, the survey provides detailed information on production, including the value of gross output, employment, and expenditures on various inputs. In addition, the ASI reports the value of imported inputs consumed and the value of exports for each establishment. We deflate all nominal variables using industry-specific price deflators.

In this section, we present key moments about the producer concentration and import and export statistics from the ASI data. We will use these results to calibrate the model in the simulation exercises.

Table 1 presents measures of producer concentration within 4-digit industrial sectors. The mean and median inverse Herfindahl index (HHI) are 22 and 13, respectively. On average, the largest producer has a market share of around 26 percent. The distribution of market shares is strongly skewed to the right. The mean producer's market share is 0.008, which is substantially larger than the median share of 0.001. The top 1 percent of producers account for over 40% of total sales. Overall, these patterns suggest a relatively high degree of concentration within sectors.

Table 1: Within-Sector Moments in the Data

Within-sector concentration	
Mean inverse HHI	22.0
Median inverse HHI	13.1
Mean top share	0.26
Median top share	0.18
Distribution of sectoral shares	
Mean share	0.008
Median share	0.001
p75 share	0.005
p95 share	0.034
p99 share	0.117
SD share	0.036
Size distribution of producers	
Fraction of sales by top 0.01 producers	0.42
Fraction of sales by top 0.05 producers	0.65
Fraction of wages by top 0.01 producers	0.25
Fraction of wages by top 0.05 producers	0.45

Table 2 reports statistics about producer concentration across sectors. The median inverse HHI across sectors is 13, and the median number of producers per sector is 63. The top 1 percent of sectors accounts for about 20 percent of total sales, while the top 5 percent accounts for about 34 percent of total labor payments.[4]

Table 3 provides several data moments related to trade. About 18 percent of manufacturing establishments are exporters, and the average import share is 21 percent. The Grubel and Lloyd (1971) index of intra-industry trade is 0.41. The index of import share dispersion, calculated using equation (35), is 0.40.

[4]Labor payments include wage, bonus, and contribution to provident and other funds

Table 2: Across-Sector Moments in the Data

Across-sector concentration	
p10 inverse HHI	2.30
p50 inverse HHI	13.1
p90 inverse HHI	48.1
p10 top share	0.07
p50 top share	0.18
p90 top share	0.60
p10 number producers	10
p50 number producers	63
p90 number producers	269
Size distribution of sectors	
Fraction sales by top 0.01 sectors	0.20
Fraction sales by top 0.05 sectors	0.47
Fraction wages by top 0.01 sectors	0.05
Fraction wages by top 0.05 sectors	0.34

Table 3: Import and Export Statistics in the Data

Aggregate fraction exporters	0.18
Aggregate import share	0.21
Index intra-industry trade	0.41
Index import share dispersion	0.40

6 Results and Discussion (in progress)

This draft is still very much work in progress and we have no results to show yet. Therefore, in this section we provide a list of the exercises we are running and the results we will present. We also discuss the intuition behind our model and place emphasis on how we expect the introduction of persistent importer-exporter relationships to alter the results relative to EMX(2015) and FHM (2025).

6.1 A Trade Liberalization Episode

We will start by studying an episode of trade liberalization, implemented through a reduction in tariffs $\tau_{in}^{j(k)}$. Our results will show an immediate reduction in the market share for the home producers of intermediates $s_{nn}^{j(k)}$. Through eq. (13), it can be seen that a lower domestic share results in an increase in the elasticity of demand $\epsilon_{nn}^{j(k)}$ faced by these producers. Through this impact on the elasticity of demand, and via the obvious inverse relationship between elasticity and markups, markups end up falling.

A reduction in markups (the highest ones charged by the less competitive domestic producers who are priced-out of the market first after trade liberalization) lowers markup dispersion. As in Atkeson and Burstein (2008), markup dispersion implies that resources are misallocated, and the larger the dispersion, the higher the degree of misallocation, and the lower aggregate productivity. By increasing competitive pressures on domestic producers and lowering markups, trade reduces the cross-sector variation in markups, reduces misallocation and raises aggregate productivity. Implicit in this is the assumption that $\sigma_n^{j(k)} > \sigma_n^k$, i.e. that the lower-tier elasticity is larger than the upper-tier because subsectors are more substitutable among each other than sectors.

In our setup, a reduction in trade costs τ brings in an additional dynamic effect. Now, through the persistent trade relationships story (modeled through the introduction of "deep habits"), the price elasticity of demand and markups are endogenous in not only market shares as in EMX and FMH, but also in the quantity of imports of intermediates by each firm f in sub-sector $j(k)$ in country n from country i. Thus, trade liberalization or pro-trade policies that raise the value of imports introduce a second channel through which elasticity rises, markups fall as does markup dispersion and the degree of misallocation, i.e. a second way in which liberalizing trade induces an increase in aggregate productivity and pro-competitive gains.

EMX conclude that trade can significantly reduce these markup distortions but only under two conditions: (1) there must be extensive misallocation across sectors, and (2) trade liberalization is such that competitive pressures on domestic producers rise significantly and their market shares respond accordingly. By introducing a second channel through which markups

are endogenous to the volume of trade, we believe our model would relax the second condition, and would therefore yield pro-competitive gains from trade even in environments in which trade shares do not react so starkingly to changes in trade costs.

We believe that we also contribute empirically. While EMX focus on the case of Taiwan, a small open economy, we work with highly disaggregated data for India, a much larger and diverse economy. Taiwan and India differ drastically from each other in the sectoral structural composition of their economies. Importantly, while EMX conclude that these conditions are met in Taiwan, we are in the process of testing whether these conditions also hold for a different small open economy, India. We are doing so using firm-level data for India (much smaller level of disaggregation - data publicly available only at the ISIC-4 level), and we are measuring the markup dispersion index $Var(\lambda)$ as well as the the cross-country correlation in productivity ρ. These indicators are important because to obtain pro-competitive gains from trade, the results in EMX (2015) require that ρ be rather high. In particular, results from trade switch from losses to gains only for $\rho \geq 0.4$. Since $Var(\lambda)$ and ρ are inversely related to each other, for an economy in which the dispersion of markups is larger than in Taiwan, the degree of cross-country correlation in productivity required for trade liberalization to be welfare-improving would be lower. More generally, gains from trade would be observed even in cases in which productivity is not so closely linked across countries and therefore, opening to trade does not necessarily expose producers to genuine *head-to-head* competition.

6.2 A Global Trade War

We will aim to replicate the exercise in FHM in which all countries raise tariffs by 50 percentage points. We will study the differences in the transmission mechanisms of retaliation when importers are exporters are "locked-in" their trade relationships.

6.3 Understanding the Role of Cross-Country Heterogeneity of Markups

We will compare the results of the baseline model with another one in which both the lower- and upper-tier elasticities $\sigma_n^{j(k)}$ and σ_n^k are assumed to be equal across countries.

6.4 Welfare Analysis

In all exercises we will assess the welfare gains from trade by comparing welfare between our model and one with fixed markups. We will also compare the increases in wages and profits in the two models.

References

[1] Roger Aliaga-Díaz and María Pía Olivero, "Macroeconomic Implications of 'Deep Habits' in Banking," *Journal of Money, Credit and Banking*, Vol. 42, No. 8, December 2010, pp. 1495–1521.

[2] Arkolakis, C., A. Costinot, D. Donaldson and A. Rodriguez-Clare (2019), "The Elusive Pro-Competitive Effects of Trade", *Review of Economic Studies*, Vol. 86, pp. 46-80.

[3] Asturias, José, Manuel García-Santana, and Roberto Ramos (2019), "Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India," *Journal of the European Economic Association*, Vol. 17(6), pp. 1881–1940.

[4] Atkeson, A. and A. Burstein (2008), "Pricing-to-Market, Trade Costs, and International Relative Prices", *American Economic Review*, Vol. 98(5), pp. 1998-2031.

[5] Bloch, H. and M. Olive (2001), "Pricing Over the Cycle", *Review of Industrial Organization*, Vol. 19, pp. 99-108.

[6] Chevalier, J. and D. Scharfstein (1995), "Liquidity Constraints and the Cyclical Behavior of Markups", *The American Economic Review - Papers and Proceedings*, Vol. 85(2), May 1995, pp. 390-396.

[7] Chevalier, J. and D. Scharfstein (1996), "Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence", *The American Economic Review*, Vol. 86(4), September 1996, 703-725.

[8] Campello, M. (2003), "Capital Structure and Product Markets Interactions: Evidence from Business Cycles", *Journal of Financial Economics*, Vol. 68, pp. 353-378.

[9] Domowitz, I., G. Hubbard and B. Petersen (1986), "Business Cycles and the Relationship Between Concentration and Price-Cost Margins", *Rand Journal of Economics*, Spring 1986, 17, pp. 1-17.

[10] Feenstra, Robert (2018), "Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity," *Journal of International Economics*, Vol. 110, 2018, pp. 16–27.

[11] Firooz, H. (2024), "The Pro-Competitive Consequences of Trade in Frictional Labor Markets", working paper.

[12] Firooz, H., G. Heins and S. Mathur (2024), "Markup Heterogeneity, Openness, and the Pro-Competitive Gains from Trade Across Countries", working paper.

[13] Galeotti, M. and F. Schiantarelli (1998), "The Cyclicality of Markups in a Model with Adjustment Costs: Econometric Evidence for US Industry", Oxford Bulletin of Economics and Statistics, Vol. 60 (2), 121-142.

[14] Galí, Jordi, "Monopolistic Competition, Business Cycles, and the Composition of Aggregate Demand," *Journal of Economic Theory* 63, 1994, 73-96.

[15] Heid, Benedikt and Frank Stähler (2024), "Structural Gravity and the Gains from Trade under Imperfect Competition," *Economic Modelling*, Vol. 131.

[16] Holmes, Thomas Wen-Tai Hsu, and Sanghoon Lee (2014), "Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade,", *Journal of International Economics*, Vol. 94(2), pp. 195–206.

[17] Hsu, Wen-Tai, Yi Lu, and Guiying Wu (2020), "Competition, Markups, and Gains from Trade: A Quantitative Analysis of China Between 1995 and 2004," *Journal of International Economics*, Vol. 122.

[18] Impullitti, Giammario, Omar Licandro, and Pontus Rendahl (2022), "Technology, Market Structure, and the Gains from Trade," *Journal of International Economics*, Vol. 135.

[19] Klemperer, Paul, "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies* 62, 1995, 515-539.

[20] Lebow, D. (1992), "Imperfect Competition and Business Cycles: An Empirical Investigation", *Economic Inquiry*, Vol. 30(1), January 1992, pp. 177-193.

[21] Ravn, M, S. Schmitt-Grohé and M. Uribe, "Deep Habits", *Review of Economic Studies* 73, 2006, 195-218.

[22] Rotemberg, Julio J., and Garth Saloner (1986), "A Supergame-Theoretic Model of Price Wars During Booms," *American Economic Review* 76, June, 390-407.

[23] Rotemberg, Julio J. and Michael Woodford (1992), "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy*, Vol. 100, December, 1153-1207.

Appendix A: Additional Model Derivations

In this appendix we show that the expression for the elasticity of demand in our model has the one in FHM embedded in it. In the case of no deep habits, i.e. ψ=0, our model boils down to FHM.

The term

$$\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}}$$

in the expression for the elasticity of demand eq. (13) is

$$\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}} = \left[\frac{\left(q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)} \right)}{q_{n,t}^{j(k)}} \right]^{-\frac{1}{\sigma_n^{j(k)}}}$$

so that, plugging this expression in eq. (13), the elasticity of demand becomes:

$$\epsilon_{inf,t}^{j(k)} = \left\{ \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \left[\frac{\left(\psi q_{inf,t}^{j(k)} - \psi q_{n,t-1}^{j(k)} \right)}{q_{n,t}^{j(k)}} \right]^{\frac{1-\sigma_n^{j(k)}}{\sigma_n^{j(k)}}} - \left(\frac{1}{\sigma_n^{j(k)}} - \frac{1}{\sigma_n^{k}} \right) s_{inf,t}^{j(k)} \right\}^{-1}$$

If ψ=0, the above expression becomes

$$\begin{aligned}
\left(\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}} \right)^{\sigma_n^{j(k)}-1} &= \left(\frac{q_{inf,t}^{j(k)}}{q_{n,t}^{j(k)}} \right)^{\frac{(1-\sigma_n^{j(k)})}{\sigma_n^{j(k)}}} \\
&= \left(\frac{q_{inf,t}^{j(k)}}{q_{n,t}^{j(k)}} \right)^{\frac{1}{\sigma_n^{j(k)}}} \frac{q_{n,t}^{j(k)}}{q_{inf,t}^{j(k)}} \\
&= \left(\frac{p_{inf,t}^{j(k)}}{\mathcal{P}_{n,t}^{j(k)}} \right)^{-1} \frac{q_{n,t}^{j(k)}}{q_{inf,t}^{j(k)}} \\
&= \frac{1}{s_{inf,t}^{j(k)}}
\end{aligned}$$

Then, plugging the above expression in the elasticity of demand:

$$\begin{aligned}
\epsilon_{inf,t}^{j(k)} &= \left\{ \frac{1}{\sigma_n^{j(k)}} s_{inf,t}^{j(k)} \frac{1}{s_{inf,t}^{j(k)}} - \left(\frac{1}{\sigma^j(k)_n} - \frac{1}{\sigma_n^k} \right) s_{inf,t}^{j(k)} \right\}^{-1} \\
&= \left\{ \frac{1}{\sigma_n^{j(k)}} - \left(\frac{1}{\sigma^j(k)_n} - \frac{1}{\sigma_n^k} \right) s_{inf,t}^{j(k)} \right\}^{-1} \\
&= \left\{ \frac{1}{\sigma_n^{j(k)}} (1 - s_{inf,t}^{j(k)}) + \frac{1}{\sigma_n^k} s_{inf,t}^{j(k)} \right\}^{-1}
\end{aligned}$$