IIMB-WP N0. 734/2026

WORKING PAPER NO: 734

Firm Dynamics, Geographic Expansion and Local Competition

Anubha Agarwal
Assistant Professor
Economics Area
Indian Institute of Management Bangalore
anubha.agarwal@iimb.ac.in

Year of Publication – March 2026

Firm Dynamics, Geographic Expansion and Local Competition*

Anubha Agarwal†

Indian Institute of Management Bangalore

March 27, 2026

Abstract

The importance of competition in local geographic markets has increased as the consumption of non-tradable goods and services has risen compared to tradable ones. Leveraging confidential matched employer-employee microdata and detailed geographic information on firms and workers, I document new evidence on the rise in firms' geographic expansion and local competition in Canada between 2001-2018. To account for these trends, I build a new dynamic general equilibrium model of firms' geographic expansion decisions incorporating multiple markets and Cournot oligopolistic competition in local markets. I estimate the model for two periods and find that three factors are crucial to match the observed trends: (1) increased innovation costs for entrepreneurs, (2) a compositional shift toward less productive and less expansion-efficient entrants, and (3) greater product differentiation between local varieties. The decentralized equilibrium is inefficient due to various externalities associated with firms' geographic expansion. This supports the need for taxing or subsidizing firms' geographic expansion costs to increase welfare. I find that subsidizing the geographic expansion of more productive and more expansion-efficient firms can substantially increase efficiency and social welfare.

Keywords: firm dynamics, geographic expansion, local market power, firm growth
JEL Classification: E20, L10, O30

*I am extremely grateful to Murat Celik, Diego Restuccia and Geuorgui Kambourov for their invaluable guidance on this project. Special thanks to Laurent Cavenaile and Mark Rempel for very helpful discussions and feedback. I thank Ufuk Akcigit, Michelle Alexopoulos, Jason Choi, Margarida Duarte, Sebastian Dyrda, Michael Ehrmann, Ying Feng, Ruben Gaetani, Sampreet Goraya, Fatih Guvenen, Stephan Heblich, Guangbin Hong, Burhanettin Kuruscu, Etienne Lalé, Rasmus Lentz, Virgiliu Midrigan, Jordi Mondria, Peter Morrow, Maarten De Ridder, João Ritto, Baxter Robinson, Joseph Steinberg, Eugene Tan, Nicolas Vincent, Alexander Whalley, Ronald Wolthoff, Ming Xu and Yingnan Zhao for insightful comments. I also thank participants of the Banff Innovation, Productivity, and Growth Conference (2024) and MacDev reading group for helpful suggestions. I gratefully acknowledge the Ph.D. students research grant from the Department of Economics, University of Toronto to access data for this project. The empirical analysis for this paper has been conducted at the Statistics Canada Research Data Centre (RDC) in Toronto. All remaining errors are my own.

†Indian Institute of Management Bangalore, Email: anubha.agarwal@iimb.ac.in

1 Introduction

Firms are central to modern market economies, influencing employment, wages, innovation and competition. They can grow by expanding into various product markets and geographic markets. While research on firm dynamics has extensively covered growth within product markets, less attention has been given to firm growth through geographic expansion, especially at local levels such as provinces and cities. The importance of firms' presence in local geographic markets has increased over time, as firms in the non-tradable sectors command an increasingly larger share of sales and employment. These firms, by their very nature, cannot grow forever by simply improving their market share within a certain geography. Instead, they must reach a larger number of new consumers by selling their goods and services in multiple geographic markets. In addition to being a crucial margin of firm growth, firms' geographic expansion is closely tied to measures of product market competition, which are important determinants of innovation, growth, and antitrust policy.

Research related to firms' geographic expansion is crucial in understanding the modern economy, but its general equilibrium consequences have been underexamined. This is because the dearth of suitable data and appropriate theoretical frameworks make this a challenging task. To tackle these challenges, I leverage confidential microdata and a new dynamic general equilibrium model to address three research questions related to firms' geographic expansion in this paper.

First, I analyze whether there has been any change in the trends related to geographic expansion of firms and local product market competition in non-tradable sectors[1] in Canada between 2001-2018. Second, I study the sources and aggregate implications of this change in firms' geographic expansion. Third, given the benefits and costs associated with firms' geographic expansion (discussed below), I ask whether a policy maker should tax or subsidize firms' geographic expansion to increase welfare, and whether this tax or subsidy should be targeted or uniform.

To answer these questions, I proceed in two steps. First, I use confidential administrative

[1]In particular, I am interested in firms' number of geographic markets within a country. See section 2 for details on the sectors included in this analysis.

microdata on Canadian firms (2001-2018). Leveraging the matched employer-employee feature of the data and a novel empirical methodology, I document that the average firm in non-tradable sectors in Canada has increased its geographic presence from 2.4 markets in 2001 to 4.5 markets in 2018.[2] These averages mask a lot of heterogeneity, since I find substantial differences in the number of geographic markets for firms in Canada. I also document evidence for a 17.1% fall in the sales-weighted average local Herfindahl–Hirschman Index (HHI) in non-tradable sectors from 0.35 in 2001 to 0.29 in 2018, and a 11.8% increase in the sales-weighted net average markup for firms in non-tradable sectors from 0.51 in 2001 to 0.57 in 2018 in Canada.[3]

Second, I build a new rich dynamic general equilibrium model to shed light on the sources and aggregate implications of the change in firms' geographic expansion. I draw on the innovation literature and model firms as entities that can enter (or exit) multiple geographic markets over their lifetime (Klette and Kortum (2004)). In turn, the geographic markets feature oligopolistic competition in quantities (Cournot) between firms (Atkeson and Burstein (2008) and Cavenaile, Celik, and Tian (2019)). At the root of the model is a Poisson process for firms' geographic expansion with an arrival rate that is a function of the number of existing markets of the firm. Heterogeneity in firm types and the stochastic process for the geographic expansion technology leads to heterogeneity in the firms' geographic presence over time.

In this framework, a firm adds new markets by spending resources to potentially expand to new geographic markets, but the likelihood of success of entering a new market depends on the number of markets it already has. Once a firm enters a new market, it changes the number of competitors in that market, hence a firm's geographic expansion comes at the expense of its competitors. A firm can be driven out of a specific market or out of all its markets when hit by exogenous shocks. There are three margins of heterogeneity in the model. First, firms differ in their time- and market-invariant, firm-level productivities, i.e., some firms are always more productive in all the geographic markets they serve. Second, firms differ in their geographic

[2]Canada is made up of 10 provinces and 3 territories, which are further divided into 293 census divisions (CDs) and 5162 census subdivisions (CSDs) according to the 2016 census. For example, the City of Toronto is a CD which consists of multiple CSDs such as Toronto, Mississauga, Markham and Vaughan. My measure of a geographic market for this analysis is a census subdivision (CSD).

[3]I find that the simultaneous increase in markup and fall in local HHI is related to the rising productivity of firms expanding into more markets over time.

expansion efficiencies, i.e., some firms are better at expanding into new geographic markets. Third, geographic markets differ in size, e.g., some markets might have more consumers than others, so it is more profitable for firms to enter such markets. These margins of heterogeneity deliver realistic model predictions and help to match the heterogeneity in firms' geographic expansion observed in the data.

Changes in firms' observed geographic expansion patterns can be attributed to a multitude of factors such as geographic expansion costs, entrepreneurs' innovation costs, and firm productivities. Since some of these mechanisms cannot be directly observed in the data, I use the model as a measurement device to infer changes in these economic mechanisms over time. I estimate the model for two periods: 2002-2005 (early period) and 2013-2016 (late period) and use the two estimated stationary equilibria to disentangle the sources of the observed trends in geographic expansion and local competition in Canada over this period. Comparing the estimated early period economy with counterfactual economies where some parameters are set to their late period values (keeping the remaining parameters at their early period values) allows me to individually disentangle the effects of different mechanisms over time.

In the model, the geographic expansion costs for firms are characterized by scale parameters and a convexity parameter. The estimations reveal that the firms' expansion cost scale parameters show a decline (signifying a fall in geographic expansion costs), but the expansion cost convexity parameter also features a decline (signifying an increase in firms' geographic expansion costs). The first set of counterfactuals suggest that the change in the costs of geographic expansion - intuitively the most likely candidate - is unable to explain the rise in the average number of markets by itself. In particular, it underexplains the increase in the geographic presence of high-type expansion efficiency firms,[4] and overexplains the same for low-type expansion efficiency firms. The change in costs explains about 40% of the decline in local HHI, but only 2% of the increase in average markups.

[4]Due to substantial differences in the number of geographic markets for firms in the data, I divide the firms into high-type and low-type productivity, and high-type and low-type expansion efficiency firms. This division is done for 'large' firms, after removing the 'fringe' firms, i.e., firms that are never present in more than one local market in the data. More details can be found in appendices 8.1 and 8.6.1.

It turns out that there are three key drivers for the increase in the geographic expansion of firms and the fall in local product market concentration in Canada: increased innovation costs for entrepreneurs, compositional shift in entrants and greater product differentiation between local varieties. The increase in innovation costs for entrepreneurs reduces both firm entry and competition from new entrants, creating a stronger incentive for firms to expand geographically. Similarly, the shift in the composition of new entrants toward less productive and less expansion-efficient firms further weakens prospective competition in local markets, prompting firms to increase their geographic expansion and intensify local competition. Greater product differentiation between local varieties enhances firms' market power in local markets, providing a strong incentive for them to expand their presence in these markets to capture higher profits.

Firms' geographic expansion has both benefits and costs for the economy in the model. Geographic expansion of firms provides the opportunity to improve geographic markets' average productivity by changing the composition of firms in those markets. More firms are preferable due to the households' love for variety and an increase in geographic expansion of firms can increase local competition, reduce markups and increase the labor share within local markets. All the aforementioned aspects are benefits of geographic expansion, but firms do not internalize these benefits since they only consider the gain in their value when making their expansion decisions. Hence, the policy maker would prefer to make geographic expansion cheaper to encourage firms to expand more. In addition to having these potential benefits, geographic expansion of firms has costs in terms of using the economy's resources that might have been spent on other uses instead. Firms do not internalize the negative effects of their expansion on the profits of other firms. This 'business stealing externality' leads to incentives for over-expansion in the economy. At the same time, more productive firms do not consider the positive effects of their expansion on the economy - they use less resources to expand, but their expansion provides benefits to the economy beyond the increase in their profits, in terms of lower markups and higher aggregate productivities in local markets. The presence of these externalities renders the decentralized equilibria inefficient, creating an opportunity for policy to improve welfare by using taxes and subsidies on firms' geographic expansion costs to influence their expansion decisions.

I find the set of taxes and subsidies on firms' geographic expansion costs that a policy maker with the goal of maximizing welfare should implement. Using my late period estimated stationary equilibrium, I find that the policy maker can increase welfare by subsidizing the geographic expansion costs of the high-type productivity and high-type expansion efficiency firms at 17.14%, and taxing the geographic expansion costs of all other firms.[5] This increases welfare (equivalent to consumption) by 1.42%, and total output by 3.09%, signifying that the benefits from the most productive firms' geographic expansion outweigh the costs to the economy. Since observing firm types in the real world can be problematic, I also consider uniform taxes and subsidies. The constrained-optimal subsidy rate is found to be 18.18%, leading to a 0.45% increase in welfare and 2.41% increase in total output - less than the welfare and output gain observed under the targeted policy. Hence, I find that efficiency and social welfare can be improved by using a targeted policy of subsidizing the geographic expansion costs of high-type productivity and high-type expansion efficiency firms and taxing the geographic expansion costs of all other types of firms.

Related Literature: This paper contributes to four strands of the literature. The first strand is the firm dynamics literature. This paper enlarges the scope of the present firm dynamics literature to include the geography of firm dynamics. Numerous papers such as Klette and Kortum (2004), Decker, Haltiwanger, Jarmin, and Miranda (2020), Argente, Baslandze, Hanley, and Moreira (2020), and Akcigit and Ates (2021) study firm dynamics, but their focus is on product markets rather than geographic markets. This paper emphasizes the role of geographic expansion in enabling the growth of firms in the non-tradable sectors. In my model, firms are defined as the portfolio of geographic markets they serve, hence firm dynamics is related to firms' presence in geographic markets.

The second strand of the literature to which this paper relates is the one on market power, innovation, growth, and antitrust policy. The dynamic general equilibrium model in this paper

[5]The subsidy rate in these experiments is constrained to range from -100% to 100%.

provides the theoretical setup to study how firms' geographic expansion outcomes can impact measures of local competition, which matters for outcomes related to innovation, growth and antitrust policy (as in Cavenaile, Celik, and Tian (2021), De Loecker, Eeckhout, and Unger (2020), Gutiérrez and Philippon (2016), Gutiérrez and Philippon (2017), and Covarrubias, Gutiérrez, and Philippon (2020)). Hence, this paper highlights the importance of understanding firms' geographic expansion decisions in evaluating antitrust policies.

The third strand of literature to which this paper contributes is the work on multi-establishment firms. Hsieh and Rossi-Hansberg (2023) is one of the most closely related papers to this one and provides good motivation for why we should care about local market concentration and firm dynamics across geographic markets. They shed light on the expansion in the number of markets per firm in non-tradable industries in the U.S., and attribute it to the availability of new fixed-cost-intensive technologies in these service industries. They show that this increase has been more pronounced for the top firms and has led to falling local market concentration in the new markets using the U.S. Census Longitudinal Business Database (1977-2013) and establishment level data from the economic censuses (1977-2012). Argente, Fitzgerald, Moreira, and Priolo (2021) highlight the importance of demand for firm growth. They analyze firms in the consumer food sector and find that entrants increase market share by entering more geographic markets and spend on marketing and advertising to reach more customers. Cao, Hyatt, Mukoyama, and Sager (2017) also find an increase in the average number of establishments per firm and a larger increase for larger firms in the service sector using data from the Quarterly Census of Employment and Wages (1990-2015).

Oberfield, Rossi-Hansberg, Sarte, and Trachter (2023) use U.S. establishment-level data to study the firms' decision of balancing the benefit of having multiple plants with the cost of operating these plants. They show that productive firms place more plants in dense high-rent locations and place fewer plants in markets with low density and low rents as compared to less productive plants. Xi (2023), Hsieh and Rossi-Hansberg (2019) and Cao et al. (2017) also highlight the importance of multi-establishments for firm growth. They emphasize that operating more establishments, instead of operating larger establishments, is the key margin for the growth of large firms. Jiang (2023) examines how information and communication technology (ICT) impacts firms' geographic organization and overall efficiency by enhancing internal communication. Using Census data on U.S. manufacturers, the author estimates a spatial equilibrium model in which firms adopt ICT, select multiple production sites, and engage in domestic trade. Franco (2024) studies firms' location decisions in imperfectly competitive markets, and its consequences for local competition. He also sheds light on the general equilibrium effects of changing firms' location choices using place-based policies. By contrast, I use a dynamic general equilibrium model where firms invest in a 'geographic expansion technology' that generates a Poisson rate of undirected expansion into another geographic market, thus affecting competition in local markets. Hsieh and Rossi-Hansberg (2023), Oberfield, Rossi-Hansberg, Sarte, and Trachter (2024), and Kleinman (2022) explore the competition between multi-establishment firms across regions. I contribute to this literature by providing additional evidence on multi-establishment firms in Canada and building a model with oligopolistic competition in local markets.[6]

The fourth strand of literature to which this paper contributes is the one related to concentration in output markets. Rossi-Hansberg, Sarte, and Trachter (2021) and Autor, Patterson, and Van Reenen (2023) document the divergence in trends in national and local product market concentration in the United States. Benkard, Yurukoglu, and Zhang (2021) measure concentration in narrowly defined product markets for a broad range of consumer goods and services in the U.S. from 1994 to 2019 and find that concentration levels are high but falling since 1994. Carlson and Mitchener (2009) assess the effects of branching on competition and stability of banking systems. I contribute to this literature by documenting novel evidence on the increase in local competition in Canada and study its link to aggregate welfare.

This paper is organized as follows: I describe the data and motivating facts in section 2. I introduce the dynamic model in section 3 and describe the model estimation in section 4. Section 5 applies my framework to the study of the sources of the increase in firms' geographic

[6] Jia (2008) and Holmes (2011) analyze the geographic expansion strategy of Walmart, Syverson (2008) uses the ready-mixed concrete industry as an example of an industry with a local market, Hortaçsu and Syverson (2015) describe the evolution of market concentration and scale in the retail industry, and Ganapati et al. (2018) analyze the expansion of warehouse and input use of the top firms in the wholesale industry.

expansion and the fall in local product market concentration in Canada. Section 6 discusses targeted and uniform policies and section 7 concludes.

2 Empirics

In this section, I describe the data and present some facts related to firms' geographic expansion and local market concentration in Canada that motivate the model developed in this paper.

2.1 Data

Studying firms' geographic expansion requires data that will allow me to observe firms' geographic presence over time. To this end, I use confidential administrative microdata on Canadian firms: the Canadian Employer Employee Dynamics Database (CEEDD) from 2001-2018, which is a set of linkable administrative tax files on firms and workers maintained by Statistics Canada. I use the following files for my analysis:

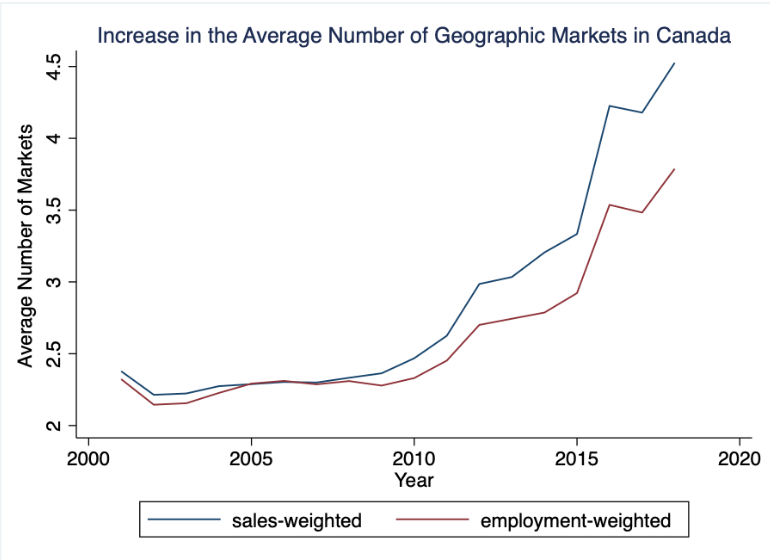
- T1 Personal Master File (T1 PMF) gives information on the demographic and financial characteristics of individuals.
- National Accounts Longitudinal Microdata File (NALMF) gives firm-level data prepared from multiple sources such as the Business Register (BR), Corporation Income Tax files (T2), General Index of Financial Information (GIFI), Statement of Remuneration Paid (T4), Remittances for Employee Income Taxes (PD7), Goods and Services Tax (GST) and the Harmonized Sales Tax (HST).
- T4 Statement of Remuneration Paid (T4) and Record of Employment (ROE) give job-level information.
- A set of sub-provincial geographic indicators that are derived using postal code information in T1 PMF give detailed information on firms' and workers' locations.

The business registry (BR) has all Canadian businesses which satisfy one of these criteria: has at least one employee, makes at least $30,000 in annual revenue, or is incorporated. The firms in this analysis cover the universe of all incorporated firms (that file T2) and unincorporated firms with employees (that file T4 and/or PD7) in Canada. I select active, for-profit firms (i.e., remove charities, government and trust & special funds) and do not consider firms in Agriculture, Mining, Utilities, Manufacturing, Education and Public Administration.[7] I use a novel empirical methodology to derive firms' geographic markets by exploiting the information on all firm-worker matches and the workers' residential locations.[8] This allows me to construct measures of number of geographic markets of each firm over time in Canada. I now describe the facts related to expansion in firms' geographic markets and fall in local market concentration that motivate the theoretical framework developed in this paper.

2.2 Motivating Facts

2.2.1 Average Number of Geographic Markets of Firms

Figure 1: Rise in Geographic Expansion in Canada



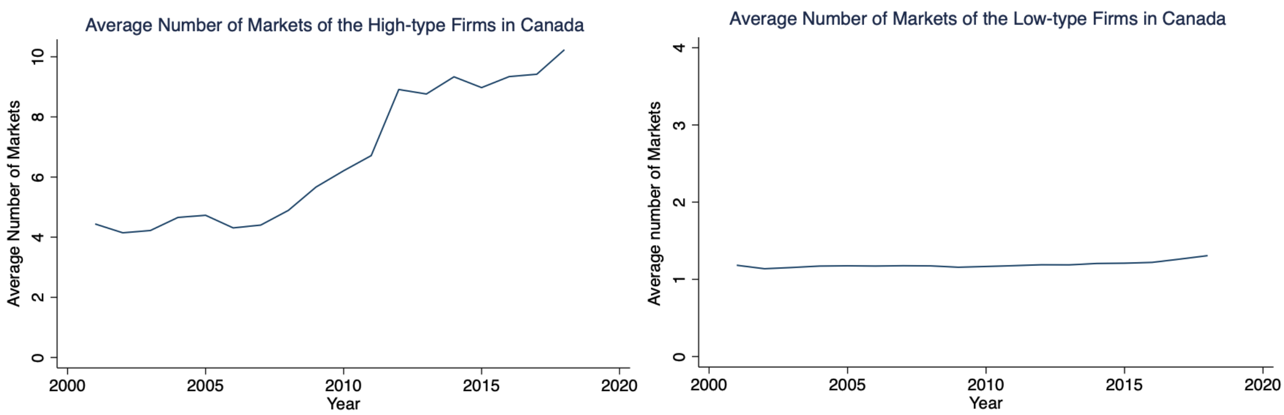
[7]They are excluded since economic activity in these sectors is tied to a specific geography, so firm growth may not come from geographic expansion. I include firms in Construction; Wholesale Trade; Retail Trade; Transportation and Warehousing (except some 4 digit NAICS codes where the description suggests that firms' operations are tied to certain geographies); Information and cultural Industries; Finance and Insurance (except central banking); Real Estate and Rental and Leasing; Professional, Scientific, and Technical Services; Management of Companies and Enterprises; Administrative and Support and Waste Management and Remediation Services; Health Care and Social Assistance; Arts, Entertainment, and Recreation; Accommodation and Food Services; Other Services (except Public Administration).

[8]See appendix 8.1 for a detailed description of my empirical methodology.

Figure 1 shows the average (sales-weighted/employment-weighted)[9] number of geographic markets of firms in Canada, i.e., for each year, I calculate the number of geographic markets for each firm within a sector, take their average to get a sector-level average and then calculate a national average from the sector-level averages using the sector-level sales/employment as weights. My measure of a local geographic market for this analysis is a census subdivision (CSD). The figure shows that an average firm in Canada was present in about 2.4 local markets in 2001, but expanded its presence to about 4.5 local markets (using a sales-weighted average) and to more than 3.5 local markets (using an employment-weighted average) in 2018.

2.2.2 Heterogeneity in the Average Number of Geographic Markets of Firms

Figure 2: Heterogeneity in the Average Number of Markets



Though figure 1 highlights the rise in the firms' average number of markets, it masks the heterogeneity in firms' average number of markets. To uncover this heterogeneity, I divide the firms in the data into different types depending on their expansion efficiencies.[10] In figure 2, I report the sales-weighted average number of markets for the high-type and low-type expansion efficiency firms. There is substantial heterogeneity in the number of geographic markets of firms.[11] An average high-type expansion efficiency firm in Canada was present in about 4.5 local markets in 2001, but expanded its presence to about 10 local markets in 2018. On the

[9] I calculate both sales-weighted and employment-weighted numbers for robustness since there is a rapid increase in the sales of a particular industry in the later period. The numbers obtained after removing firms in this industry still show an increasing trend.

[10] see Appendix 8.6.1 for the detailed description.

[11] This heterogeneity is robust to calculating sales-weighted and employment-weighted averages.

other hand, an average low-type expansion efficiency firm in Canada was present in about 1.18 local markets in 2001, but expanded its presence to about 1.3 local markets in 2018. This heterogeneity highlights the fact that there may be persistent differences between firms that have made it harder for some of them to expand geographically.

2.2.3 Local Competition

Figure 3: Fall in the Average Local HHI in Canada

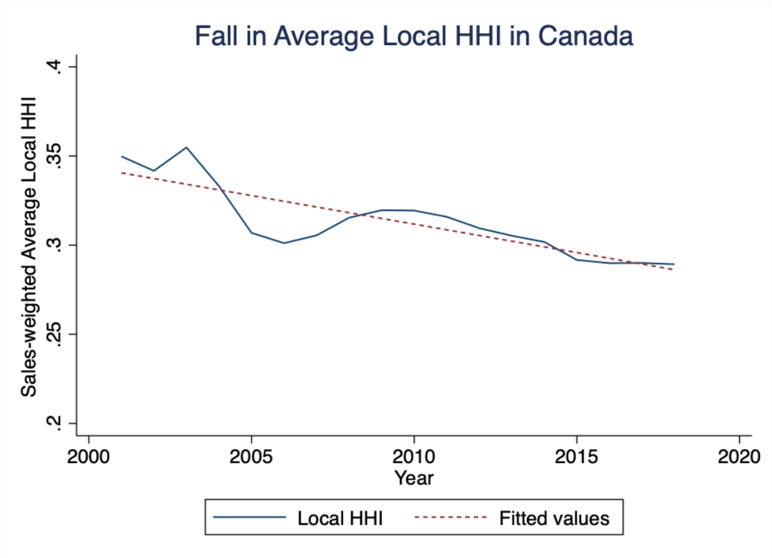


Figure 3 shows the increase in local competition in Canada. Each point on the graph denotes the sales-weighted average local HHI in Canada.[12] The figure shows that the sales-weighted average local HHI in Canada decreased from 0.35 in 2001 to 0.29 in 2018, signifying that the average local market became more competitive over this period. This aligns with the finding of an increase in the average number of markets for firms in Canada, suggesting that, on average, firms' expansion into more markets has resulted in a higher number of firms in each local market and less local market concentration overall.

[12]See appendix 8.6.1 for a description of the average local HHI calculation.

2.2.4 Average Markup

Figure 4: Rise in the Average Markup in Canada

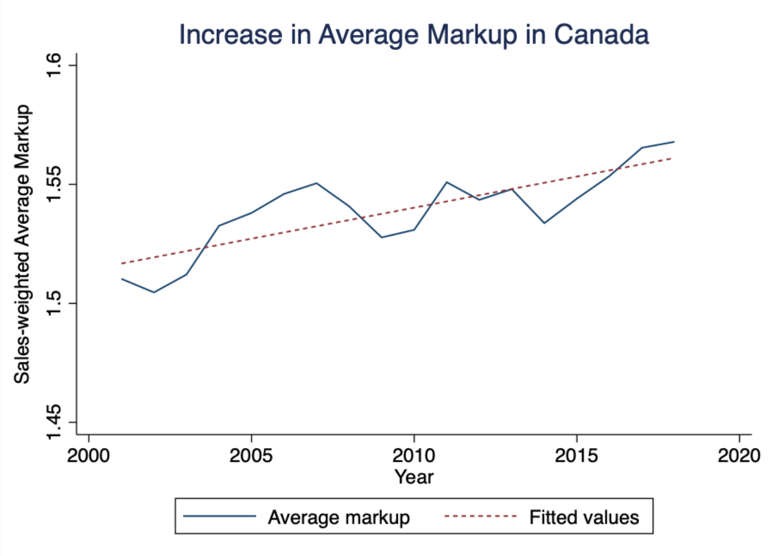


Figure 4 shows the rise in the sales-weighted average markup for Canadian firms between 2001-2018. The average markup weighted by industry-level sales increased from 1.51 in 2001 to 1.57 in 2018,[13] denoting an increase in firms' market power in terms of charging higher markups.

2.2.5 Discussion

The descriptive graphs presented above are informative about the trends in the firms' geographic expansion, but cannot shed light on the sources of these trends or their aggregate implications for the economy. In addition, the finding of an increase in the average markup coupled with the finding of a fall in average local HHI suggests that firm productivities might have changed in a way that makes it easier to charge higher markups, while being present in more competitive markets over time. These trends highlight the need for a theoretical framework to identify the relevant economic mechanisms that can help explain the observed trends and evaluate relevant economic policies. Hence, I build a model to explain the increase in geographic expansion, the decline in local product market concentration, and the increase in average markups, using it to analyze the sources and aggregate implications of these trends in Canada.

[13] The total cost weighted and total cost of sales weighted markups also have similar levels and changes in average markups.

3 Model

3.1 Environment

Preferences Time is continuous and indexed by t $\in R_+$. There is an infinitely-lived representative consumer who discounts the future at rate ρ and maximizes lifetime utility:

$$U = \int_0^\infty e^{-\rho t} \ln(C_t) dt \tag{1}$$

where C_t is the consumption of the final good at time t and the price of the final good is normalized to 1. The household supplies one unit of labor inelastically and earns an endogenously determined wage rate w_t in exchange. The household owns all the assets in the economy and faces the following budget constraint:

$$\dot{\mathcal{A}}_t = r_t \mathcal{A}_t + w_t - C_t \tag{2}$$

where $\mathcal{A}_t$ is household wealth and r_t is the rate of return on assets.

Final Good Production The economy consists of a unit continuum of geographic markets indexed by j. Final good Y_t is produced competitively using inputs from these geographic markets:

$$\ln Y_t = \left(\int_0^1 \omega_j \ln(y_{jt}) dj \right) \tag{3}$$

where y_{jt} is production in market j at time t and ω_j is a time-invariant market specific demand shifter (which could of a high-type or low-type)[14].

Geographic Market Production Each geographic market has an endogenous number of large firms having exogenously given time-invariant productivities. Firms can be high-type or low-type productivity firms ($N_{jt} \in \{1, ..., \bar{N}\}, N_{jt} = N_{Ljt} + N_{Hjt}$ where N_{Ljt} is the num-

[14] $\omega_j \in \{\omega_H, \omega_L\}$ where I use j to index the geographic market as well as the type (high/low) of the geographic market demand shifter for notational ease.

ber of low-type productivity firms in market j at time t and N_{Hjt} is the number of high-type productivity firms in market j at time t). Firms in each market compete in quantities, each producing a differentiated variety as the result of a static Cournot game. Each market also has a competitive fringe composed of a mass m_{jt} of small firms producing a homogeneous good. All small firms in the competitive fringe are price takers since there is a continuum of small firms and their products are homogeneous. The markets without any large firms also have non-zero production due to the competitive fringe's output. Output in market j at time t is a CES aggregate of the firms' production:

$$y_{jt} = \left[\sum_{n=1}^{N_{jt}} y_{njt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \left[N_{Hjt} y_{Hjt}^{\frac{\sigma-1}{\sigma}} + N_{Ljt} y_{Ljt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{4}$$

where y_{njt} is the production of the n^{th} large firm (which could be a high-type or a low-type productivity firm) in market j at time t, y_{csjt} is the production of small firm s in the fringe in market j at time t, $\tilde{y}_{cjt} = \int_{F_{jt}} y_{csjt} ds$ is the production of the competitive fringe in market j at time t, F_{jt} is the set of small firms in the fringe in market j at time t and σ is the elasticity of substitution between varieties in the market.

Variety Production Large firms' productivity types are indexed by i_A. They could be of high-type ($i_A = H$) or of low-type ($i_A = L$).[15] Out of N_{jt} firms in market j at time t, N_{Hjt} firms have 'high' productivity A_{Hjt} and $N_{Ljt} = N_{jt} - N_{Hjt}$ firms have 'low' productivity A_{Ljt}. They produce their variety in each market using a linear production technology in labor:

$$y_{i_A jt} = A_{i_A jt} l_{i_A jt}, \quad i_A \in \{L, H\} \tag{5}$$

where $A_{i_A jt}$ is the productivity of the type i_A large firm in market j and $l_{i_A jt}$ is the labor demand of the type i_A large firm in market j.

[15] These indices actually mean $i_A(n)$ since they map every n^{th} firm to a type i_A.

Similarly for every small firm s in the competitive fringe:

$$y_{csjt} = A_{sjt} l_{csjt} \tag{6}$$

I assume that every small firm in the fringe in a given market has the same productivity A_{cjt}.

Geographic Expansion Technology In addition to having heterogeneous productivities (indexed by i_A), large firms also differ in their expansion efficiency type (indexed by i_χ)[16] which can be of a high-type or low-type ($i_\chi \in \{H, L\}$).[17] A large firm is defined by the set of local geographic markets it serves (indexed by k), so it adds new markets using a 'geographic expansion technology' (referred to as 'expansion' from now on) which is a function of the firm type and the firm's existing number of markets at time t (K_t). Each large firm can invest in expansion to increase its presence in multiple geographic markets. To generate a Poisson rate of success in expansion $x_{i_A,i_\chi,t}$, a firm of type (i_A, i_χ) must pay a cost in units of the final good equal to:

$$R_{i_A,i_\chi,t}(K_t) = \chi_{i_\chi,t}\left(\frac{X_{i_A,i_\chi,t}}{K_t}\right)^{\phi} K_t Y_t = \chi_{i_\chi,t}\left(x_{i_A,i_\chi,t}\right)^{\phi} K_t Y_t, \quad i_A \in \{L, H\},\ i_\chi \in \{L, H\} \tag{7}$$

where $\chi_{i_\chi,t} > 0$ is the expansion cost scale parameter, $\phi > 1$ is the large firms' expansion cost convexity parameter and K_t is the number of existing geographic markets of the firm at time t.[18]

Exit of Incumbent Large Firms δ is the exogenous exit probability of a geographic market and δ_f is the exogenous shut-down probability (exit from the economy) for a large firm in the economy.

Entry of New Large Firms A firm in the competitive fringe cannot become a large firm

[16] These indices actually mean $i_\chi(n)$ since they map every n^{th} firm to a type i_χ.

[17] A high-type firm could have a high or low expansion efficiency ($i_A = H, i_\chi \in \{L, H\}$), i.e., the types i_A and i_χ are independent of each other.

[18] Firms' geographic expansion costs include expenses related to acquiring new infrastructure, machinery, and equipment, as well as costs for market research and legal or regulatory compliance.

and cannot exit a market or the economy. There is a mass one of entrepreneurs who can pay a cost $\psi z_t^2 Y_t$ to generate a Poisson rate z_t of starting a new firm (upon successful innovation, they form a new large firm that is randomly assigned its type and randomly allocated to a geographic market). In order to keep the mass of entrepreneurs unchanged, I assume that they sell their firm on a competitive market at its full value and remain in the set of entrepreneurs.[19]

3.2 Equilibrium

Consumer's Problem Household lifetime utility maximization gives the standard Euler equation:

$$\frac{\dot{C}_t}{C_t} = r_t - \rho \tag{8}$$

Final Good Producers The final good market is competitive. Given σ and prices $\{P_t, p_{njt}\}$, the representative final good producer chooses the quantity of each variety in each geographic market to maximize profit. The corresponding optimization problem gives the inverse demand function for variety 'n' in geographic market 'j':

$$\max_{[\{y_{njt}\}_{n=1}^{N_{jt}}, \tilde{y}_{cjt}]_{j=0}^{1}} \exp\left(\frac{\sigma}{\sigma-1}\int_0^1 \omega_j \ln\left[\sum_{n=1}^{N_{jt}} y_{njt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}\right] dj\right) - \int_0^1 \left(\sum_{n=1}^{N_{jt}} p_{njt} y_{njt} + p_{cjt}\tilde{y}_{cjt}\right) dj \tag{9}$$

where p_{njt} is the price of variety n in market j at time t and p_{cjt} is the price of the competitive fringe variety in market j at time t. This leads to the following inverse demand function:

$$p_{njt} = Y_t \omega_j \frac{y_{njt}^{-\frac{1}{\sigma}}}{\sum_{k=1}^{N_{jt}} y_{kjt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}} \tag{10}$$

Also,

$$\frac{y_{njt}}{y_{kjt}} = \left(\frac{p_{kjt}}{p_{njt}}\right)^{\sigma} \tag{11}$$

[19] This gives the same observed equilibrium as assuming that successful entrepreneurs are replaced by new entrepreneurs after new business creation. The representative household gets all entrepreneurial profits, which would remain the same regardless of which assumption is chosen.

where y_{njt} should be replaced by $\tilde{y}_{cjt}$ for the competitive fringe.

Variety Producers Large firms in each local market compete in quantities. Each large firm maximizes profit:

$$\max_{y_{njt}} \; p_{njt}y_{njt} - w_t\frac{y_{njt}}{A_{njt}} = Y_t\omega_j\frac{y_{njt}^{\frac{\sigma-1}{\sigma}}}{\sum_{k=1}^{N_{jt}} y_{kjt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}} - \frac{wy_{njt}}{A_{njt}} \tag{12}$$

This gives the following best response function:

$$y_{njt} = \left[\frac{\sigma-1}{\sigma}A_{njt}\frac{Y_t\omega_j}{w_t}\frac{\sum_{k\neq n} y_{kjt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}}{\left(\sum_{k=1}^{N_{jt}} y_{kjt}^{\frac{\sigma-1}{\sigma}} + \tilde{y}_{cjt}^{\frac{\sigma-1}{\sigma}}\right)^2}\right]^{\sigma} \tag{13}$$

$$= \frac{\sigma-1}{\sigma}A_{njt}\frac{Y_t\omega_j}{w_t}\frac{\sum_{k\neq n}\left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}}{\left[\sum_{k=1}^{N_{jt}}\left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}\right]^2} \tag{14}$$

The total production of the competitive fringe is:

$$\tilde{y}_{cjt} = A_{cjt}\frac{\frac{Y_t\omega_j}{w_t}}{\sum_{k=1}^{N_{jt}}\left(\frac{y_{kjt}}{\tilde{y}_{cjt}}\right)^{\frac{\sigma-1}{\sigma}} + 1} \tag{15}$$

Relative production between each large firm's variety and the competitive fringe within the geographic market can be written as:

$$\left(\frac{y_{njt}}{y_{kjt}}\right)^{\frac{1}{\sigma}} = \frac{A_{njt}}{A_{kjt}}\frac{\sum_{l\neq n}\left(\frac{y_{ljt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}}{\sum_{l\neq k}\left(\frac{y_{ljt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}} \tag{16}$$

$$\left(\frac{y_{njt}}{\tilde{y}_{cjt}}\right)^{\frac{1}{\sigma}} = \frac{\sigma-1}{\sigma}\frac{A_{njt}}{A_{cjt}}\frac{\sum_{l\neq n}\left(\frac{y_{ljt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}}{\sum_{l=1}^{N_{jt}}\left(\frac{y_{ljt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}} \tag{17}$$

Given the number of high-type and low-type productivity large firms and the relative productivities of these firms and the competitive fringe, I use the following three equations to obtain all unknown production ratios within each geographic market:

$$\left(\frac{y_{Hjt}}{y_{Ljt}}\right)^{\frac{1}{\sigma}} = \frac{A_{Hjt}}{A_{Ljt}} \frac{(N_H - 1) + N_L \left(\frac{y_{Ljt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}}}{(N_L - 1)\left(\frac{y_{Ljt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}} + N_H + \left(\frac{\tilde{y}_{cjt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}}} \tag{18}$$

$$\left(\frac{y_{Hjt}}{\tilde{y}_{cjt}}\right)^{\frac{1}{\sigma}} = \frac{\sigma-1}{\sigma} \frac{A_{Hjt}}{A_{cjt}} \frac{(N_H - 1) + N_L \left(\frac{y_{Ljt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}}}{\left(N_H + N_L\left(\frac{y_{Ljt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{Hjt}}\right)^{\frac{\sigma-1}{\sigma}}\right)} \tag{19}$$

$$\left(\frac{y_{Ljt}}{\tilde{y}_{cjt}}\right)^{\frac{1}{\sigma}} = \frac{\sigma-1}{\sigma} \frac{A_{Ljt}}{A_{cjt}} \frac{N_H \left(\frac{y_{Hjt}}{y_{Ljt}}\right)^{\frac{\sigma-1}{\sigma}} + (N_L - 1) + \left(\frac{\tilde{y}_{cjt}}{y_{Ljt}}\right)^{\frac{\sigma-1}{\sigma}}}{\left(N_H\left(\frac{y_{Hjt}}{y_{Ljt}}\right)^{\frac{\sigma-1}{\sigma}} + N_L + \left(\frac{\tilde{y}_{cjt}}{y_{Ljt}}\right)^{\frac{\sigma-1}{\sigma}}\right)} \tag{20}$$

Variety Prices Variety prices (p_{njt}) can then be derived as a function only of relative productivities in the geographic market:

$$p_{njt} = \left(\frac{\sigma}{\sigma-1}\right)\left(\frac{w_t}{A_{njt}}\right) \frac{\left(\sum_{k=1}^{N_{jt}} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}\right)}{\sum_{k\neq n} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}} \tag{21}$$

Variety Markups Firms charge varying markups over marginal cost depending on the number of firms and their relative productivities (which determines relative outputs):

$$markup_{njt} = \frac{p_{njt}}{\frac{w_t}{A_{njt}}} = \left(\frac{\sigma}{\sigma-1}\right) \frac{\left(\sum_{k=1}^{N_{jt}} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}\right)}{\sum_{k\neq n} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}} \tag{22}$$

Variety Profits Profits charged by the large firms also depend on relative productivities in the market:

$$\pi_{njt} = \frac{Y_t \omega_j}{\left[\sum_{k=1}^{N_{jt}} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}\right]^2} \cdot \frac{\sigma + \sum_{k \neq n} \left(\frac{y_{kjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{\tilde{y}_{cjt}}{y_{njt}}\right)^{\frac{\sigma-1}{\sigma}}}{\sigma} \tag{23}$$

Large Firms' Value Functions and Expansion Decisions Static profits within each geographic market only depend on the number of high-type and low-type productivity large firms, their relative productivities and the market demand shifter. The relevant state variables for a large firm of type (i_A, i_χ) in market 'k' at time 't' are its number of geographic markets at time t (K_t) and the vector $\{\Theta_{kt}\}_{k=1}^{K_t} = \{N_{Hkt}, N_{Lkt}, \omega_k\}_{k=1}^{K_t}$ denoting the number of high-type and low-type productivity large firms and the demand shifter in each of its individual markets.[20] Since $K_t = |(N_{Hkt}, N_{Lkt}, \omega_k)_{k=1}^{K_t}|$, I drop state variable K_t and time subscripts unless needed. A market can have at most $\bar{N}$ firms (a firm's entry into a market fails if it enters a market which already has $\bar{N}$ firms). There is an exogenous exit probability δ for every geographic market and an exogenous shut-down probability δ_f (the large firm exits the economy in this case). e_H and e_L denote the successful expansion into a current market by large firms with high-type productivity and low-type productivity respectively. A large firm of type (i_A, i_χ) chooses an expansion rate $x_{i_A,i_\chi}(\{\Theta_k\}_{k=1}^{K})$ of expanding into another geographic market to maximize the value of the firm.[21]

High-type productivity large firm[22] The value function for a type (H, i_χ) large firm is:

[20]The index on ω can be understood as j(k), as $\omega_j : [0,1] \to \{H, L\}, k \in \{1, 2, ..., K\}, j(k) : \mathcal{Z}_+ \to [0,1], \{j(k) \mid k \in \{1, 2, ...K\}\} \subseteq [0,1]$.

[21]Choosing x_{i_A,i_χ} is equivalent to choosing X_{i_A,i_χ} since $x_{i_A,i_\chi} = X_{i_A,i_\chi}/K$. Expansion investment can lead to expansion into any geographic market with the same probability (undirected expansion).

[22]See Appendix 8.2 for the corresponding value function for a low-type productivity large firm.

$$
\begin{aligned}
&rV_{H,i_\chi}(\{\Theta_k\}_{k=1}^K) \\
&= \max_{x_{H,i_\chi}} \sum_{k=1}^{K} \pi_H(\Theta_k)Y - \chi_{i_\chi}\left(x_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)\right)^\phi KY \\
&+ x_{H,i_\chi}K[\mathbb{E}\,V_{H,i_\chi}(\{\Theta_k\}_{k=1}^{K+1}) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \delta_f[0 - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] + \delta\sum_{l=1}^{K}[V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K - \Theta_l) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_H\sum_{l=1}^{K}[V_{H,i_\chi}((\{\Theta_k\}_{k=1}^K - \Theta_l) \cup (\min(N_{Hl}+1, \bar{N} - N_{Ll}), N_{Ll}, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_L\sum_{l=1}^{K}[V_{H,i_\chi}((\{\Theta_k\}_{k=1}^K - \Theta_l) \cup (N_{Hl}, \min(N_{Ll}+1, \bar{N} - N_{Hl}), \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta+\delta_f)\sum_{l=1}^{K} N_{Ll}[V_{H,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl}, N_{Ll}-1, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta+\delta_f)\sum_{l=1}^{K}(N_{Hl}-1)[V_{H,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl}-1, N_{Ll}, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \dot{V}_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)
\end{aligned} \tag{24}
$$

A "∪" indicates a set with the following element added. The profit in market k normalized by total output Y is denoted by $\pi_H(N_{Hk}, N_{Lk}, \omega_k)$. The first line is the high-type productivity large firm's expected flow profit (from all current geographic markets) minus the total cost of expansion investment. The second line is the change in firm value due to successful geographic expansion by the firm which happens at Poisson rate x_{H,i_χ}. If the firm successfully enters another market, it increases its number of markets K by 1. The third line is the change in value due to exogenous exit: with probability δ_f, the firm exits the economy and gets a value of 0, and with probability δ it exits a local market and gets the value associated with being present in one less market (and this can happen in any of the K markets it is already present in). The fourth and fifth lines show the change in value from any other firm entering a market in which this firm is already present. e_H denotes the probability of another high-type productivity large firm (which could have high-type or low-type expansion efficiency) arriving in a market and this can happen in any of the K markets this firm is in. In this case, the number of high-type productivity firms increases by 1 (as long as the total number of firms in the market does not exceed $\bar{N}$). e_L denotes the probability of a low-type productivity large firm (which could have high-type or low-type expansion efficiency) arriving in a market. This can happen in any of the K markets this firm is in and then the number of low-type productivity large firms increases by 1 (as long as the total number of firms in the market does not exceed $\bar{N}$). The sixth and seventh lines denote the event in which other firms in the high-type productivity large firm's markets exit or lose their product line. The sixth line denotes the event in which a low-type productivity large firm leaves one of the existing markets. This can be any of the N_L large firms in any of the K existing markets of the firm, and then the number of low-type productivity large firms decreases by 1. The seventh line denotes the event in which another high-type productivity large firm leaves one of the existing markets. This can be any of the $(N_H - 1)$ large firms in any of the K existing markets of the firm, and then the number of high-type productivity large firms decreases by 1. The last line denotes the growth in firm value.

Theorem 1 *The* (i_A, i_χ) *type large firm's value function can be written as* $V_{i_A,i_\chi}(\{\Theta_k\}_{k=1}^K) = \sum_{k=1}^K V_{i_A,i_\chi}(\Theta_k) = \sum_{k=1}^K v_{i_A,i_\chi}(\Theta_k)Y = \sum_{N_H=0}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{i_A,i_\chi}(\Theta)v_{i_A,i_\chi}(\Theta)Y$ *where*

- $v_{i_A,i_\chi}(\Theta)$ *are time-invariant scalars in a stationary equilibrium*
- $m_{i_A,i_\chi}(\Theta)$ *are integers denoting the number of markets the large firm is in with* N_H *high-type productivity large firms,* N_L *low-type productivity large firms and* ω_k *market demand shifter*

i.e., the value function is linear in the number of geographic markets, total output and some scalars.

Using equation 8 and the theorem above for the high-type productivity large firms, I can write[23]

[23] See appendix 8.3 for the proof in the case of a high-type productivity firm.

$$[\rho+\delta_f+\delta+e_H+e_L]\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\Theta)$$

$$= \sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}}$$
$$\Big(m_{H,i_\chi}(\Theta)\pi_H(\Theta)-\chi_{i_\chi}x_{H,i_\chi}^{\phi}m_{H,i_\chi}(\Theta)+x_{H,i_\chi}m_{H,i_\chi}(\Theta)\hat{\mathbb{E}}(v_{H,i_\chi})$$
$$+e_H m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\min(N_H+1,\bar{N}-N_L),N_L,\omega_k)+e_L m_{H,i_\chi}(\Theta)v_{H,i_\chi}(N_H,\min(N_L+1,\bar{N}-N_H),\omega_k)$$
$$+(\delta+\delta_f)m_{H,i_\chi}(\Theta)\left(N_L[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H,N_L-1,\omega_k)]+(N_H-1)[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H-1,N_L,\omega_k)]\right)\Big) \tag{25}$$

The optimal level of geographic expansion of the high-type productivity large firm is:

$$x_{H,i_\chi}=\frac{X_{H,i_\chi}}{K}=\left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1}\sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)}\sum_{\omega_l\in\{\omega_H,\omega_L\}}\mu_{N_{Hl},N_{Ll},\omega_l}v_{H,i_\chi}(N_{Hl}+1,N_{Ll},\omega_l)}{\phi\chi_{i_\chi}}\right]^{\frac{1}{\phi-1}}=\left[\frac{\hat{\mathbb{E}}(v_{H,i_\chi})}{\phi\chi_{i_\chi}}\right]^{\frac{1}{\phi-1}} \tag{26}$$

Similarly, the optimal level of geographic expansion of the low-type productivity large firm is:

$$x_{L,i_\chi}=\frac{X_{L,i_\chi}}{K}=\left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1}\sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)}\sum_{\omega_l\in\{\omega_H,\omega_L\}}\mu_{N_{Hl},N_{Ll},\omega_l}v_{L,i_\chi}(N_{Hl},N_{Ll}+1,\omega_l)}{\phi\chi_{i_\chi}}\right]^{\frac{1}{\phi-1}}=\left[\frac{\hat{\mathbb{E}}(v_{L,i_\chi})}{\phi\chi_{i_\chi}}\right]^{\frac{1}{\phi-1}} \tag{27}$$

where $\mu_{N_{Hl},N_{Ll},\omega_l}$ is the mass of geographic markets with N_H high-type productivity large firms, N_L low-type productivity large firms and ω_l market demand shifter. Note that the optimal geographic expansion rate of a large firm is independent of the expansion rate choices of other firms in its geographic markets, but depends on the choices of other firms in the economy (which determines the market-state distribution $\mu_{N_{Hl},N_{Ll},\omega_l}$ through general equilibrium effects, hence affecting the geographic expansion rate of the given firm). In addition, the optimal geographic expansion rate of a large firm falls with an increase in the large firms' expansion cost convexity parameter (ϕ) and expansion cost scale parameter χ_{i_χ}. I focus on the Markov Perfect equilibrium at the intersection of all firms' best response functions.[24]

[24] This is common in innovation models such as Aghion, Harris, Howitt, and Vickers (2001).

Entrepreneurs and Entry into a Market Recall that there is a mass one of entrepreneurs who can pay an R&D cost $\psi z^2 Y$ to generate a Poisson rate z of starting a new firm. Upon successful innovation, they form a new large firm that is randomly assigned its type (i_A, i_χ) and randomly allocated to a geographic market. I assume that a successful entrepreneur immediately sells the firm in a competitive market. The expected selling price of a new firm of a successful entrepreneur with type (H, i_χ) is given by $W_{H,i_\chi} = \sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} V_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l) = \hat{\mathbb{E}}(v_{H,i_\chi})$ where $\mu_{N_{Hl},N_{Ll},\omega_l}$ is the mass of geographic markets with N_H high-type productivity large firms, N_L low-type productivity large firms and ω_l market demand shifter.[25] Then the expected selling price of a new firm of a successful entrepreneur is $\mathbb{E}(W) = (\Lambda_{HH}\hat{\mathbb{E}}(v_{HH}) + \Lambda_{HL}\hat{\mathbb{E}}(v_{HL}) + \Lambda_{LH}\hat{\mathbb{E}}(v_{LH}) + \Lambda_{LL}\hat{\mathbb{E}}(v_{LL}))$, where Λ_{i_A,i_χ} is the probability of getting assigned type (i_A, i_χ) upon successful entry in the economy such that $\sum_{i_A}\sum_{i_\chi}\Lambda_{i_A,i_\chi} = 1$.

The value of being an entrepreneur (S) can be written as:

$$\rho S = \max_z z\,\mathbb{E}(W) - \psi z^2 Y \tag{28}$$

Guessing and verifying that in a stationary equilibrium, $S = sY$, I get

$$z = \frac{\mathbb{E}(W)}{2\psi Y} = \frac{\Lambda_{HH}\hat{\mathbb{E}}(v_{HH}) + \Lambda_{HL}\hat{\mathbb{E}}(v_{HL}) + \Lambda_{LH}\hat{\mathbb{E}}(v_{LH}) + \Lambda_{LL}\hat{\mathbb{E}}(v_{LL})}{2\psi} \tag{29}$$

Equilibrium Definition I focus on the unique Markov Perfect equilibrium of the economy. An equilibrium is defined by a set of allocations $\{Y_t, C_t\}$, policy functions $\Big([\{y_{L(n)jt}\}_{n=1}^{N_{Ljt}}]_{j=0}^{1}$, $[\{y_{H(n)jt}\}_{n=1}^{N_{Hjt}}]_{j=0}^{1}$, $[\{\tilde{y}_{cjt}\}]_{j=0}^{1}$, $[\{l_{L(n)jt}\}_{n=1}^{N_{Ljt}}]_{j=0}^{1}$, $[\{l_{H(n)jt}\}_{n=1}^{N_{Hjt}}]_{j=0}^{1}$, $[\{l_{csjt}\}]_{j=0}^{1}$, $[\{x_{L(n)i_\chi jt}\}_{n=1}^{N_{Ljt}}]_{j=0}^{1}$, $[\{x_{H(n)i_\chi jt}\}_{n=1}^{N_{Hjt}}]_{j=0}^{1}, z_t\Big)_{t=0}^{\infty}$ for $i_\chi \in \{L, H\}$, prices $\Big([\{\{p_{L(n)jt}\}_{n=1}^{N_{Ljt}}]_{j=0}^{1}$, $[\{\{p_{H(n)jt}\}_{n=1}^{N_{Hjt}}]_{j=0}^{1}, [\{p_{cjt}\}]_{j=0}^{1}, w_t, r_t\}\Big)_{t=0}^{\infty}$, a set of vectors $\{N_{Hjt}, N_{Ljt}, \omega_j\}_{j=1}^{\infty}$ that denote the number of high-type productivity large firms, low-type productivity large firms and market de-

[25] Similarly, the expected selling price of a new firm of a successful entrepreneur with type (L, i_χ) is given by $W_{L,i_\chi} = \sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} V_{L,i_\chi}(N_{Hl}, N_{Ll}+1, \omega_l) = \hat{\mathbb{E}}(v_{L,i_\chi})$

mand shifters in each of the 'j' geographic markets in the economy, market-state distribution[26] $\{\{\{\{\{\zeta_{\zeta_{N_{HH},N_H-N_{HH},N_{LH},N_L-N_{LH},\omega_k}}\}_{N_H=0}^{\bar{N}}\}_{N_{HH}=0}^{N_H}\}_{N_L=0}^{\bar{N}-N_H}\}_{N_{LH}=0}^{N_L}\}_{\omega_j\in\{\omega_H,\omega_L\}}$ and firm-size distribution[27] $\gamma_{i_A,i_\chi}(j), i_A \in \{L,H\}, i_\chi \in \{L,H\},\ j \in \{1,2,...,\infty\}$ such that, $\forall t \geq 0$:

1. The representative household maximizes her lifetime utility subject to the intertemporal budget constraint, hence aggregate consumption C_t grows at rate $r_t - \rho$ in the stationary equilibrium.
2. Given prices, the final good producer maximizes profit.
3. Given $(N_{Hjt}, N_{Ljt}, \omega_j)$, large incumbent firms choose y_{i_Ajt} to maximize profit in each local market.
4. The competitive fringe in each local market equates $p_{cjt} = MC$ to get $\tilde{y}_{cjt}$.
5. Given K_t and $\{N_{Hkt}, N_{Lkt}, \omega_k\}_{k=1}^{K_t}$, large incumbent firms of type (i_A, i_χ) choose expansion policy x_{i_A,i_χ} to maximize firm value.
6. Entrepreneurs break even in expectation: they choose z_t to maximize their value.
7. The real wage rate w_t clears the labor market.
8. Resource constraint is satisfied: $Y_t = C_t + \psi z_t^2 Y_t +$
$+ \int_0^1 \left[N_{HHjt}.\chi_H.x_{HHt}^{\phi}.Y_t + N_{HLjt}.\chi_L.x_{HLt}^{\phi}.Y_t + N_{LHjt}.\chi_H.x_{LHt}^{\phi}.Y_t + N_{LLjt}.\chi_L.x_{LLt}^{\phi}.Y_t\right]dj$

4 Estimation

In this section, I describe the estimation of my model using moments from CEEDD. The model is estimated for two periods: 2002-2005 (referred to as the early period) and 2013-2016 (referred to as the late period). The two estimated stationary equilibria are used in the quantitative application of the model in Section 5 which disentangles the sources of firms' geographic expansion observed in Canada over 2001-2018.

[26]This maps to $\{\{\{\mu_{N_{Hj},N_{Lj},\omega_j}\}_{N_{Hj}=0}^{\bar{N}}\}_{N_{Lj}=0}^{\bar{N}-N_H}\}_{\omega_j\in\{\omega_H,\omega_L\}}$. Refer to appendix 8.4 for details.

[27]Here firm-size refers to the number of geographic markets, and not employment as in Klette and Kortum (2004). Refer to appendix 8.5 for details.

Fifteen parameters must be determined: $\rho, s_{\omega_H}, \omega_H/\omega_L, A_H/A_L, \delta, \psi, \delta_f, A_H/A_c, \chi_H, \chi_L, \phi,$ $\lambda_{HH}, \lambda_{HL}, \lambda_{LL}, \sigma$. The consumer discount rate ρ is set to 0.04, which implies a real interest rate of 4% when the growth rate is 0% (as in my model). The share of high-type markets s_{ω_H} is set to be equal to the share set in the data, and the high-type market demand shifter relative to the low-type market ω_H/ω_L is derived using population ratios of local markets (CSDs) once s_{ω_H} is set.[28] The relative productivity of high-type productivity firms with respect to the low-type productivity firms A_H/A_L is obtained by assuming that the top 40% of the firms are high-type productivity firms and calculating the ratio of average labor productivities of the two types of firms.[29] The rest of the parameters are structurally estimated following a simulated method of moments (SMM) approach.[30] The success of the SMM estimation depends on choosing moments that are sensitive to changes in the structural parameters. All these parameters are jointly estimated to match the following targeted data moments: the sales-weighted average local HHI, firm exit rate, high-type expansion efficiency firms' average number of markets, low-type expansion efficiency firms' average number of markets, expansion expenditure ratio, proportion of HH-type firms, proportion of HL-type firms, proportion of LH-type firms, average markup, standard deviation of markups, average market share of top 4 firms, standard deviation of the number of markets and the labor share.[31]

Table 1 reports the values of the externally set parameters, whereas table 2 provides the parameter estimates of the internally (jointly) estimated parameters for both estimations. Table 3 shows the values of the targeted moments in the data and the estimated model. The model matches the thirteen data moments well for both the estimations. The Jacobian matrix of the model moments with respect to the model parameters in percentage terms is displayed in Table 8.

[28] For example, $s_{\omega_H} = 0.05$ implies that the top 5% markets in Canada are designated to be high-type markets, and the ratio of their average population to the average population of the remaining low-type markets is the number given in table 1. In this case, I trimmed 10% of the markets on each side due to the extreme variation in populations of CSDs in Canada.

[29] More details can be found in appendix 8.6.1.

[30] See appendix 8.6 for a description of the estimation algorithm and the objective function.

[31] A more detailed description of the data moments and which moments help in identifying which parameters is provided in appendix 8.6.

Table 1: Externally Set Parameters

Parameter	*Description*	*Early period*	*Late period*
ρ	household discount rate	0.04	0.04
s_{ω_H}	share of H-type markets	0.05	0.05
ω_H/ω_L	H-type market demand shifter relative to L-type	4.9335	5.1122
A_H/A_L	relative productivity of H-type firms w.r.t. L-type firms	2.8706	6.4477

Table 2: Internally Estimated Parameters

Parameter	*Description*	*Early period*	*Late period*
δ	exogenous market exit rate	0.0058	0.0348
ψ	entrepreneurs' innovation cost scale	0.9446	1.1377
δ_f	exogenous firm shut-down rate	0.0267	0.0096
A_H/A_c	relative productivity of H-type firms w.r.t. fringe	1.8681	1.9213
χ_H	H-type expansion efficiency firms' expansion cost scale	223.2071	91.1950
χ_L	L-type expansion efficiency firms' expansion cost scale	681.7752	152.4749
ϕ	large firms' expansion cost convexity	3.7599	3.2612
λ_{HH}	probability of being HH-type upon successful entry	0.0699	0.0463
λ_{HL}	probability of being HL-type upon successful entry	0.1948	0.1478
λ_{LL}	probability of being LL-type upon successful entry	0.5337	0.7515
σ	elasticity of substitution between varieties in a market	31.6488	20.6766

Notes: The estimation is conducted using a simulated method of moments (SMM) approach. This table reports the estimated parameters for the two estimations: early period (2002-2005) and late period (2013-2016).

4.1 Change in Parameters

As reported in table 2, the estimated parameter values in the two periods are quite different, which captures the structural changes in the Canadian economy throughout this time period. The elasticity of substitution within a local market (σ) decreases from 31.65 to 20.68. This parameter signifies the extent of local product market competition between large firms and the competitive fringe, and the fall over the two periods signifies an increase in market power since a lower elasticity of substitution between local goods implies that a given firm's goods become more unique.[32] This increase in market power is coupled with an increase in the relative productivity of high-type productivity firms with respect to the low-type productivity firms (A_H/A_L) which increases from 2.87 to 6.45. In addition, the relative productivity of high-type productivity firms with respect to the fringe (A_H/A_c) increases from 1.87 to 1.92, but this increase is smaller than the increase in A_H/A_L, implying an increase in productivity of the competitive fringe (since A_L is normalized to 1). Thus, the high-type productivity large firms

[32] This is in line with the finding of increased market power in the U.S. over the last decades as in De Loecker et al. (2020).

Table 3: Data and Model Moments

	Early period		*Late period*	
Target Moments	*Data*	*Model*	*Data*	*Model*
sales-weighted average local HHI	0.334	0.305	0.297	0.297
firm exit rate (%)	4.108	3.137	5.799	3.708
average number of markets of H-type firms	4.161	4.108	7.291	7.241
average number of markets of L-type firms	1.160	1.158	1.206	1.206
expansion expenditure ratio	0.047	0.047	0.074	0.072
proportion of HH-type firms	0.151	0.159	0.158	0.148
proportion of HL-type firms	0.249	0.270	0.242	0.247
proportion of LL-type firms	0.351	0.415	0.358	0.564
average markup	1.521	1.529	1.545	1.543
std. dev. markup	0.219	0.262	0.247	0.292
average market share of top 4 firms	0.876	0.865	0.873	0.832
std. dev. number of markets	4.755	5.785	7.082	10.444
labor share	0.643	0.678	0.656	0.677

Notes: The estimation is conducted using a simulated method of moments (SMM) approach. This table reports the estimated and actual moments for the two estimations: early period (2002-2005) and late period (2013-2016).

can charge higher markups due to reduced competition in local markets. Turning to parameters governing the costs of geographic expansion, I find that there is a decrease in the expansion cost scale parameters (χ_H and χ_L) from 223.21 to 91.20 and 681.76 to 152.47 respectively, though this effect is opposed by the fall in the large firms' expansion cost convexity (ϕ) from 3.76 to 3.26. The changes in χ_H and χ_L reduce the cost of geographic expansion, whereas the fall in ϕ allows expansion to be more concentrated across firms and increases expansion costs on average. The entrepreneurs' innovation cost scale ψ increases from 0.94 to 1.14 and is coupled with a decrease in the probability of being a high-type productivity firm on entry, hence pointing to a decline in business dynamism in Canada.[33]

5 Quantitative Experiments

I rely on the early (2002-2005) and late (2013-2016) period estimations shown in Section 4 to conduct a quantitative application of my framework. In particular, I use the two estimated

[33]This is in line with the overall decline in business dynamism documented for the U.S., as in Decker, Haltiwanger, Jarmin, and Miranda (2016).

stationary equilibria to disentangle the economic mechanisms underlying the increase in firms' geographic expansion and the fall in local market concentration observed in Canada over this period. To do this, I compare the estimated early period economy against counterfactual economies in which I set the values of selected parameters to their late period estimates. These experiments show how the economy would have looked if there was a change only in a subset of the parameters, which in turn allows me to understand the primary drivers of the increase in firms' geographic expansion in Canada. They are also helpful in understanding the sources of the time trends in important statistics such as local product market concentration, average markup, and labor share in Canada. I discuss these experiments in detail below.

5.1 Geographic Expansion Cost

Table 4: Changing the Geographic Expansion Cost Parameters

	Early	Late χ_H, ϕ	% change	Late χ_L, ϕ	% change	Late χ_H, χ_L, ϕ	% change	Late	% change
sales-weighted avg. local hhi	0.305	0.303	-0.695	0.292	-4.218	0.301	-1.048	0.297	-2.560
firm exit rate (%)	3.137	3.173	1.140	3.126	-0.362	3.123	-0.471	3.708	18.199
avg. no. of markets of H firms	4.108	4.207	2.413	1.246	-69.654	2.952	-28.143	7.241	76.288
avg. no. of markets of L firms	1.158	1.029	-11.141	1.613	39.285	1.313	13.444	1.206	4.151
expansion expenditure ratio	0.047	0.052	10.678	0.044	-7.697	0.050	5.483	0.072	53.180
proportion of HH firms	0.159	0.160	0.675	0.099	-37.853	0.143	-10.066	0.148	-6.788
proportion of HL firms	0.270	0.230	-14.771	0.325	20.463	0.292	8.393	0.247	-8.269
proportion of LL firms	0.415	0.443	6.789	0.418	0.904	0.410	-1.143	0.564	35.944
avg. markup	1.529	1.529	0.006	1.521	-0.535	1.530	0.018	1.543	0.881
std. dev. markup	0.262	0.276	5.100	0.315	20.127	0.283	7.740	0.292	11.269
avg. market share, top 4 firms	0.865	0.845	-2.308	0.778	-10.081	0.834	-3.538	0.832	-3.801
std. dev. number of markets	5.785	9.869	70.596	1.364	-76.419	5.517	-4.635	10.444	80.538
labor share	0.678	0.680	0.396	0.693	2.243	0.682	0.599	0.677	-0.056
total output	3.110	3.062	-1.530	2.909	-6.450	3.037	-2.336	6.970	124.115
wage	2.107	2.083	-1.140	2.016	-4.352	2.071	-1.751	4.720	123.990
total expansion expenditure	0.147	0.160	8.984	0.127	-13.650	0.151	3.019	0.504	243.301
total entry expenditure	0.039	0.040	0.260	0.058	46.466	0.047	18.287	0.069	75.831
total consumption	2.924	2.863	-2.082	2.725	-6.802	2.840	-2.883	6.397	118.785
consumption share	0.940	0.935	-0.560	0.937	-0.377	0.935	-0.560	0.918	-2.378

This table reports the changes in model moments when setting the geographic expansion cost parameters to their estimated levels in the late period while keeping other parameters fixed at their estimated values in the early period. 'Early' denotes the early period stationary equilibrium, and 'Late' denotes the late period stationary equilibrium.

In the first set of experiments, I set the parameters governing the firms' geographic expansion costs to their late period levels. The results are shown in table 4, where the first column displays the benchmark values of chosen model moments in the early period, and the last two columns show the results of changing all the parameters to their late period values to provide comparisons. All the remaining columns show how the moments change in each exercise. In columns 2 and 3, I show the model moments and % changes (as compared to the early period's stationary equilibrium) associated with changing the high-type expansion efficiency firms' expansion cost scale (χ_H) and large firms' expansion cost convexity (ϕ) to their late period values. Unsurprisingly, the fall in χ_H increases the average number of markets for the high-type expansion efficiency firms, but this increase is reduced due to a fall in ϕ. This leads to a slight increase in the average number of markets of the high-type expansion efficiency firms and expansion expenditure ratio, but a 11.14% fall in the average number of markets of the low-type expansion efficiency firms. This is associated with a fall in the average local HHI by 0.69%, but it is insufficient in generating the overall fall of 2.56% observed throughout the period. Total output and wages fall in this economy as compared to the benchmark economy, and this change is also insufficient to match the 0.88% increase in markups observed over this period.

Columns 4 and 5 highlight the model moments and % changes (as compared to the early period's stationary equilibrium) associated with changing the low-type expansion efficiency firms' expansion cost scale (χ_L) and large firms' expansion cost convexity (ϕ) to their late period values. This leads to an increase in the average number of markets of the low-type expansion efficiency firms even though the falls in χ_L and ϕ work in opposite directions. This change over-explains the fall in local HHI and cannot explain the increase in markups. Total output, wages and consumption fall instead of increasing as observed in the late period.

Finally, columns 6 and 7 show the model moments and % changes (as compared to the early period's stationary equilibrium) associated with changing the high-type expansion efficiency firms' expansion cost scale (χ_H), low-type expansion efficiency firms' expansion cost scale (χ_L) and large firms' expansion cost convexity (ϕ) to their late period values. Since the fall in χ_L is much more stark in comparison to the fall in χ_H, this combination of parameters cannot generate the increase in the average number of markets for the high-type or low-type expansion efficiency large firms. The increase in the expansion expenditure ratio is much smaller than the total increase, the fall in local HHI is less than half the total fall in local HHI, and the increase in average markups in a fraction of the total increase shown in the last column. Since the large firms have lower incentives to enter local markets due to a flat increase in markups, the high-type large firms are dis-incentivized to expand as compared to the low-type firms in this counterfactual economy. To sum up, these experiments highlight that a change in the costs of geographic expansion are insufficient to explain the increase in firms' geographic expansion and the fall in local market concentration in Canada from 2001 to 2018.

5.2 Firm Productivities

Table 5: Changing the Productivity Parameters

	Early	Late A_H/A_L	% change	Late A_H/A_c	% change	Late $A_H/A_L, A_H/A_c$	% change	Late	% change
sales-weighted avg. local hhi	0.305	0.305	0.005	0.313	2.703	0.313	2.706	0.297	-2.560
firm exit rate (%)	3.137	3.137	-0.002	3.137	0.002	3.137	-0.002	3.708	18.199
avg. no. of markets of H firms	4.108	4.115	0.169	4.035	-1.765	4.043	-1.571	7.241	76.288
avg. no. of markets of L firms	1.158	1.158	0.008	1.157	-0.055	1.157	-0.044	1.206	4.151
expansion expenditure ratio	0.047	0.047	0.021	0.048	2.250	0.048	2.276	0.072	53.180
proportion of HH firms	0.159	0.159	0.069	0.159	0.200	0.160	0.325	0.148	-6.788
proportion of HL firms	0.270	0.270	0.094	0.269	-0.088	0.270	0.028	0.247	-8.269
proportion of LL firms	0.415	0.414	-0.037	0.415	-0.019	0.414	-0.079	0.564	35.944
avg. markup	1.529	1.529	-0.002	1.551	1.411	1.551	1.406	1.543	0.881
std. dev. markup	0.262	0.262	-0.065	0.274	4.596	0.274	4.506	0.292	11.269
avg. market share, top 4 firms	0.865	0.865	0.028	0.873	0.968	0.873	1.004	0.832	-3.801
std. dev. number of markets	5.785	5.856	1.228	5.893	1.868	5.854	1.188	10.444	80.538
labor share	0.678	0.678	-0.003	0.669	-1.216	0.669	-1.219	0.677	-0.056
total output	3.110	6.987	124.654	3.111	0.042	6.990	124.763	6.970	124.115
wage	2.107	4.734	124.647	2.083	-1.175	4.679	122.024	4.720	123.990
total expansion expenditure	0.147	0.330	124.701	0.150	2.293	0.337	129.879	0.504	243.301
total entry expenditure	0.039	0.089	124.521	0.041	4.037	0.092	133.541	0.069	75.831
total consumption	2.924	6.568	124.654	2.920	-0.125	6.561	124.388	6.397	118.785
consumption share	0.940	0.940	-0.000	0.939	-0.167	0.939	-0.167	0.918	-2.378

This table reports the changes in model moments when setting the productivity parameters to their estimated levels in the late period while keeping other parameters fixed at their estimated values in the early period. 'Early' denotes the early period stationary equilibrium, and 'Late' denotes the late period stationary equilibrium.

In the next set of experiments, I set the parameters governing the firms' productivities to their late period levels. The results are shown in table 5, where the first column displays the benchmark values of chosen model moments in the early period, and the last two columns show the results of changing all the parameters to their late period values. All the remaining columns show how the moments change in each exercise. Columns 2 and 3 show the results when only the relative productivity of high-type productivity firms with respect to the low-type productivity firms (A_H/A_L) is increased from 2.8706 to the late period value of 6.4477. This leads to a small increase in the average number of markets for the high-type expansion efficiency firms, and a marginal increase in the average number of markets for the low-type expansion efficiency firms, both leading to a small increase in the expansion expenditure ratio. All these changes are a small fraction of the total changes observed in the last column, showing that the increase in A_H/A_L is not a major factor explaining the observed trends, though it is able to match the total increase in output and wages. In addition, this increase in A_H/A_L has marginal effects on the measures of market power. Columns 4 and 5 highlight the results for when I increase the relative productivity of high-type productivity firms with respect to the fringe (A_H/A_c) to its late period value, while holding all other parameters at their early period values. This translates to a lower productivity of the fringe A_c, hence increasing markups within local markets. This in turn leads to lower incentives for geographic expansion, hence reducing the average number of markets of both types of firms and increasing average local HHI. The average market share of top 4 firms also increases much more than the value observed in the late period.

Finally, in columns 6 and 7. I show the results obtained when I increase A_H/A_L and A_H/A_c to their late period values. Average markups increase by 1.4%, dis-incentivizing the large firms to gain profits by expanding to new geographic markets. Hence, local HHI increases and average market share of top 4 firms increases as the high-type productivity firms become more productive. This results in an increase in total output and consumption, both of which closely match the actual increase observed in the late period.

All these counterfactual experiments highlight that the change in productivity parameters are unable to match the increase in the average number of markets for firms in Canada, but an increase in the relative productivity of the high-type productivity firms with respect to the low-type productivity firms and fringe help in explaining the increase in markups observed over this period in Canada.[34] The estimated decline in the relative productivity of the low-type productivity firms can be linked to several potential mechanisms such as advantages of big firms in accessing and utilizing data, changes in advertising and brand value, increasing innovation and imitation costs, or decline in knowledge spillovers. This highlights the need to dig deeper into the causes and implications of falling productivity of such firms, since lower productivity impacts these firms' incentives for geographic expansion.

Table 6: Changing Other Parameters

	Early	Late ψ, λ	% change	Late σ	% change	Late δ, δ_f	% change	Late	% change
sales-weighted avg. local hhi	0.305	0.304	-0.184	0.294	-3.610	0.297	-2.654	0.297	-2.560
firm exit rate (%)	3.137	3.164	0.847	3.139	0.054	3.755	19.689	3.708	18.199
avg. no. of markets of H firms	4.108	17.145	317.409	4.119	0.279	2.901	-29.366	7.241	76.288
avg. no. of markets of L firms	1.158	1.109	-4.237	1.157	-0.033	1.124	-2.882	1.206	4.151
expansion expenditure ratio	0.047	0.053	13.039	0.047	-0.579	0.055	16.849	0.072	53.180
proportion of HH firms	0.159	0.111	-30.037	0.156	-1.683	0.190	19.322	0.148	-6.788
proportion of HL firms	0.270	0.226	-16.004	0.269	-0.245	0.264	-2.209	0.247	-8.269
proportion of LL firms	0.415	0.618	49.033	0.417	0.500	0.397	-4.337	0.564	35.944
avg. markup	1.529	1.517	-0.838	1.527	-0.125	1.524	-0.327	1.543	0.881
std. dev. markup	0.262	0.226	-13.705	0.264	0.765	0.297	13.239	0.292	11.269
avg. market share, top 4 firms	0.865	0.911	5.303	0.850	-1.723	0.809	-6.431	0.832	-3.801
std. dev. number of markets	5.785	29.864	416.242	8.718	50.698	2.893	-49.998	10.444	80.538
labor share	0.678	0.677	-0.134	0.679	0.265	0.687	1.415	0.677	-0.056
total output	3.110	3.224	3.677	3.141	1.012	2.980	-4.171	6.970	124.115
wage	2.107	2.182	3.538	2.134	1.280	2.048	-2.815	4.720	123.990
total expansion expenditure	0.147	0.172	17.195	0.147	0.428	0.164	11.976	0.504	243.301
total entry expenditure	0.039	0.017	-58.138	0.040	0.551	0.055	40.350	0.069	75.831
total consumption	2.924	3.036	3.832	2.954	1.048	2.761	-5.582	6.397	118.785
consumption share	0.940	0.942	0.150	0.940	0.035	0.926	-1.472	0.918	-2.378

This table reports the changes in model moments when setting the parameters of interest to their estimated levels in the late period while keeping other parameters fixed at their estimated values in the early period. 'Early' denotes the early period stationary equilibrium, and 'Late' denotes the late period stationary equilibrium.

5.3 Other Parameters

In these set of experiments, I first analyze the effects of changing parameters related to firms' entry ($\psi, \lambda_{HH}, \lambda_{HL}, \lambda_{LH}, \lambda_{LL}$), The results are shown in columns 2 and 3 of table 6. Compared to the values in the early period, the increase in entrepreneurs' innovation cost scale (ψ) and the fall in the probability of being a high-type productivity firm on entry ($\lambda_{HH}, \lambda_{HL}$) leads to a very large push in the average number of markets of the high-type expansion efficiency firms. The increase observed in the average number of markets of the high-type expansion efficiency firms is over-explained, while the increase in the average number of markets of the low-type expansion efficiency firms, and the increase in average markups cannot be explained by these changes. The increase in expansion of the high-type expansion efficiency firms leads to an increase in the expansion expenditure ratio and a fall in the local HHI, but none of these changes are able to match the numbers in the late period well. As expected, the total entry expenditure falls whereas total output has a modest increase. There is a large rise in the standard deviation of markets as the divergence between high-type and low-type expansion efficiency firms increases in terms of the number of geographic markets.

[34] This is consistent with the literature regarding the decline in business dynamism and "winner-takes-most" dynamics in papers such as Decker et al. (2016) and Autor, Dorn, Katz, Patterson, and Van Reenen (2020).

Columns 4 and 5 of table 6 show the effects of decreasing only the elasticity of substitution between varieties in a market (σ) to its late period value. A reduction in the elasticity of substitution indicates greater product differentiation or a stronger preference among consumers for unique varieties, thereby enhancing the market power of firms selling these differentiated products. The average markup decreases slightly by 0.13% due to the general-equilibrium changes in the stationary distribution (μ_{N_H,N_L,ω_k}) towards low markup markets. The increase in market power favors the large firms having higher productivity and expansion efficiency. This encourages them to enter more geographic markets, hence increasing the average number of markets for these high-type firms and reducing the average local HHI in local markets. The low-type large firms are discouraged to enter more markets due to the increased market power, hence reducing the total expansion expenditure ratio in this economy. This experiment reveals the importance of market power in affecting firms' incentives to expand in more geographic markets. If the elasticity of substitution had not fallen, there would be less incentive to expand to more markets by more productive firms, which would imply that local markets would have had higher product market concentration.

Finally, columns 6 and 7 show of table 6 show the results for changing the firm exit parameters δ, δ_f to their late period values. As seen in table 2, exogenous market exit rate δ increases from 0.0058 to 0.0348 and exogenous firm shut-down rate δ_f decreases from 0.0267 to 0.0096. The increase in market exit rate discourages firms from entering more local markets, hence the firms' average number of geographic markets falls. These changes are able to match the firm exit rate moment very well, but they are unable to match the increase in average markup, output, and wages. Hence, these set of parameters are not helpful in explaining the increase in geographic expansion of firms in Canada.

6 Policy Experiments

The model features many inefficiencies, which make the decentralized equilibria suboptimal. For example, even though firms' geographic expansion is desirable as it is closely linked to declining local product market concentration and increasing growth, firms have an incentive to over-expand due to the business stealing externality in the model. Firms are always better off when entering a new geographic market, but do not take into account the negative effects of cannibalizing other incumbent firms' profits when entering the market. This leads to over-expenditure on geographic expansion and a reduction in consumption.[35] In addition, firms' geographic expansion offers the potential to boost the average productivity of geographic markets by altering the composition of firms within them. More firms are beneficial due to households' preference for variety, and an increase in geographic expansion can enhance local competition, reduce markups, and raise the labor share in local markets. While these are all advantages of geographic expansion, firms do not account for these benefits when making expansion decisions, as they only consider the gains to their own value. These inefficiencies of the decentralized equilibria can be removed by imposing suitable taxes or subsidies on firms' geographic expansion costs in the economy. In this section, I highlight the role of such taxes and subsidies on firms' geographic expansion costs that can be implemented by a policy maker to increase consumption (equivalent to welfare) in the economy.[36]

Table 7: Policy Experiments

Policy	Subsidy on HH firms	Subsidy on HL firms	Subsidy on LH firms	Subsidy on LL firms	Consumption
Targeted	17.14%	-100%	-100%	-100%	6.4875
Uniform	18.18%	18.18%	18.18%	18.18%	6.4255

6.1 Targeted Policy

Since the large firms engaging in geographic expansion differ in terms of their productivities and expansion efficiencies, I first consider the case in which the policy maker can target firms and impose type-specific taxes or subsidies on the firms' expansion costs to increase consumption in the economy. Row 1 of table 7 shows the taxes/subsidies and aggregate consumption attained in case of this targeted policy. The policy maker subsidizes the high-type productivity and high-type expansion efficiency (HH-type) firms by imposing a 17.14% subsidy rate on their

[35] A similar argument can be made for firms entering the economy: there may be over-entry as entrepreneurs don't consider the business stealing externality when deciding to enter the economy. This implies that the constrained optimal policy would be to tax entrepreneurship if the objective of the planner is to maximize consumption.

[36] For each of the exercises highlighted below, I take the stationary equilibrium of the late period and find the policies that maximize consumption in the economy.

geographic expansion, and taxes all other types of firms.[37] This is because the geographic expansion of the HH-type firms is more productive and efficient than the expansion of all other types of firms. A HH-type firm uses less resources to expand, and delivers more once it enters a new geographic market. On the other hand, the costs of expansion outweigh the benefits from expanding into more geographic markets for all other types of firms. Hence, the economy can boost consumption by limiting the expansion of these three types of firms (HL-type, LH-type, LL-type) through taxes on their geographic expansion costs.

6.2 Uniform Policy

It would be ideal to observe firm types, but it may be difficult to do so in reality. If the policy maker cannot observe the type of the firm, i.e., she does not know whether the firm is a high-type or a low-type productivity firm, and does not know whether the firm is a high-type or a low-type expansion efficiency firm, she must choose a single tax or subsidy rate that will be imposed on all the four types of firms (HH-type, HL-type, LH-type, LL-type) to increase consumption in the economy. Row 2 of table 7 shows the subsidy rate and aggregate consumption attained in case of this uniform policy. Interestingly, the policy maker subsidizes the geographic expansion costs of all firms, and this leads to a lower consumption than what can be obtained in the case of the targeted policy. A uniform subsidy would incentivize all large firms to expand geographically, and not just those who are 'better' at it, hence limiting the overall effectiveness of the policy. This highlights the importance of being able to discern firm types,[38] in order to implement a targeted policy.

7 Conclusion

Firm expansion into multiple geographic markets is a crucial margin of firm growth, but its general equilibrium aspects have been understudied. This margin has assumed greater importance

[37]A subsidy of -100% represents maximum taxation, as the subsidy rate in this experiment is constrained to range from -100% to 100%.

[38]A policy maker could use statistics such as firms' existing number of markets, market shares and expenditures on geographic expansion as signals of firm types.

with the rise in the consumption share of non-tradables relative to tradable goods and services. In addition, firms' geographic expansion is closely tied to measures of local competition and innovation.

In this paper, I overcome the challenges associated with research on the geography of firm dynamics by using detailed Canadian microdata and a novel empirical methodology to infer firms' number of geographic markets over time. I document a set of facts related to the nature of firms' geographic expansion and find that there has been a heterogeneous increase in the geographic presence of firms in Canada - the average number of geographic markets of firms has increased, but the increase has been larger for more 'expansion-efficient' firms. In addition, there has been a decline in local product market concentration as shown by the decline in the sales-weighted average local HHI, hence suggesting that firms' geographic expansion has increased competition in local markets in Canada. I use this motivating evidence to build a new model of firms' geographic expansion to analyze the sources and aggregate implications of the geographic expansion of firms. In my dynamic general equilibrium model, firms are heterogeneous in their productivity and geographic expansion efficiencies. They spend resources to potentially expand to new geographic markets (which differ in size and feature oligopolistic competition).

At the aggregate level, the model is consistent with the observed trends in the geography of firm dynamics. At the market level, it can account for the decrease in local market concentration due to firms' geographic expansion and offer realistic market dynamics with firm heterogeneity and endogenous firm entry. At the firm level, it captures the strategic interactions between firms in geographic product markets (that determines non-degenerate relative output, profits, labor demand etc. distributions endogenously). It also captures a firm's dynamic decisions to enter more geographic markets over time, optimally chosen in response to its existing number of geographic markets, the aggregate expansion rates chosen by other firms in the economy, and the general equilibrium distribution of new markets that the firm can expand into, which are heterogeneous in terms of the number and type of firms present in them, as well as in size.

Counterfactual experiments using the estimated model highlight the role of factors such as the innovation costs of entrepreneurs, geographic expansion costs of incumbent firms, firm productivities and the elasticity of substitution between local varieties in affecting the geographic expansion of firms. For example, I find that a fall in the expansion costs of incumbent firms is unable to generate the increase in the number of geographic markets by itself, but the rise in entrepreneurs' entry costs, the fall in the probability of being a high-type productivity firm and the fall in the elasticity of substitution between local varieties between the early and late periods are crucial to match the observed trends of increased geographic expansion and reduced local product market concentration in Canada.[39]

This paper also highlights the role of policy in the face of inefficiency of the decentralized equilibria. The constrained optimal policy to deal with the externalities faced by the firms is to impose suitable taxes or subsidies on their geographic expansion costs to increase welfare in the economy. The constrained optimal targeted policy is to subsidize the geographic expansion costs of the high-type productivity and high-type expansion efficiency firms since they are more productive at production and expansion. In contrast, a policy maker should tax the geographic expansion costs of all other types of firms since their costs of expansion outweigh the benefits to the economy. The uniform constrained optimal policy is to subsidize the geographic expansion costs of all firms.

The analysis in this paper enhances our understanding of firm dynamics and enlarges the scope of the firm dynamics literature to include firms' decisions related to geographic expansion. This research has direct implications for antitrust policy, as firms' geographic expansion is closely tied to product market concentration outcomes that have been a source of much debate and discussion for economists and policy makers in the recent decades. Firms' expansion into new local markets is expected to have major effects on incentives for firm innovation, as well as on local labor markets. I expect future research on these topics to be both promising and valuable.

[39] This paper remains agnostic about the underlying forces (such as policy, technology, and regulation) that may have driven these changes, and I leave a more detailed investigation of these factors to future research.

References

Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies*, *68*(3), 467–492.

Akcigit, U., & Ates, S. T. (2021). Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, *13*(1), 257–298.

Argente, D., Baslandze, S., Hanley, D., & Moreira, S. (2020). Patents to products: Product innovation and firm dynamics.

Argente, D., Fitzgerald, D., Moreira, S., & Priolo, A. (2021). How do firms build market share? *Available at SSRN 3831706*.

Atkeson, A., & Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, *98*(5), 1998–2031.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Autor, D., Patterson, C., & Van Reenen, J. (2023). *Local and national concentration trends in jobs and sales: The role of structural transformation* (Tech. Rep.). National Bureau of Economic Research.

Benkard, C. L., Yurukoglu, A., & Zhang, A. L. (2021). *Concentration in product markets* (Tech. Rep.). National Bureau of Economic Research.

Cao, D., Hyatt, H. R., Mukoyama, T., & Sager, E. (2017). Firm growth through new establishments. *Available at SSRN 3361451*.

Carlson, M., & Mitchener, K. J. (2009). Branch banking as a device for discipline: Competition and bank survivorship during the great depression. *Journal of Political Economy*, *117*(2), 165–210.

Cavenaile, L., Celik, M. A., & Tian, X. (2019). Are markups too high? competition, strategic innovation, and industry dynamics. *Competition, Strategic Innovation, and Industry Dynamics (August 1, 2019)*.

Cavenaile, L., Celik, M. A., & Tian, X. (2021). The dynamic effects of antitrust policy on growth and welfare. *Journal of Monetary Economics*, *121*, 42–59.

Covarrubias, M., Gutiérrez, G., & Philippon, T. (2020). From good to bad concentration? us industries over the past 30 years. *NBER Macroeconomics Annual*, *34*(1), 1–46.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2016). Declining business dynamism: What we know and the way forward. *American Economic Review*, *106*(5), 203–207.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2020). Changing business dynamism and productivity: Shocks versus responsiveness. *American Economic Review*, *110*(12), 3952–3990.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

Franco, S. (2024). *Output market power and spatial misallocation.* The University of Chicago.

Ganapati, S., et al. (2018). *The modern wholesaler: Global sourcing, domestic distribution, and scale economies* (Tech. Rep.).

Gutiérrez, G., & Philippon, T. (2016). *Investment-less growth: An empirical investigation* (Tech. Rep.). National Bureau of Economic Research.

Gutiérrez, G., & Philippon, T. (2017). *Declining competition and investment in the us* (Tech. Rep.). National Bureau of Economic Research.

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica*, *79*(1), 253–302.

Hortaçsu, A., & Syverson, C. (2015). The ongoing evolution of us retail: A format tug-of-war. *Journal of Economic Perspectives*, *29*(4), 89–112.

Hsieh, C.-T., & Rossi-Hansberg, E. (2019). *The industrial revolution in services (no. w25968).* National Bureau of Economic Research.

Hsieh, C.-T., & Rossi-Hansberg, E. (2023). The industrial revolution in services. *Journal of Political Economy Macroeconomics*, *1*(1), 3-42. Retrieved from https://doi.org/10.1086/723009 doi: 10.1086/723009

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

Jiang, X. (2023). Information and communication technology and firm geographic expansion.

Kleinman, B. (2022). *Wage inequality and the spatial expansion of firms* (Tech. Rep.). Princeton Working Paper.

Klette, T. J., & Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of political economy*, *112*(5), 986–1018.

Loecker, J. D., & Warzynski, F. (2012). Markups and firm-level export status. *American economic review*, *102*(6), 2437–2471.

Oberfield, E., Rossi-Hansberg, E., Sarte, P.-D., & Trachter, N. (2023). Plants in space. *Journal of Political Economy*.

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

Rossi-Hansberg, E., Sarte, P.-D., & Trachter, N. (2021). Diverging trends in national and local concentration. *NBER Macroeconomics Annual*, *35*(1), 115–150.

Syverson, C. (2008). Markets: Ready-mixed concrete. *Journal of Economic Perspectives*, *22*(1), 217–234.

Xi, X. (2023). Multi-establishment firms, misallocation, and productivity. *Journal of Economic Dynamics and Control*, *154*, 104705.

8 Appendices

8.1 Empirical Methodology

In this section, I outline my empirical methodology for determining the number of local geographic markets for firms in Canada.

Canada is divided into 13 provinces and territories, which are further divided into 293 census divisions (CDs) and 5162 census subdivisions (CSDs) according to the 2016 census.[40] For example, the City of Toronto is a CD which consists of multiple CSDs such as Toronto, Mississauga, Markham and Vaughan. I observe the firm's province, census division, census subdivision and postal code, but take the census subdivision (CSD) to be the measure of a geographic market in this analysis.

In the data, a firm is defined at the enterprise level (i.e., at the level of the entity filing taxes in Canada), so I only observe the location of its main headquarters. Hence, I utilize the joint information on firm-worker matches and the reported locations of firms and workers to infer the possible locations of all establishments of the firms. The CEEDD data masks the names of firms, workers and locations for confidentiality, but an example of my empirical methodology is as follows. If the data shows that a firm's headquarters is in Toronto, Toronto is one of its geographic markets. Additionally, if I observe employees matched to this firm in Ottawa, it suggests that Ottawa may also be one of its geographic markets. For every firm, I first derive all such possible geographic markets based on firm-worker matches and their respective locations. Firm-worker matches where the geography of a worker is not the same as the reported geography of the firm headquarters are prospective new establishments at that geography level (CSD in this case). Each match gives me the number of employees in each of these new geographic markets. Then, I sort each firm's geographic markets based on their number of matched employees in decreasing order (starting from the market with the highest number of employees). I calculate the distances between each market and previous bigger markets (having a greater number of matched employees) and re-assign the matched workers of the current market to the previous

[40]See https://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/chap12-eng.cfm for more information.

bigger market if the distance between the two markets is less than the threshold distance (50km for the analysis presented in this paper).[41] Even though the 'work from home' effect is not a major concern since the data is pre-covid, I run some robustness checks to account for it. For example, I treat another market as a new geographic market only if there are 'enough' number of firm-worker matches in that market. In the analysis presented in this paper, I do not consider CSDs with only 1 matched worker.[42]

Hence, the firm-worker matches and detailed geographic information on firms and their workers allow me to infer the firms' geographic markets for each year. Even though this methodology allows me to get information on all local geographic markets of the firms, I only use the number of geographic markets of firms for the main analysis in this paper. The information at the geographic market level (such as those on the average market shares within local markets) is used to construct measures of local product market concentration as described in appendix 8.6.1 below.

Go back to section 2.1.

8.2 Value Function of a Low-type Productivity Large Firm

The value function for a firm of type (L, i_χ) (where L denotes its productivity type and $i_\chi \in \{H, L\}$ denotes its expansion efficiency type) is:

[41] The threshold distance is chosen as a conservative measure of the average commuting distance in Canada. More information on the average commuting distances and times can be found at https://www150.statcan.gc.ca/n1/daily-quotidien/190225/dq190225a-eng.htm. This exercise is equivalent to constructing a measure of commuting zones within Canada.

[42] Changing the threshold for how many matched workers in a new location are required for it to be called a new geographic market does not change the trend of increasing average number of geographic markets.

$$
\begin{aligned}
&rV_{L,i_\chi}(\{\Theta_k\}_{k=1}^K) \\
&= \max_{X_{L,i_\chi}} \sum_{k=1}^{K} \pi_L(N_{Hk}, N_{Lk}, \omega_k)Y - \chi_{i_\chi}(\frac{X_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)}{K})^\phi K^\alpha Y \\
&+ X_{L,i_\chi}[\mathbb{E}\, V_{L,i_\chi}(\{\Theta_k\}_{k=1}^{K+1}) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \delta_f[0 - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] + \delta \sum_{l=1}^{K}[V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K - \Theta_l) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_H \sum_{l=1}^{K}[V_{L,i_\chi}((\{\Theta_k\}_{k=1}^K - \Theta_l) \cup (\min(N_{Hl}+1, \bar{N} - N_{Ll}), N_{Ll}, \omega_l)) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_L \sum_{l=1}^{K}[V_{L,i_\chi}((\{\Theta_k\}_{k=1}^K - \Theta_l) \cup (N_{Hl}, \min(N_{Ll}+1, \bar{N} - N_{Hl}), \omega_l)) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta+\delta_f) \sum_{l=1}^{K} N_{Hl}[V_{L,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl}-1, N_{Ll}, \omega_l)) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta+\delta_f) \sum_{l=1}^{K} (N_{Ll}-1)[V_{L,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl}, N_{Ll}-1, \omega_l)) - V_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \dot{V}_{L,i_\chi}(\{\Theta_k\}_{k=1}^K)
\end{aligned}
\tag{30}
$$

Go back to section 3.2.

8.3 Proof of Theorem 1

The value function of a high-type productivity large firm is:

$$
\begin{aligned}
& rV_{H,i_\chi}(\{\Theta_k\}_{k=1}^K) \\
&= \max_{x_{H,i_\chi}} \sum_{k=1}^{K} \pi_H(\Theta_k)Y - \chi_{i_\chi}(x_{H,i_\chi}(\{\Theta_k\}_{k=1}^K))^{\phi} KY \\
&+ x_{H,i_\chi} K[\mathbb{E}\, V_{H,i_\chi}(\{\Theta_k\}_{k=1}^{K+1}) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \delta_f[0 - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] + \delta \sum_{l=1}^{K} [V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K - \{\Theta_l\}) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_H \sum_{l=1}^{K} [V_{H,i_\chi}((\{\Theta_k\}_{k=1}^K - \{\Theta_l\}) \cup (\min(N_{Hl}+1, \bar{N} - N_{Ll}), N_{Ll}, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ e_L \sum_{l=1}^{K} [V_{H,i_\chi}((\{\Theta_k\}_{k=1}^K - \{\Theta_l\}) \cup (N_{Hl}, \min(N_{Ll}+1, \bar{N} - N_{Hl}), \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta + \delta_f) \sum_{l=1}^{K} N_{Ll}[V_{H,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl}, N_{Ll} - 1, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ (\delta + \delta_f) \sum_{l=1}^{K} (N_{Hl} - 1)[V_{H,i_\chi}((\{\Theta\}_{k=1}^K - \Theta_l) \cup (N_{Hl} - 1, N_{Ll}, \omega_l)) - V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)] \\
&+ \dot{V}_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)
\end{aligned} \tag{31}
$$

I guess and verify that in a stationary equilibrium,

$V_{H,i_\chi}(\{\Theta_k\}_{k=1}^K) = \sum_{k=1}^{K} v_{H,i_\chi}(\{\Theta_k\})Y = \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)Y$

Hence, $\sum_{k=1}^{K} \pi_H(\{\Theta_k\})Y = \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \pi_H(\Theta)Y$,

$\dot{V}_{H,i_\chi}(\{\Theta_k\}_{k=1}^K) = gV_{H,i_\chi}(\{\Theta_k\}_{k=1}^K) = g \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)Y$,

$\mathbb{E}\, V_{H,i_\chi}(\{\Theta_k\}_{k=1}^{K+1}) = \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)Y$

$+ \sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)Y$

$$\Longrightarrow r \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y$$

$$\begin{aligned}
&= \max_{x_{H,i_\chi}} \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \pi_H(\Theta) Y - \chi_{i_\chi} \left(\frac{X_{H,i_\chi}(\{\Theta_k\}_{k=1}^K)}{K}\right)^\phi KY \\
&+ X_{H,i_\chi} [\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l) Y] \\
&+ \delta_f [0 - \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y] \\
&+ \delta \sum_{l=1}^{K} [-v(\Theta_l) Y] \\
&+ e_H \sum_{l=1}^{K} [-v(\Theta_l) Y + v(\min(N_{Hl}+1, \bar{N} - N_{Ll}), N_{Ll}, \omega_l) Y] \\
&+ e_L \sum_{l=1}^{K} [-v(\Theta_l) Y + v(N_{Hl}, \min(N_{Ll}+1, \bar{N} - N_{Hl}), \omega_l) Y] \\
&+ (\delta + \delta_f) \sum_{l=1}^{K} N_{Ll} [\big(- v(\Theta_l) Y + v(N_{Hl}, N_{Ll}-1, \omega_l) Y \big)] \\
&+ (\delta + \delta_f) \sum_{l=1}^{K} (N_{Hl}-1) [\big(- v(\Theta_l) Y + v(N_{Hl}-1, N_{Ll}, \omega_l) Y \big)] \\
&+ g \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y
\end{aligned} \tag{32}$$

Since $\sum_{k=1}^{K} v_{H,i_\chi}(\{\Theta_k\}) Y = \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y$, I can replace
$\sum_{l=1}^{K} v(\Theta_l) Y = \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y$
Also, $e_H \sum_{l=1}^{K} v(\min(N_{Hl}+1, \bar{N} - N_{Ll}), N_{Ll}, \omega_l) Y =$
$e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N} - N_L), N_L, \omega_k) Y,$
$e_L \sum_{l=1}^{K} v(N_{Hl}, \min(N_{Ll}+1, \bar{N} - N_{Hl}), \omega_l) Y =$
$e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N} - N_H), \omega_k) Y,$
$(\delta + \delta_f) \sum_{l=1}^{K} N_{Ll} [-v(\Theta_l) Y + v(N_{Hl}, N_{Ll}-1, \omega_l) Y] =$
$(\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} N_L [-m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y + m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_{Hl}, N_{Ll}-1, \omega_l) Y],$
$(\delta + \delta_f) \sum_{l=1}^{K} (N_{Hl}-1) [-v(\Theta_l) Y + v(N_{Hl}-1, N_{Ll}, \omega_l) Y] =$
$(\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} N_L [-m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) Y + m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_{Hl}-1, N_{Ll}, \omega_l) Y]$

Using the above to convert $\sum_{l=1}^{K}$ terms to $m_{H,i_\chi}()v_{H,i_\chi}()$ terms and dividing by Y,

$$
\begin{aligned}
& r \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
&= \max_{x_{H,i_\chi}} \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)\pi_H(\Theta) - \chi_{i_\chi} X_{H,i_\chi}^{\phi}(\{\Theta_k\}_{k=1}^{K}) K^{1-\phi} \\
&+ X_{H,i_\chi}[\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)] \\
&- \delta_f \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) - \delta \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
&- e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
&+ e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N}-N_L), N_L, \omega_k) \\
&- e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
&+ e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N}-N_H), \omega_k) \\
&+ (\delta+\delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} N_L[-m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) + m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, N_L-1, \omega_k)] \\
&+ (\delta+\delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} (N_H-1)[-m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) + m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H-1, N_L, \omega_k)] \\
&+ g \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)
\end{aligned}
\tag{33}
$$

The first order condition with respect to x_{H,i_χ} is

$$
-\phi \chi_{i_\chi} X_{H,i_\chi}^{\phi-1} K^{1-\phi} + \sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l) = 0 \tag{34}
$$

$$\implies x_{H,i_\chi} = \frac{X_{H,i_\chi}}{K} = \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{1}{\phi-1}} \tag{35}$$

At the optimum, equation 33 can be written as

$$[r + \delta_f + \delta + e_H + e_L - g] \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)$$

$$\begin{aligned}
&= \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \pi_H(\Theta) - \chi_{i_\chi} X^{\phi}_{H,i_\chi}(\{\Theta_k\}_{k=1}^{K}) K^{1-\phi} \\
&+ X_{H,i_\chi}[\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)] \\
&+ e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N}-N_L), N_L, \omega_k) \\
&+ e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N}-N_H), \omega_k) \\
&+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} N_L m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H, N_L-1, \omega_k)] \\
&+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} (N_H-1) m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H-1, N_L, \omega_k)]
\end{aligned} \tag{36}$$

From the Euler equation of the household, I know that $\frac{\dot{C}_t}{C_t} = r_t - \rho$. In the BGP, all variables grow at a constant rate 'g', hence $g = r - \rho$ or $r - g = \rho$.

$$
\begin{aligned}
\implies [\rho + \delta_f + \delta + e_H + e_L] \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
= \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta)\pi_H(\Theta) - \chi_{i_\chi} X^{\phi}_{H,i_\chi}(\{\Theta_k\}_{k=1}^{K}) K^{1-\phi} \\
+ X_{H,i_\chi}[\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)] \\
+ e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N}-N_L), N_L, \omega_k) \\
+ e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N}-N_H), \omega_k) \\
+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} N_L m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H, N_L-1, \omega_k)] \\
+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} (N_H-1) m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H-1, N_L, \omega_k)]
\end{aligned}
\tag{37}
$$

Substituting X_{H,i_χ} from equation 35 into the equation above, I get

$$[\rho+\delta_f+\delta+e_H+e_L]\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\Theta)$$

$$\begin{aligned}
&=\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)\pi_H(\Theta)\\
&-\chi_{i_\chi}\left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1}\sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)}\sum_{\omega_l\in\{\omega_H,\omega_L\}}\mu_{N_{Hl},N_{Ll},\omega_l}v_{H,i_\chi}(N_{Hl}+1,N_{Ll},\omega_l)}{\phi\chi_{i_\chi}}\right]^{\frac{\phi}{\phi-1}} K\\
&+\left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1}\sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)}\sum_{\omega_l\in\{\omega_H,\omega_L\}}\mu_{N_{Hl},N_{Ll},\omega_l}v_{H,i_\chi}(N_{Hl}+1,N_{Ll},\omega_l)}{\phi\chi_{i_\chi}}\right]^{\frac{1}{\phi-1}} K\\
&\sum_{N_{Hl}=0}^{\bar{N}-1}\sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)}\sum_{\omega_l\in\{\omega_H,\omega_L\}}\mu_{N_{Hl},N_{Ll},\omega_l}v_{H,i_\chi}(N_{Hl}+1,N_{Ll},\omega_l)\\
&+e_H\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\min(N_H+1,\bar{N}-N_L),N_L,\omega_k)\\
&+e_L\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)v_{H,i_\chi}(N_H,\min(N_L+1,\bar{N}-N_H),\omega_k)\\
&+(\delta+\delta_f)\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} N_L m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H,N_L-1,\omega_k)]\\
&+(\delta+\delta_f)\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} (N_H-1) m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H-1,N_L,\omega_k)]
\end{aligned} \tag{38}$$

Using $K=\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}((N_H,N_L,\omega_k))=\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)$ in the expression above, I get

$$
\begin{aligned}
&[\rho + \delta_f + \delta + e_H + e_L] \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta) \\
&= \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \pi_H(\Theta) \\
&- \chi_{i_\chi} \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{\phi}{\phi-1}} \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \\
&+ \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{1}{\phi-1}} \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) \\
&\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H, \omega_L\}} \mu_{N_{Hl}, N_{Ll}, \omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l) \\
&+ e_H \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N}-N_L), N_L, \omega_k) \\
&+ e_L \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N}-N_H), \omega_k) \\
&+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} N_L m_{H,i_\chi}(\Theta) [-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H, N_L-1, \omega_k)] \\
&+ (\delta + \delta_f) \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H, \omega_L\}} (N_H-1) m_{H,i_\chi}(\Theta) [-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H-1, N_L, \omega_k)]
\end{aligned}
\tag{39}
$$

$$\implies [\rho + \delta_f + \delta + e_H + e_L] \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\Theta)$$

$$= \sum_{N_H=1}^{\bar{N}} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{\omega_k \in \{\omega_H,\omega_L\}}$$

$$\Bigg(m_{H,i_\chi}(\Theta)\pi_H(\Theta) - \chi_{i_\chi} \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{\phi}{\phi-1}} m_{H,i_\chi}(\Theta)$$

$$+ \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{1}{\phi-1}} m_{H,i_\chi}(\Theta) \tag{40}$$

$$\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)$$

$$+ e_H m_{H,i_\chi}(\Theta) v_{H,i_\chi}(\min(N_H+1, \bar{N}-N_L), N_L, \omega_k)$$

$$+ e_L m_{H,i_\chi}(\Theta) v_{H,i_\chi}(N_H, \min(N_L+1, \bar{N}-N_H), \omega_k)$$

$$+ (\delta + \delta_f) N_L m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H, N_L - 1, \omega_k)]$$

$$+ (\delta + \delta_f)(N_H - 1) m_{H,i_\chi}(\Theta)[-v_{H,i_\chi}(\Theta) + v_{H,i_\chi}(N_H - 1, N_L, \omega_k)] \Bigg)$$

I know from equation 35 that

$$x_{H,i_\chi} = \frac{X_{H,i_\chi}}{K} = \left[\frac{\sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l)}{\phi \chi_{i_\chi}} \right]^{\frac{1}{\phi-1}} \tag{41}$$

Define[43]

$$\hat{\mathbb{E}}(v_{H,i_\chi}) = \sum_{N_{Hl}=0}^{\bar{N}-1} \sum_{N_{Ll}=0}^{\bar{N}-(N_{Hl}+1)} \sum_{\omega_l \in \{\omega_H,\omega_L\}} \mu_{N_{Hl},N_{Ll},\omega_l} v_{H,i_\chi}(N_{Hl}+1, N_{Ll}, \omega_l) \tag{42}$$

Using equations 41 and 42, equation 40 can be written as

[43] $\hat{\mathbb{E}}(v_{H,i_\chi})$ is the expectation with special $\mu_{N_{Hl},N_{Ll},\omega_l}$ terms, i.e., those such that $N_{Hl} + N_{Ll} < \bar{N}$

$$[\rho+\delta_f+\delta+e_H+e_L]\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}} m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\Theta)$$

$$=\sum_{N_H=1}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{\omega_k\in\{\omega_H,\omega_L\}}$$

$$\Big(m_{H,i_\chi}(\Theta)\pi_H(\Theta)-\chi_{i_\chi}x^{\phi}_{H,i_\chi}m_{H,i_\chi}(\Theta)+x_{H,i_\chi}m_{H,i_\chi}(\Theta)\hat{\mathbb{E}}(v_{H,i_\chi})$$

$$+e_H m_{H,i_\chi}(\Theta)v_{H,i_\chi}(\min(N_H+1,\bar{N}-N_L),N_L,\omega_k)+e_L m_{H,i_\chi}(\Theta)v_{H,i_\chi}(N_H,\min(N_L+1,\bar{N}-N_H),\omega_k)$$

$$+(\delta+\delta_f)m_{H,i_\chi}(\Theta)\left(N_L[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H,N_L-1,\omega_k)]+(N_H-1)[-v_{H,i_\chi}(\Theta)+v_{H,i_\chi}(N_H-1,N_L,\omega_k)]\right)\Big) \tag{43}$$

Go back to section 3.2.

8.4 Market-state Distribution and Expansion Rates

$\zeta_{N_{HH},N_{HL},N_{LH},N_{LL},\omega_k}$ denotes the share of geographic markets $j\in[0,1]$ having N_{HH} high-type productivity and high-type expansion efficiency large firms, N_{HL} high-type productivity and low-type expansion efficiency large firms, N_{LH} low-type productivity and high-type expansion efficiency large firms, N_{LL} low-type productivity and low-type expansion efficiency large firms and $\omega_k\in\{\omega_H,\omega_L\}$ market demand shifter. μ_{N_H,N_L,ω_k} denotes the share of geographic markets $j\in[0,1]$ having N_H high-type productivity large firms, N_L low-type productivity large firms and $\omega_k\in\{\omega_H,\omega_L\}$ market demand shifter.

Then the share of high-type markets in the economy is:

$$\sum_{N_H=0}^{\bar{N}}\sum_{N_{HH}=0}^{N_H}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{N_{LH}=0}^{N_L}\zeta_{N_{HH},N_H-N_{HH},N_{LH},N_L-N_{LH},\omega_H}=\sum_{N_H=0}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\mu_{N_H,N_L,\omega_H}=s_{\omega_H}$$

and the share of low-type markets in the economy is:

$$\sum_{N_H=0}^{\bar{N}}\sum_{N_{HH}=0}^{N_H}\sum_{N_L=0}^{\bar{N}-N_H}\sum_{N_{LH}=0}^{N_L}\zeta_{N_{HH},N_H-N_{HH},N_{LH},N_L-N_{LH},\omega_L}=\sum_{N_H=0}^{\bar{N}}\sum_{N_L=0}^{\bar{N}-N_H}\mu_{N_H,N_L,\omega_L}=s_{\omega_L}=1-s_{\omega_H}$$

The total geographic expansion of the high-type productivity large firms is:

$$e_H=e_{HH}+e_{HL} \tag{44}$$

where

$$e_{HH} = \sum_{N_H=0}^{\bar{N}} \sum_{N_{HH}=0}^{N_H} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{N_{LH}=0}^{N_L} \sum_{\omega_k \in \{\omega_H, \omega_L\}} \zeta_{N_{HH}, N_H - N_{HH}, N_{LH}, N_L - N_{LH}, \omega_k} N_{H,H} x_{H,H} + \Lambda_{HH} z \quad (45)$$

and

$$e_{HL} = \sum_{N_H=0}^{\bar{N}} \sum_{N_{HH}=0}^{N_H} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{N_{LH}=0}^{N_L} \sum_{\omega_k \in \{\omega_H, \omega_L\}} \zeta_{N_{HH}, N_H - N_{HH}, N_{LH}, N_L - N_{LH}, \omega_k} N_{H,L} x_{H,L} + \Lambda_{HL} z \quad (46)$$

Similarly, the total geographic expansion of the low-type productivity large firms is:

$$e_L = e_{LH} + e_{LL} \quad (47)$$

where

$$e_{LH} = \sum_{N_H=0}^{\bar{N}} \sum_{N_{HH}=0}^{N_H} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{N_{LH}=0}^{N_L} \sum_{\omega_k \in \{\omega_H, \omega_L\}} \zeta_{N_{HH}, N_H - N_{HH}, N_{LH}, N_L - N_{LH}, \omega_k} N_{L,H} x_{L,H} + \Lambda_{LH} z \quad (48)$$

and

$$e_{LL} = \sum_{N_H=0}^{\bar{N}} \sum_{N_{HH}=0}^{N_H} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{N_{LH}=0}^{N_L} \sum_{\omega_k \in \{\omega_H, \omega_L\}} \zeta_{N_{HH}, N_H - N_{HH}, N_{LH}, N_L - N_{LH}, \omega_k} N_{L,L} x_{L,L} + \Lambda_{LL} z \quad (49)$$

Total geographic expansion in the economy: $e = e_H + e_L = e_{HH} + e_{HL} + e_{LH} + e_{LL}$.

Go back to section 3.2.

8.5 Firm-size Distributions

$\gamma_{i_A, i_\chi}(j)$ denotes the mass of type $\{i_A, i_\chi\}$ large firms present in j geographic markets.

Then $\sum_{i_A=H,L} \sum_{i_\chi=H,L} \sum_{j=1}^{\infty} \gamma_{i_A, i_\chi}(j) = \sum_{j=1}^{\infty} \gamma_{HH}(j) + \sum_{j=1}^{\infty} \gamma_{HL}(j) + \sum_{j=1}^{\infty} \gamma_{LH}(j) + \sum_{j=1}^{\infty} \gamma_{LL}(j)$ is the mass of firms in the economy.

$\sum_{i_A=H,L} \sum_{i_\chi=H,L} \sum_{j=1}^{\infty} j\gamma_{i_A, i_\chi}(j) = \sum_{j=1}^{\infty} j\gamma_{HH}(j) + \sum_{j=1}^{\infty} j\gamma_{HL}(j) + \sum_{j=1}^{\infty} j\gamma_{LH}(j) + \sum_{j=1}^{\infty} j\gamma_{LL}(j)$ is the mass of establishments/plants in the economy.

For exposition, assume that a high-type productivity and high-type expansion efficiency large firm can be in at most 5 markets, then the inflow-outflow equations for $\sum_{j=1}^{\infty} \gamma_{HH}(j)$ can be written as:

	Inflow	Outflow
$j = 1$	$z.\Lambda_{HH} + \gamma_{HH}(2).\delta.2$	$\gamma_{HH}(1).x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(1)\delta + \gamma_{HH}(1)\delta_f$
$j = 2$	$\gamma_{HH}(1).x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(3).\delta.3$	$\gamma_{HH}(2).2x.(1-\hat{\mu}) + \gamma_{HH}(2).2.\delta + \gamma_{HH}(2).\delta_f$
$j = 3$	$\gamma_{HH}(2).2x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(4).\delta.4$	$\gamma_{HH}(3).3x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(3).3.\delta + \gamma_{HH}(3).\delta_f$
$j = 4$	$\gamma_{HH}(3).3x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(5).5.\delta$	$\gamma_{HH}(4).4x.(1-\hat{\mu}) + \gamma_{HH}(4).4.\delta + \gamma(4).\delta_f$
$j = 5$	$\gamma_{HH}(4).4x_{HH}.(1-\hat{\mu})$	$\gamma_{HH}(5).5.\delta + \gamma_{HH}(5).\delta_f$

where $\hat{\mu} = \sum_{N_H=0}^{\bar{N}} (\mu^H_{N_H,\bar{N}-N_H} + \mu^L_{N_H,\bar{N}-N_H})$

Since inflow=outflow in the steady state, we can write this as a system of equations to solve for $\gamma_{HH}(j)$. In general, for $j = 1$,

$$\dot{\gamma}_{HH}(1) = z.\Lambda_{HH} + \gamma_{HH}(2).\delta.2 - \gamma_{HH}(1)[x_{HH}.(1-\hat{\mu}) + \delta + \delta_f] = 0 \tag{50}$$

$$\implies \gamma_{HH}(2) = \frac{\Big(\gamma_{HH}(1)[x_{HH}.(1-\hat{\mu}) + \delta + \delta_f] - (z.\Lambda_{HH})\Big)}{\delta.2}$$

for $j \geq 2$,

$$\dot{\gamma}_{HH}(j) = \gamma_{HH}(j-1).(j-1)x_{HH}.(1-\hat{\mu}) + \gamma_{HH}(j+1).(j+1).\delta - \gamma_{HH}(j)[j.x_{HH}.(1-\hat{\mu}) + j.\delta + \delta_f] = 0 \tag{51}$$

$$\implies \gamma_{HH}(j+1) = \frac{\Big(\gamma_{HH}(j)[j.x_{HH}.(1-\hat{\mu}) + j.\delta + \delta_f]\Big) - \Big(\gamma_{HH}(j-1).(j-1)x_{HH}.(1-\hat{\mu})\Big)}{(j+1).\delta}$$

and for $j = \bar{j}$

$$\dot{\gamma}_{HH}(j) = \gamma_{HH}(j-1).(j-1)x_{HH}.(1-\hat{\mu}) - \gamma_{HH}(j)[j.\delta + \delta_f] = 0$$

$$\implies \gamma_{HH}(j) = \frac{\gamma_{HH}(j-1).(j-1)x_{HH}.(1-\hat{\mu})}{[j.\delta + \delta_f]}$$

In addition,

$$\sum_{j=1}^{\infty} j.\gamma_{HH}(j) = \sum_{N_H=0}^{\bar{N}} \sum_{N_{HH}=0}^{N_H} \sum_{N_L=0}^{\bar{N}-N_H} \sum_{N_{LH}=0}^{N_L} (\zeta_{N_{HH},N_H-N_{HH},N_{LH},N_L-N_{LH},\omega_H} + \zeta_{N_{HH},N_H-N_{HH},N_{LH},N_L-N_{LH},\omega_L})N_{HH} \tag{52}$$

Similarly we can solve for $\gamma_{HL}(j), \gamma_{LH}(j), \gamma_{LL}(j)$ by replacing x_{HH} with x_{HL}, x_{LH}, x_{LL} respectively, $z.\Lambda_{HH}$ by $z.\Lambda_{HL}, z.\Lambda_{LH}, z.\Lambda_{LL}$ and N_{HH} by N_{HL}, N_{LH}, N_{LL} respectively.

Go back to section 3.2.

8.6 Estimation Algorithm and Identification

I pick the maximum number of large firms in a local market to be 4, which delivers 70 unique market states. Fifteen parameters need to be determined in the model: $\rho, \omega_H/\omega_L, A_H/A_L, s_{\omega_H}, \delta, \psi, \delta_f, A_H/A_c, \chi_H, \chi_L, \phi, \lambda_{HH}, \lambda_{HL}, \lambda_{LH}, \sigma$. The consumer discount rate (ρ) is set to 0.04, which implies a real interest rate of 4% when the growth rate is 0% (as in my model). The share of high-type markets (s_{ω_H}) is set to be equal to the share set in the model, and the high-type market demand shifter relative to the low-type market demand shifter (ω_H/ω_L) is derived using population ratios of local markets (CSDs) once s_{ω_H} is set. The relative productivity of high-type productivity firms with respect to the low-type productivity firms (A_H/A_L) is obtained by assuming that the top 40% of the firms are high-type productivity firms and calculating the ratio of average labor productivities of the two types of firms. The rest of the parameters are structurally estimated following a simulated method of moments approach. In this section, I discuss the data moments I use to discipline the parameter values, and highlight which moments help identify which parameters. The associated Jacobian matrix is presented in Table 8.

8.6.1 Data Moments

Since the CEEDD consists of the universe of Canadian firms, I select the active, for-profit firms in non-tradable industries[44] to calculate the aggregate moments in the data. As explained in

[44] I include firms in Construction; Wholesale Trade; Retail Trade; Transportation and Warehousing (except some 4 digit NAICS codes where the description suggests that firms' operations are tied to certain geographies); Information and cultural Industries; Finance and Insurance (except central banking); Real Estate and Rental and Leasing; Professional, Scientific, and Technical Services; Management of Companies and Enterprises; Administrative and Support and Waste Management and Remediation Services; Health Care and Social Assistance;

appendix 8.1, I use the matched employer-employee feature of CEEDD to derive the number of geographic markets of firms in Canada.

I classify firms with at most one local market (CSD) throughout their period of operation as 'fringe firms'. The remaining firms are classified as 'large firms', which are further divided into high-type and low-type based on productivity and expansion efficiency. Productivity is measured using firm-level labor productivity (calculated as value added per total number of employees), while expansion efficiency is assessed based on the number of geographic markets, as follows:

- Firm types based on productivity: I regress firm level labor productivities on industry (4 digit NAICS level) and year fixed effects. I sort the residuals of this regression in decreasing order by year. I categorize the top 40% of firms as high-type productivity firms and the bottom 60% as low-type productivity firms.
- Firm types based on expansion efficiency: I regress firm level number of geographic markets on firm age, and industry (4 digit NAICS level) and year fixed effects. I sort the residuals of this regression in decreasing order by year. I classify the top 40% of firms as high-type expansion efficiency firms and the bottom 60% as low-type expansion efficiency firms.

I will now discuss the data moments I calculate for the 'large firms', which are used to inform the parameter values in the model estimations. For each moment, I calculate annual values using the methods outlined below, and then average them across the periods 2002-2005 and 2013-2016 to obtain the data moments for the early and late period estimations, respectively:

- **Sales-weighted average local HHI**: The Herfindahl-Hirschman Index (HHI) is calculated by squaring the market share of every firm competing in a market and then summing the resulting numbers. I calculate the HHI within each 4 digit NAICS industry within each local market (CSD), and take the average for each 4 digit NAICS industry using

Arts, Entertainment, and Recreation; Accommodation and Food Services; Other Services (except Public Administration). My sample excludes firms operating in Agriculture, Mining, Utilities, Manufacturing, Education and Public Administration as economic activity in these sectors is tied to a specific geography.

the CSD sales as weights. I aggregate these 4 digit NAICS industry level HHIs to 2 digit NAICS industry level average HHIs using the 4 digit NAICS industry sales as weights. Then I obtain an economy-wide average using the respective 2 digit NAICS industry level sales as weights.

- **Firm exit rate**: I first get the exit year of each large firm based on the last year the firm is observed in the data (I do not consider 2018 as that is the last year in the data). Then I calculate the exit rate as the total count of exiting firms over the total number of firms in that year.

- **Average number of markets of H-type firms**: I calculate the average number of markets for the high-type expansion efficiency firms and get their average at the 2 digit NAICS level. Then I get the economy-wide sales-weighted and employment-weighted averages using sales and employment of the 2 digit NAICS industry as weights respectively. I use the employment-weighted numbers in the estimations since the sales-weighted numbers show a very steep increase in the late period due to the rapid increase in sales of a particular 2 digit NAICS industry.

- **Average number of markets of L-type firms**: I calculate the average number of markets for the low-type expansion efficiency firms and get their average at the 2 digit NAICS level. Then I get the economy-wide sales-weighted and employment-weighted averages using sales and employment of the 2 digit NAICS industry as weights respectively. I use the sales-weighted numbers in the estimations (the employment-weighted numbers are very similar).

- **Expansion expenditure ratio**: Since the CEEDD does not have direct information on firms' expenditure on geographic expansion, I construct my own measures using information on expenditures on buildings and machinery and equipment (M&E). I calculate the flow variable versions of expenditures on buildings and M&E each year using information on the stocks of expenditures on buildings and M&E in the current year and the consecutive year. Then I calculate the firms' geographic expansion expenditure by summing their flow expenditures on buildings and M&E. I divide the total geographic expansion expenditure of the firms who are in my sample for constructing this variable by the share of Canadian GDP attributable to these firms to derive the expansion expenditure ratio.

- **Proportion of HH-type firms**: The methodology described above for obtaining the high-type and low-type productivity and high-type and low-type expansion efficiency firms assigns one of these types to each firm: HH-type, HL-type, LH-type, LL-type, where the first H/L denotes their productivity type and the second H/L denotes their expansion efficiency type. I use this firm-type assignment in the data to derive counts of firms assigned the HH-type and divide them by the total number of large firms to get the proportion of HH-type firms.

- **Proportion of HL-type firms**: I use the firm-type assignment in the data to derive counts of firms assigned the HL-type and divide them by the total number of large firms to get the proportion of HL-type firms.

- **Proportion of LL-type firms**: I use the firm-type assignment in the data to derive counts of firms assigned the LL-type and divide them by the total number of large firms to get the proportion of LL-type firms.

- **Average Markup**: I estimate firm-level markups by 2 digit NAICS industry in the data based on the markup estimation micro-approach proposed by Loecker and Warzynski (2012) using all the firms (including the fringe firms). Then I get 2 digit NAICS industry level averages of markups, which I aggregate to an economy-wide average using the 2 digit NAICS industry sales as weights. For robustness, I also derive total cost weighted and total cost of sales weighted average markups using the 2 digit NAICS industry total cost and total cost of sales as weights respectively, but use the sales-weighted average markup in my estimations.

- **Standard deviation of markup**: Once I obtain the firm-level markups using the method described above, I also get 2 digit NAICS industry level standard deviations of markups. Then I aggregate them to an economy-wide standard deviation of markups using the 2 digit NAICS industry sales as weights. For robustness, I also derive total cost weighted and total cost of sales weighted standard deviation of markups using the 2 digit NAICS industry total cost and total cost of sales as weights respectively, but use the sales-weighted standard deviation of markup in my estimations.

- **Average market share of top 4 firms**: I calculate the market share of the top 4 firms within each 2 digit NAICS industry within each local market (CSD), and take the average for each 2 digit NAICS industry. Then I obtain an economy-wide average market share of top 4 firms using the 2 digit NAICS industry level averages.
- **Standard deviation of the number of markets**: I calculate the average number of markets for all types of firms and get their average and standard deviation at the 2 digit NAICS level. Then I get the economy-wide sales-weighted and employment-weighted averages and standard deviations using sales and employment of the 2 digit NAICS industry as weights respectively. I use the sales-weighted standard deviation of the number of markets in the estimations.
- **Labor Share**: I obtain the Canadian labor share estimates from the 'Share of Labour Compensation in GDP at Current National Prices for Canada' series on FRED.[45]

8.6.2 Identification and the Estimation Algorithm

Table 8: Identification: Jacobian Matrix

	δ	ψ	δ_f	A_H/A_c	χ_H	χ_L	ϕ	λ_{HH}	λ_{HL}	λ_{LL}	σ
sales-weighted avg. local hhi	-0.079	0.044	-0.022	1.045	-0.069	0.005	0.395	-0.006	-0.023	-0.047	0.109
firm exit rate	0.759	-0.040	0.267	0.013	-0.071	0.123	-0.372	-0.045	-0.026	0.133	0.000
avg. no. of markets of H firms	-0.703	1.362	-0.110	-0.390	-1.410	0.520	5.613	-0.907	-1.130	-0.274	-0.009
avg. no. of markets of L firms	0.048	0.040	0.039	-0.032	0.408	-0.364	0.317	-0.019	0.108	-0.117	-0.001
expansion expenditure ratio	0.402	0.233	0.111	0.925	-0.255	0.046	0.552	-0.077	-0.247	-0.114	0.021
proportion of HH firms	0.162	0.333	-0.035	-0.061	-0.607	0.293	2.190	0.690	-0.431	-0.779	-0.001
proportion of HL firms	-0.284	-0.038	-0.126	0.053	0.473	-0.564	0.587	-0.104	0.896	-0.559	0.001
proportion of LL firms	0.075	-0.065	0.058	-0.007	-0.044	0.159	-0.776	-0.131	-0.274	0.487	-0.000
avg. markup	-0.012	-0.002	-0.003	0.517	-0.005	-0.000	0.020	0.001	0.001	0.001	0.007
std. dev. markup	0.354	-0.333	0.097	1.237	0.395	-0.039	-2.421	0.048	0.175	0.338	-0.016
avg. market share, top 4 firms	-0.191	0.171	-0.052	0.447	-0.208	0.020	1.254	-0.025	-0.090	-0.175	0.049
std. dev. number of markets	-3.557	0.589	-2.277	1.043	-1.234	0.650	4.049	-0.672	-0.379	1.182	0.022
labor share	0.044	-0.028	0.012	-0.462	0.041	-0.003	-0.239	0.004	0.015	0.030	-0.010

Notes: The table shows the Jacobian matrix associated with the estimation of the late period. Each entry of the matrix is the percentage change in each moment given one percent increase in each parameter.

[45]The data can be accessed at https://fred.stlouisfed.org/series/LABSHPCAA156NRUG.

The model is highly non-linear, but some parameters are more important for certain moments even though all parameters affect all the moments. The success of the SMM estimation depends on model identification. Hence, it is important to ensure that the chosen moments are sensitive to variations in the structural parameters.

Table 8 shows the Jacobian matrix associated with the estimation of the model for the late period. Each entry of the matrix shows the percentage change in each moment given one percent increase in each parameter, hence giving an insight into which data moments are most informative in helping to identify each parameter: (i) The exogenous market exit rate is mainly identified by the firm exit rate and expansion expenditure ratio. If the market exit rate is higher, the firms expect to be hit by this shock more often and spend more resources to expand into more geographic markets. Higher market exit is also positively correlated with higher firm exit rate; (ii) The entrepreneurs' innovation cost scale ψ is identified using firm exit, average number of markets for both types of firms and the standard deviation of the number of markets. if ψ increases, firm entry falls (hence exit falls since firm entry rate equals firm exit rate in the stationary equilibrium) and incumbent firms expand into more markets since they can earn higher profits due to lesser entry. The standard deviation of the number of markets increases as the dispersion between firms' number of markets increases due to the difference in expansion rates of the firms; (iii) The exogenous firm shut-down rate is highly identified using firm exit rate and the standard deviation of markets. If δ_f increases, firm exit rate increases as a higher number of firms are hit by the firm shut-down shock. In addition, this implies a fall in the standard deviation of the number of markets since this shock decreases the variance in firms' number of markets; (iv) The average markup and the labor share are helpful in identifying the relative productivity of high-type productivity firms with respect to the competitive fringe A_H/A_c since a higher A_H/A_c implies reduced competition from fringe firms and a within-market share reallocation to large firms, which generates a higher average markup and lower labor share; (v) The high-type expansion efficiency large firms' expansion cost scale χ_H is mainly identified by the average number of markets of the high-type expansion efficiency firms and the standard deviation of the number of markets. A higher χ_H implies a higher cost scale of expansion for the high-type expansion efficiency firms. This reduces the number of markets of the high-type expansion efficiency firms and the dispersion between the firms' number of markets as the difference between the high-type and low-type expansion efficiency firms' number of markets falls due to this change; (vi) The low-type expansion efficiency large firms' expansion cost scale χ_L is closely identified by the average number of markets of the low-type expansion efficiency firms and the standard deviation of the number of markets. A higher χ_L implies a higher cost scale of expansion for the low-type expansion efficiency firms. This reduces the number of markets of the low-type expansion efficiency firms, but increases the dispersion between the firms' number of markets as the difference between the high-type and low-type expansion efficiency firms' number of markets widens due to this change; (vii) The average number of markets for high-type and low-type expansion efficiency firms and the standard deviation of the number of markets are helpful in identifying the large firms' expansion cost convexity ϕ. As ϕ increases, expansion costs fall, hence increasing the number of markets for both types of firms and the standard deviation of the number of markets; (viii) The probability of being HH-type upon successful entry λ_{HH} is very closely identified by the proportion of HH firms and the number of number of markets of the high-type expansion efficiency firms. When λ_{HH} increases, the proportion of HH firms increases, but the number of markets of the high-type expansion efficiency firms falls since these firms expect to encounter more high-type productivity firms in the new markets which discourages their geographic expansion; (ix) The probability of being HL-type upon successful entry λ_{HL} is very closely identified by the proportion of HL firms and the number of number of markets of the high-type expansion efficiency firms. When λ_{HL} increases, the proportion of HL firms increases, but the number of markets of the high-type expansion efficiency firms falls since these firms expect to encounter more high-type productivity firms in the new markets which discourages their geographic expansion; (x) The probability of being LL-type upon successful entry λ_{LL} is very closely identified by the proportion of LL firms, proportion of HH firms and the proportion of HL firms. When λ_{LL} increases, the proportion of LL firms increases, but the proportion of HH firms and the proportion of HL firms fall; (xi) The average market share of top 4 firms and the standard deviation of markups are most helpful in identifying the elasticity of substitution between local varieties σ. Larger σ implies higher substitution among local varieties, which leads to lower market power, hence a lower standard deviation of markups. At the same time, it helps the most productive firms gain more market share, hence increasing the average market share of top 4 firms.

SMM proceeds in the following way: I solve the dynamic problem, get the policy functions and use the policy functions to calculate the model moments for an arbitrary value of parameter vector $\theta = \{\delta, \psi, \delta_f, A_H/A_c, \chi_H, \chi_L, \phi, \lambda_{HH}, \lambda_{HL}, \lambda_{LH}, \sigma\}$. The simulated moments estimator is defined as the solution to the minimization of:

$$\hat{\theta} = arg\ min_{\theta} \sum_{k=1}^{11} \left[\frac{moment_k^{model}(\theta) - moment_k^{data}(\theta)}{moment_k^{data}(\theta)} \right]^2 \hat{\Omega}_k \tag{53}$$

where moment$_k^{model}$ is the value of moment k in the model and moment$_k^{data}$ is the value of the moment in the data with weight $\hat{\Omega}_k$. I use the differential evolution algorithm to minimize the objective function and restart the program with different initial conditions to ensure that the estimator converges to the global minimum.

Go back to section 4.