

# Three Essays in Macroeconomics with a Focus on Market Power

by

Hassan Faryaar

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial  
fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics

Carleton University

Ottawa, Ontario

©2020

Hassan Faryaar

# Abstract

This thesis includes three essays on macroeconomics. The first essay (Chapter 2) focuses on the impacts of the rising market power. The second and third essays (Chapters 3 and 4) concentrate on estimating the measure of market power, i.e. markups. In Chapter 2, I provide a unified explanation for some key secular stylized facts in the US economy using a heterogeneous-household model along with the rise of market power. In Chapter 3, using the novel database of T2-Longitudinal Employment Analysis Program (T2-LEAP), I estimate firm-level markups in Canada. In Chapter 4, I introduce a new approach to estimate firm-level markups in the US. The results of Chapters 3 and 4 indicate that markups in the US and Canada increased mildly. This finding is consistent with my model in Chapter 2 that a moderate increase in markups can explain well the new stylized facts of the US economy.

The market power, as measured by markups, has risen moderately over the last 35 years in the US corporate sector. Over the same time, the labour and capital shares of income and the real rate of return on capital have declined, while the wealth-output ratio has increased. In the second chapter, I argue that in the presence of income inequality, the rise of market power can be one of the main drivers of these trends. To

jointly study the trends, I combine the incomplete market model of [Aiyagari \(1994\)](#) with the imperfect competition model of [Dixit and Stiglitz \(1977\)](#). The results suggest that the increasing profit of the corporate sector and its unequal distribution among households, who have a heterogeneous marginal propensity to consume, provide a unified explanation for the above stylized facts of the US economy.

In the third chapter, I examine the evolution of markups in the Canadian economy. The literature on firm-level markup estimation suffers from two problems: The lack of a reliable measure of a variable input and the sample selection bias due to using only publicly-traded firms. To address these issues, I use the T2-LEAP database, which has the universe of Canadian firms (both publicly-traded and non-publicly-traded) and the record of wage bill for each firm. The wage bill is more reliable than the commonly used measures of variable inputs, i.e. the cost of goods sold and operating expenses. The result indicates that the average gross (value-added) output markups increased mildly from 6% (24%) in 2002 to 9% (29%) in 2015. I further find that the operating expenses is a better measure than the cost of goods sold when the wage bill is not available, and excluding non-publicly-traded firms leads to an increase in the estimated average markup.

In the fourth chapter, I introduce a new approach to estimate firm-level markups from accounting data. In the existing literature, the estimates of firm-level markups vary significantly depending on the choice of measure of variable input obtained from accounting data. The new approach, which uses economic profit rates and returns to scale of firms, is less sensitive to the choice of the measure of variable inputs. The findings suggest that, in the US, the average markups increased by six percentage

points, from 7% in 1980 to around 13% in 2018. Over the same period, the average economic profit rate increased by five percentage points, from 4% to 9%, and production technology exhibits constant returns to scale.

# Acknowledgements

First of all, I would like to express my sincere gratitude to my thesis supervisors, Minjoon Lee and Patrick Coe. This dissertation would not have been accomplished without their thoughtful input and feedback, ongoing support and sincere encouragement.

I would like to thank my committee members, Hashmat Khan (Carleton University) and Lilia Karnizova (University of Ottawa), for their feedback and guidance throughout this dissertation. I would also like to thank the external examiner, Danny Leung (Statistics Canada) and the internal examiner, Shi Li (Carleton University) for their helpful insights and recommendations.

I am very thankful to the Canadian Centre for Data Development and Economic Research (CDER) at Statistics Canada for providing me with secure access to the novel database of T2-LEAP. Without granting this funding, I would not have been able to complete my second paper.

I would like to express my gratitude to my parents, who were the first to pave the path for me and whose love and guidance follow me in every project I pursue. Finally, I would like to extend my deep gratitude to my wife, the love of my life, Azam

Kazemiyan. She has always been there throughout this long journey to support and motivate me.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 The Rise of Market Power, Inequality and the New Stylized Facts in the US Economy</b>	<b>7</b>
2.1 Introduction . . . . .	7
2.2 The new stylized facts . . . . .	13
2.3 Model . . . . .	16
2.3.1 Model description . . . . .	16
2.3.2 Households . . . . .	18
2.3.3 The progressive tax system . . . . .	20
2.3.4 Firms . . . . .	22
2.3.5 Equilibrium . . . . .	23
2.3.6 The accounting of assets . . . . .	25
2.3.7 The labour share . . . . .	26

2.3.8	Computation strategy . . . . .	27
2.4	Results . . . . .	31
2.4.1	Calibration . . . . .	31
2.4.2	The initial distribution of wealth and factor shares . . . . .	32
2.4.3	An increase in the markup . . . . .	33
2.4.4	Results . . . . .	34
2.5	Conclusion . . . . .	39
<b>3</b>	<b>The Estimates of Firm-Level Markups in Canada Using a Novel Data to Overcome Measurement Issues</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Methodology . . . . .	48
3.3	The T2-LEAP database . . . . .	52
3.4	The challenges in measuring a variable input . . . . .	56
3.5	Results . . . . .	60
3.5.1	The evolution of firm-level markups . . . . .	60
3.5.2	Changing the measure of variable inputs . . . . .	64
3.5.3	Excluding non-publicly-traded firms . . . . .	67
3.5.4	The dispersion of markups and the industry-level markups . . . . .	70
3.5.5	Micro-data vs. macro-data markups . . . . .	72
3.6	Conclusion . . . . .	75
<b>4</b>	<b>Firm-Level Markups and Accounting Data: A New Approach</b>	<b>78</b>
4.1	Introduction . . . . .	78

4.2	Methodology . . . . .	82
4.3	Economic profits and accounting data . . . . .	88
4.3.1	Economic profits . . . . .	88
4.3.2	Data . . . . .	91
4.4	Results . . . . .	92
4.4.1	The evolution of markups . . . . .	92
4.4.2	The dispersion of markups . . . . .	95
4.4.3	The industry-level markups . . . . .	96
4.5	Total-input v.s. single-input approach . . . . .	98
4.6	Conclusion . . . . .	101
<b>5</b>	<b>Conclusion</b>	<b>103</b>
	<b>References</b>	<b>107</b>
<b>A</b>	<b>Appendix for Chapter 2</b>	<b>121</b>
A.1	Firms . . . . .	121
A.1.1	Consumption good producers . . . . .	121
A.1.2	Intermediate good producers . . . . .	122
A.2	The rise of market power and its measurement . . . . .	124
<b>B</b>	<b>Appendix for Chapter 3</b>	<b>127</b>
B.1	Wage bill and the US markups . . . . .	127
B.1.1	Generating intermediate inputs . . . . .	129
B.2	Returns to scale . . . . .	134

B.3	The value-added markups at the industry level . . . . .	137
B.4	The gross markups at the industry level . . . . .	138
<b>C</b>	<b>Appendix for Chapter 4</b>	<b>139</b>
C.1	Estimating markups using wage bill . . . . .	139

# List of Tables

2.1	Calibration . . . . .	32
2.2	The initial distribution of wealth and the stylized facts in the baseline model and data . . . . .	33
2.3	The stylized facts in data and in the model . . . . .	39
3.1	A sample income statement . . . . .	57
A.1	The estimated increase of markups for the US economy . . . . .	126
A.2	The estimated value of markup for the US economy . . . . .	126
B.1	The average returns to scale (ARTS) . . . . .	135
B.2	The value-added returns to scale (RS) and the elasticity of variable inputs ( $\beta_l$ ) at the industry level for the benchmark model . . . . .	136

# List of Figures

2.1	The labour share . . . . .	15
2.2	The capital share . . . . .	15
2.3	The wealth-output ratio . . . . .	16
2.4	The rate of return on capital . . . . .	16
2.5	The pre-tax profit share of GVA . . . . .	16
2.6	The overall picture of the model. . . . .	17
2.7	The algorithm of the model . . . . .	30
2.8	The flow of logic behind the model . . . . .	35
3.1	The average markups using both publicly-traded and non-publicly-traded firms in Canada. . . . .	47
3.2	The average gross markups, considering both publicly-traded and non-publicly traded firms (source: T2-LEAP) . . . . .	66
3.3	The average gross markups for publicly-traded firms (source: Compustat) . . . . .	69
3.4	The evolution of the distribution (percentiles) of markups . . . . .	71
3.5	The average of the value-added markups for the major sectors. . . . .	72
3.6	The aggregate labour share of GDP for Canada . . . . .	74

3.7	The evolution of aggregate markups for Canada . . . . .	75
4.1	The average markups for the US . . . . .	82
4.2	Markups using different measures of production function inputs . . . . .	95
4.3	The evolution of the distribution (percentiles) of markups for the US . . . . .	96
4.4	The average markups for the major sectors in the US. . . . .	98
4.5	The average markups using the two different approaches . . . . .	101
B.1	The US average markups using the wage bill. The markups are weighted based on the value-added of firms. . . . .	129
B.2	The comparison of the US markups using the wage bill, the operating expenses, and the COGS. . . . .	129
B.3	The US markups using the COGS and operating expenses for the whole observation of Compustat . . . . .	130
B.4	The average ratio of intermediate inputs . . . . .	132
B.5	The average markups estimated using the two different costs of intermediate inputs . . . . .	133
B.6	The average of the value-added markups at the industry level (source: T2-LEAP, 2002-2015). . . . .	137
B.7	The average of the gross markups at the industry level (source: T2-LEAP, 2002-2015). . . . .	138

# Chapter 1

## Introduction

The rise of market power, as measured by markups, has received increasing attention in recent years. In this dissertation, I contribute to the literature in the following ways. First, I examine the macroeconomic implications of the increase in the market power in the presence of realistic household heterogeneity in income and wealth (Chapter 2). Second, I estimate the changes in market power in Canada and in the US using a novel data (Chapter 3) and a new methodology (Chapter 4) that are matched with my findings in the second chapter.

The market power has risen in the US corporate sector since the early 1980s. Over the same time, the labour and capital shares of income and the real rate of return on capital have declined, while the capital-output ratio has increased. The literature on these trends provides separate partial explanations for each of these trends rather than a common driver among all. In the first essay (Chapter 2), I argue that a moderate rise of market power is enough to provide a unified explanation for these new stylized facts of the US economy when income inequality is considerably high,

and households are heterogeneous.”

In Chapter 2, to study inequality in an imperfectly-competitive market, I present a new model by combining the incomplete market model of [Aiyagari \(1994\)](#) with the imperfect competition model of [Dixit and Stiglitz \(1977\)](#). My model, however, is different from that in [Aiyagari \(1994\)](#) in the following ways. First, households in [Aiyagari \(1994\)](#) only save in the form of bonds, while in this model, households have an extensive margin portfolio choice, i.e. they participate either in the stock or bond market, though not both at the same time, in each period. I segment stockholders from bondholders by introducing a stock market participation cost. Second, the market in [Aiyagari \(1994\)](#) is perfectly competitive, while firms in this model are monopolistically-competitive, so they set their prices above the marginal cost and generate profits.

To provide a unified explanation for the aforementioned stylized facts, I examine the economic impacts of an increase in the market power and the profit of firms by raising the value of markup in the Dixit-Stiglitz firms. When profit share increases, ceteris paribus, the labour and capital shares decrease by the same percent (the *direct effects*). However, the wealthy, who benefit more from the increasing profit, have lower marginal propensity to consume (MPC) and, consequently, save more in the form of capital. The increase in the aggregate capital pushes down the rental rate of capital through the marginal product of capital (MPK). Therefore, an increase in the profit share of income decreases the rental rate of capital while it increases the wealth-output ratio (the *indirect effects*). Labour and capital are complementary because the elasticity of substitution between labour and capital is less than unity. When

the elasticity is less than unity, the increase in the capital-to-labour ratio results in a proportionally larger decrease in the relative price of capital-to-labour. Therefore, the *indirect effects* reduce the capital share more than the labour share. In other words, the rise of market power provides a unified explanation for the key stylized facts of the economy.

The next two essays focus on estimating changes in markups using firm-level data. Recently, a wave of “production function” approach has been used to estimate markups using accounting data. In this approach, the markup is defined as the product of two terms, the output elasticity of variable inputs, and the ratio between revenue and the cost of variable inputs. Since markup estimation depends on the measure of a single variable input, this approach is also called a “single-input” approach (Basu (2019)). This approach allows for precise and rigorous estimates of markups (Berry, Gaynor and Scott Morton (2019)), and its use of firm-level data rather than aggregate data enables researchers to capture the heterogeneity of firms (De Loecker and Eeckhout (2018b)). Despite these advantages, the current literature of the firm-level markup estimation suffers from two problems due to data availability constraints. First, this approach does not use a reliable measure of variable input for production functions because the variable inputs are not directly available in accounting data. Second, the literature focuses only on publicly-traded firms. In Chapter 3, I address the above issues by estimating firm-level markups using the novel database of Canadian firms, i.e. the T2-Longitudinal Employment Analysis Program (T2-LEAP) database.

The T2-LEAP database has two notable advantages over other databases that are

widely used in the literature, such as Compustat. First, T2-LEAP provides a separate report for the wage bill of all firms. The wage bill is more reliable than some alternative measures of variable inputs used in the literature, i.e. the *cost of goods sold* (*COGS*) and the *operating expenses*. The reason is that the *COGS* undermeasures variable inputs and the *operating expenses* overmeasures variable inputs. Using the *COGS* (De Loecker and Eeckhout (2018a), Diez, Leigh and Tambunlertchai (2018) and De Loecker, Eeckhout and Unger (2020)) and the *operating expenses* (Traina (2018), Karabarbounis and Neiman (2018), and Flynn, Gandhi and Traina (2019)) as the measures of (the cost of) variable inputs results in biased estimates of markups. The *COGS* undermeasures variable inputs, and therefore, overestimates markups, while the *operating expenses* overmeasures variable inputs, which results in underestimating markups. The literature reports up to 50 percentage points discrepancies in the estimated markups, most of which can be attributed to the choice of the measure of variable inputs.<sup>1</sup> Second, T2-LEAP contains the information about financial statements on both publicly-traded and non-publicly-traded firms. The publicly-traded firms are usually bigger in size and may have larger markups. Hence, their estimated markups cannot be used as a proxy for the entire economy. Using the wage bill as the measure of variable inputs and considering both types of firms as reported in T2-LEAP, I find that the average gross (value-added) output markups increased mildly from 6% (24%) in 2002 to 9% (29%) in 2015.

The sensitivity of markups to the choice of measure of variable inputs and the

---

<sup>1</sup>Using the *COGS* (*operating expenses*) as a measure of variable inputs, the literature finds that markups increased from around 20% (10%) in 1980 to more than 60% (15%) in recent years.

lack of a good measure of variable inputs in accounting data make the “single-input” approach a less convincing method for estimating markups. For example, using aggregate data, researchers argue that the large markups estimated using firm-level data are not consistent with some aggregate variables such as profit share and returns to scale ([Ahmad, Fernald and Khan \(2019\)](#)), profit share and labour share ([Basu \(2019\)](#)), and no evidence for a quick rise of prices or plunge of marginal costs at the economy-wide level ([Syverson \(2019\)](#)). However, [De Loecker, Eeckhout and Unger \(2020\)](#) dispute the above results arguing that estimates from the aggregate data cannot capture the heterogeneity of firms. In Chapter 4, I introduce a new approach of estimating markups that captures the heterogeneity of firms, while it is less sensitive to the choice of measure of variable inputs.

In the new approach, I estimate markups by combining the two methodologies that are applied in the literature of estimating markups, i.e. the methodology that uses aggregate data and the one that uses firm-level data. In this approach, the markup is defined as a function of economic profit rates and returns to scale of firms. I attain the economic profit rates from income statements and balance sheets of firms. I obtain returns to scale by estimating the production function of firms using accounting data. I call this method the “total-input” approach because I use total inputs of firms from accounting data instead of relying on a single choice of variable input of production. Therefore, the result is less sensitive to the choice of measure of variable inputs. Using this approach that allows for firm heterogeneity while being less sensitive to the measure of variable inputs, I find that markups in the US increased from 7% in 1980 to 13% in 2018. This mild increase suggests that the wide gap between the

findings in the current literature that uses firm versus aggregate data is more likely due to the use of *COGS* as the measure of variable inputs, instead of heterogeneities across firms. I further find that the production technology across industries exhibits constant returns to scale over the same period of time.

# Chapter 2

## The Rise of Market Power, Inequality and the New Stylized Facts in the US Economy

### 2.1 Introduction

Since the early 1980s, the US corporate sector has faced the following new stylized facts. The labour share and the capital share have declined (see Figures 2.1 and 2.2).<sup>1</sup> The wealth-output ratio has increased (see Figure 2.3) while the rate of return on capital has decreased (see Figure 2.4). Over the same time, the market power and consequently, the profit share of the corporate sector has risen while the income

---

<sup>1</sup>The labour share is defined as the share of value added that paid out to the workers. Mathematically, the labour share,  $S_n$ , is equal to  $S_n = \frac{W_t N_t}{W_t N_t + R_t K_t + \Pi_t}$  where  $W_t$  is wage rate,  $N_t$  is total employment,  $R_t$  is the rental rate of capital,  $K_t$  is the value of capital, and finally,  $\Pi_t$  is the profit. The numerator of the ratio is the total payment to the workers and the denominator is the total value added of the economy. One can define the capital share and the profit share by the same method. See Subsection 2.3.7.

inequality has become more severe (for the rise of market power, see Appendix A.2 and for the corporate profit, see Figure 2.5). In this chapter, I argue that the rise of market power can provide a unified explanation for the above stylized facts when inequality is considerably high and households are heterogeneous in terms of their labour income.

The rise of market power that increases the profit of firms can affect the economy through two channels: the *direct effects* and the *indirect effects*. Based on the *direct effects*, when profit share increases, ceteris paribus, the labour and capital shares decrease by the same percent.<sup>1</sup> However, in the US economy, due to the large inequality in wealth and stock holdings, the increase in the profit has been distributed unequally among households.<sup>2</sup> The wealthy, who benefit more from the increasing profit, have lower marginal propensity to consume (MPC) and hence, save more in the form of capital. The increase in aggregate capital pushes down the rate of return on capital through the marginal product of capital (MPK). I call this process the *indirect effects* of the increase in the profit share. Labour and capital are complementary inputs because the elasticity of substitution between labour and capital is less than unity.<sup>3</sup> When the elasticity is less than unity, the increase in the capital-to-labour ratio results in a proportionally larger decrease in the relative price of capital-to-labour. Therefore, the *indirect effects* reduce the capital share more than the labour share. A model with homogeneous households can only capture the

---

<sup>2</sup>According to Wolff (2017), between 1983 to 2016, the top 10% held on average more than 80% of the stocks and the World Inequality Database (WID) reports that the bottom 50% of the US households hold almost zero wealth since the 1960s.

<sup>3</sup>Empirical evidence supports that the elasticity of substitution between labour and capital is less than one in the US (see Antras (2004) and Knoblich, Roessler and Zwerschke (2020)).

*direct effects*. In this chapter, I will show the importance of the *indirect effects* by using a model with heterogeneous households.

To study inequality in an imperfectly-competitive market, I present a model by combining the incomplete market model of [Aiyagari \(1994\)](#) with the imperfect competition model of [Dixit and Stiglitz \(1977\)](#). My model, however, is different from Aiyagari in the following ways. First, households in Aiyagari only hold one type of asset, while in my model, households have an extensive margin portfolio choice, i.e. they participate either in the stock or bond market, though not both at the same time, in each period. I segment stockholders from bondholders by modelling the stock market participation cost. Second, the market in Aiyagari is perfectly competitive, while firms in my model are monopolistically-competitive, so they set their prices above the marginal cost and generate profits. In other words, I combine the incomplete market model of [Aiyagari \(1994\)](#) with the imperfect competition model of [Dixit and Stiglitz \(1977\)](#). Combining these two features and adding portfolio choice and progressive tax enable my model to explain the secular trends while matching data on wealth distribution and stock market participation. For a general picture of the model, see [Figure 2.6](#).

To provide a unified explanation for the above stylized facts, I increase the market power by raising the value of markup in the Dixit-Stiglitz firms. As a result, the profit of firms increases and more of that will flow to wealthy households with lower MPC.<sup>4</sup>

---

<sup>4</sup>In a standard life-cycle permanent income model, consumption is proportional to permanent income because households can smooth their consumption by borrowing or lending their resources. Therefore, there is no heterogeneity in the MPC among households. However, in an incomplete market model, such as [Aiyagari \(1994\)](#), where households may face liquidity constraints and income uncertainty, the MPC is heterogeneous across households. In particular, the liquidity-constrained households have a higher MPC because they are hand-to-mouth and cannot smooth their consump-

It raises the total wealth and saving in the form of capital. The increase in aggregate capital pushes down the rental rate of capital. Therefore, an increase in the profit share of income decreases not only the labour and the capital shares but also the rental rate of capital while increases the wealth-output ratio. In other words, the rise of market power in the presence of high income inequality and heterogeneous households will provide a unified explanation to the stylized facts of the US economy.

This chapter contributes to the literature in the following ways. Most importantly, for the first time, I provide a unified explanation to the four stylized facts of the US economy. The literature presents separate partial explanations for each of these stylized facts. For example, [Elsby, Hobijn and Şahin \(2013\)](#), [Karabarbounis and Neiman \(2014\)](#), [Dorn, Katz, Patterson, Van Reenen et al. \(2017\)](#), [Autor, Dorn, Katz, Patterson and Van Reenen \(2017\)](#), [Kaymak and Schott \(2018\)](#) and [Martinez \(2018\)](#) study the decline of labour share (**S1**). [Barkai \(2020\)](#) and [De Loecker and Eeckhout \(2017\)](#) state that due to the rise of market power, not only the labour share (**S1**) but also the capital share (**S2**) have been declined. Regarding the third stylized facts, [Piketty \(2014\)](#), [Piketty and Zucman \(2014\)](#), [Piketty, Saez and Zucman \(2017\)](#), [Wolff \(2017\)](#) and [Auclert and Rognlie \(2018\)](#) and [Eggertsson, Robbins and Wold \(2018\)](#) investigate the increase in the wealth-output<sup>5</sup> ratio (**S3**). [Auclert and Rognlie \(2017\)](#) find that the rise of the top 1% increases the saving significantly, which leads to a decline between 0.45 pp to 0.85 pp in the long-run real interest rate (**S4**). Likewise,

---

tion. For the lower MPC of the wealthy, see also [McCarthy \(1995\)](#), [Carroll \(1997\)](#), [Jappelli and Pistaferri \(2014\)](#), and [Carroll, Slacalek, Tokunaka and White \(2017\)](#).

<sup>5</sup>In this chapter, similar to [Piketty \(2014\)](#), [Piketty and Zucman \(2014\)](#) and [Piketty, Saez and Zucman \(2017\)](#), the expression of wealth-output is equivalent to the capital-output ratio. See the accounting framework in Subsection 2.3.6. However, [Brun and González \(2017\)](#) and [Eggertsson, Robbins and Wold \(2018\)](#) disentangle wealth from working capital.

in this chapter, the decline in the real rate of return on capital obtained endogenously due to an increase in the saving (and wealth) of the rich. In contrast to [Auclert and Rognlie \(2017\)](#), in this chapter, the saving increases for any household who can pay the stock market participation cost, which is not necessarily the top percentile.

Some recent independent works have been done to answer some similar questions. [Brun and González \(2017\)](#) study the effect of an increase in Tobin's Q and changes of tax policies on some of the above variables. [Caballero, Farhi and Gourinchas \(2017\)](#), use an accounting framework to explain the effect of technical change, and increase in rent and risk premia on a group of stylized facts. [Eggertsson, Robbins and Wold \(2018\)](#) explore how the increase in markup and exogenous decrease in the real interest rate affect some similar stylized facts. [Farhi and Gourio \(2018\)](#) study the joint impacts of rising market power, rising unmeasured intangibles, and rising risk premia on increasing savings supply and slowing down of technological growth. [Stansbury and Summers \(2020\)](#) find that reducing worker power can provide a unified explanation for increasing market power, sluggish wage growth and declining labour share. However, this chapter is different from these papers, both theoretically and empirically.

The second contribution is about the problem of endogeneity. In the factor share literature, the fall in the rate of return on capital is considered exogenous. However, in this chapter, I claim that the decline in the rate of return on capital, at least partially, is endogenous when profit share increases and income inequality is severe. The endogeneity of return on capital may raise the possibility of an econometric issue. In particular, any regression model with the rate of return on capital as an

independent variable may suffer from the problem of endogeneity if the market is imperfect and inequality is high.<sup>6</sup>

Third, by finding the endogenous decline in the rate of return on capital, this chapter connects the two dominant strands of the literature. More precisely, in [Karabarbounis and Neiman \(2014\)](#), falling the rate of return on capital (which is equivalent to the decline in the relative price of investment) is the primary driver of labour share decline, while [Barkai \(2020\)](#), [Eggertsson, Robbins and Wold \(2018\)](#) and [De Loecker and Eeckhout \(2017\)](#) find that the rise of market power is the main driver of labour share decline. However, in this chapter, I show that when inequality is high, the increase in the markup (and profit) in the second strand leads to a decline in the rate of return on capital, which is the driver of the first strand of the literature. Hence, both strands are connected to some degree.

Finally, this chapter claims that the increase in profits provides a good explanation for the “factorless income,” i.e. the gap between the US GDP and the payments to input factors. [Karabarbounis and Neiman \(2019\)](#) explain that the issue of “factorless income” can be addressed by three possibilities: unmeasured capital, an increase in economic profits, and a decline in the real rate of capital at least since the 1980s. The authors cast doubt on the economic profit explanation and mostly focus on the real interest rate explanation. They believe that the decline in the real interest rate can be the mechanical driver of the increase in profits since the two variables have

---

<sup>6</sup>For example, the issue probably can be referred to the main regression model of the [Karabarbounis and Neiman \(2014\)](#)  $\frac{S_{L,j}}{1-S_{L,j}}\hat{S}_{L,j} = \gamma + (\sigma - 1)\hat{R}_j + u_j$ . In their model, change in labour share,  $\hat{S}_{L,j}$ , is a function of change in  $\hat{R}_j$  that is completely due to some exogenous shocks. Nonetheless, when market power rises and inequality is severe, change in  $\hat{R}_j$ , at least partially, can be endogenous.

negative comovement in data. In this chapter, however, I highlight the possibility of causality in the opposite direction. In particular, I show that when income inequality is high, an increase in profit causes a decline in the rate of return on capital.

## 2.2 The new stylized facts

To demonstrate the new stylized facts, I compare the stylized facts of interest obtained from the average value of data in the first decade (1982-1991) against the ones attained from the average value of data in the last decade (2007-2016).

**Stylized fact 1: The labour share decline.** The labour share in the non-financial corporate sector has declined by 9%, from 64% of Gross Value Added (GVA) in the first decade to 58% of GVA in the last decade (see Figure 2.1). Payment to workers consists of wages and salaries of employees, plus social and private benefit plans and other fringe benefits. Data are taken from the Bureau of Labour Statistics (BLS) and the Bureau of Economic Analysis (BEA).

**Stylized fact 2: The capital share decline.** The capital share has declined from 25% of GVA in the first decade to 17% of GVA in the last decade (see Figure 2.2). The capital share declined more than the labour share due to decrease in the rate of return of capital, which outweighs the rise of aggregate capital. Data are extracted from BEA (table 1.4 of fixed assets in NIPA) and Moody's (retrieved from FRED).

**Stylized fact 3: The increase in wealth-output ratio.** The net worth over output has increased from 210% in the first decade to 300% in the last decade (see

Figure 2.3). The net worth consists of equity (directly and indirectly held), pensions, real state, financial assets, consumer durables and proprietors' equity in non-corporate subtracting total liabilities. Considering the trend and the fluctuations of corporate equity over GVA, one can conclude that the trend of net worth over output mostly is coming from the behaviour of the corporate sector. Data are obtained from Survey of Consumer Finance (SCF) (z1 visualization table) and BEA.<sup>7</sup>

**Stylized fact 4: The decline in the rate of return on capital.** As discussed by Eggertsson, Robbins and Wold (2018), the corporate firms mostly own capital rather than rent them from bondholders, and consequently, it is difficult to observe a rental rate of capital. Hence, following the works of Barkai (2020) and Eggertsson, Robbins and Wold (2018) that are based on Hall and Jorgenson (1969), I obtain the opportunity cost of capital, which is defined as the required rate of return on capital

$$R = (R^b - E[\pi] + \delta), \quad (2.2.1)$$

where  $R$  is the required rate of return on capital, and  $R^b$  is the corporate bond yield that is obtained as an average of Moody's Aaa and Baa.  $E[\pi]$  is the expected inflation rate based on a five-year moving average, and  $\delta$  is the depreciation rate of capital in the non-financial corporate sector.<sup>8</sup> The rate of return on capital declined from 15% in the first decade to 10% in the last decade (see Figure 2.4). Data on the rate

---

<sup>7</sup>Based on data from Piketty, Saez and Zucman (2017), Eggertsson, Robbins and Wold (2018) find that the wealth-output ratio increased from 250% in 1980 to almost 400% in 2015.

<sup>8</sup>By considering the relative price of investment over consumption, Eggertsson, Robbins and Wold (2018) expanded the equation to  $R = \frac{P_k}{P_c}(R^b - E[\pi] + \delta + (1 - \delta)E(g_{pk}))$  where  $P_k$  is the price of investment,  $P_c$  is the price of consumption and  $E(g_{pk})$  is the gain from holding capital instead of a consumption good.

of return on bonds is obtained from Moody's that is retrieved from Fred. Data for the stock of capital and the inflation rate are taken from the table of fixed assets in NIPA.

**The rise of profit share.** The market power, and hence the profitability of firms, which is measured by the level of markup, has increased in the non-financial corporate sector of the US since the early 1980s. In particular, the pre-tax profit has increased from 10% of GVA in the first decade to 13% of GVA in the last decade (see Figure 2.5).<sup>9</sup> I obtain the data for the GVA and pre-tax profit with inventory valuation and capital consumption adjustment from BEA and BLS, respectively. In Chapter 4, I estimate the rise of market power in the US economy using firm-level data. Also, see Appendix A.2 for a literature review of the studies that measure the rising market power in the US.

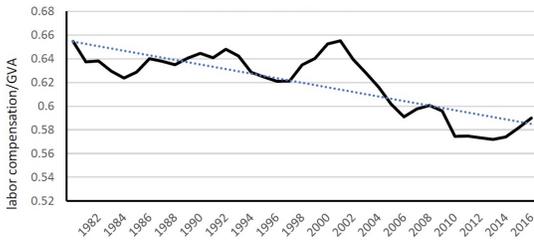


Figure 2.1: The labour share

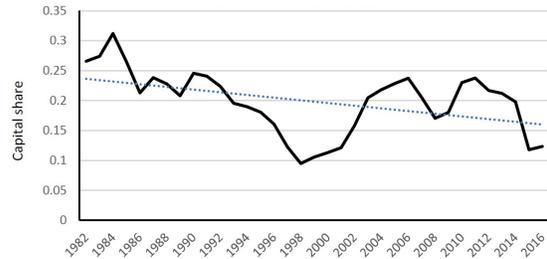


Figure 2.2: The capital share

<sup>9</sup>Due to the change of tax policies, after-tax profit increased even more. For more details about the effect of tax cut on factor shares, see [Brun and González \(2017\)](#).

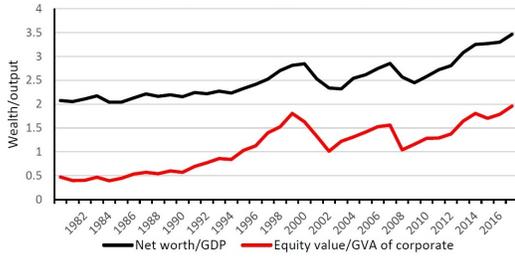


Figure 2.3: The wealth-output ratio

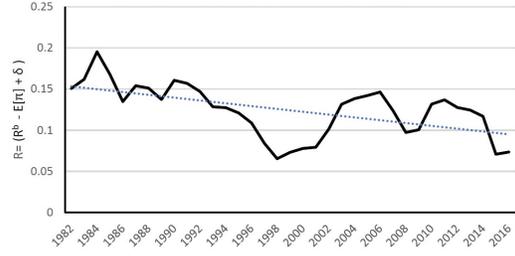


Figure 2.4: The rate of return on capital

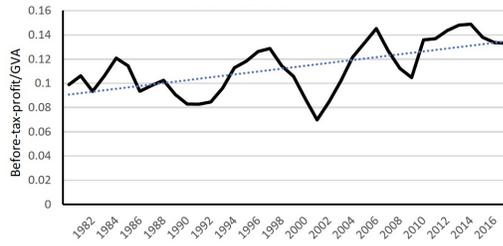


Figure 2.5: The pre-tax profit share of GVA

## 2.3 Model

### 2.3.1 Model description

This chapter combines the incomplete market model of [Aiyagari \(1994\)](#) with the imperfect competition model of [Dixit and Stiglitz \(1977\)](#) to study the effect of exogenous markup change on explaining the macroeconomic stylized facts. Households are different in terms of their productivity (labour income) which follows a Markov process. They allocate their saving either in bonds (and partially become the owner of capital) or stocks (and become the owner of firms and partially the owner of capital). While the return to the capital of bondholders are determined by the MPK, stockholders receive the profit of firms on the top of MPK since they are the owners of firms.<sup>10</sup>

<sup>10</sup>For more explanation, see Equations 2.3.16 to 2.3.18 in Subsection 2.3.6, the accounting framework.

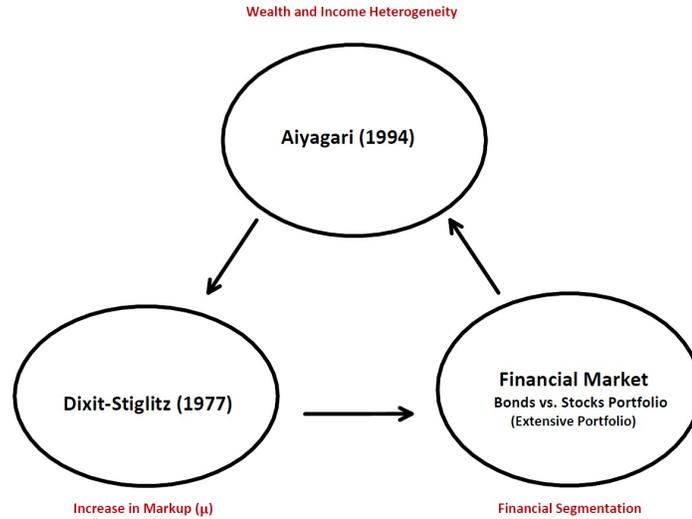


Figure 2.6: The overall picture of the model.

Notes: Households with heterogeneous income supply their labour in an imperfect market with monopolistically competitive firms and invest their saving in a financial market. In the financial market, the wealthy have higher income and lower MPC. Therefore, they can pay the stock market participation cost and benefit proportionally more than others from the increasing profit of the firms.

The firms finance their investment either by issuing stocks or bonds. For simplicity, I assume that the price of stocks is constant, and therefore, financing through stocks means issuing new stocks at a constant price. There is no aggregate technology shock in this economy; hence we have to assume that stock return is risk-free. Participation in the stock market is costly. To segment stockholders from bondholders, similar to the [Gomes and Michaelides \(2005\)](#), stock buyers have to pay a participation cost in order to participate in the stock market. Intermediate firms in Dixit-Stiglitz framework have downward input demand functions, and they set prices above their marginal costs. When markup increases, the firms become more profitable and the return to stocks increases. Figure 2.6 provides a visual summary of the model.

### 2.3.2 Households

Households are heterogeneous in terms of their productivity (income) which leads to heterogeneity in asset holding. The productivity shock follows three states,  $z_{j,t} = [z_{low}, z_{mid}, z_{hi}]$ , which is randomly drawn from the Markov process with transition matrix  $\Omega(z'|z)$ . Households supply their labour inelastically, and face uninsured idiosyncratic shock that affects their productivity level in each period (year).<sup>11</sup> They maximize their expected utility as follows:

$$\max_{b_{j,t+1}, s_{j,t+1}, c_{j,t}} E_0 \sum_{t=0}^{\infty} \beta^t U(c_{j,t}) \quad (2.3.1)$$

subject to

$$s_{j,t+1} + b_{j,t+1} + c_{j,t} = z_{j,t}W_t + (1 + R_t^b)b_{j,t} + (1 + R_t^s)s_{j,t} - T(m_{j,t}) - F(W_t)I_p, \quad (2.3.2)$$

where the utility function has a simple log form,  $U(c_{j,t}) = \log c_{j,t}$ , which is a special case of CRRA utility function.  $I_p$  is a dummy variable that is equal to one when households want to buy stocks and zero otherwise.  $W_t$  is the annual wage (salary) per efficiency unit of labour.  $F(W_t)$  is the participation cost that the household has to pay in order to invest in the stock market in each period. The stock market participation cost can be interpreted as the opportunity cost that occurs because of the time spending on collecting and analyzing the financial information and the paper

---

<sup>11</sup>I make the assumption of fixed labour supply due to computational cost. The assumption of fixed labour supply should not affect the result much as long as the income and substitution effects cancel out each other. For example, [Barro and Furman \(2018\)](#) apply the same assumption to study the macroeconomic effects of tax reform.

works associated with entering the financial market (see [Vissing-Jorgensen \(2002\)](#), [Gomes and Michaelides \(2005\)](#), [Alan \(2006\)](#), and [Khorunzhina \(2013\)](#)). Similar to [Gomes and Michaelides \(2005\)](#), the total participation cost is proportional to the expected annual salary of the household,  $E(z)W_t$ . Thus, it is defined as a function of the expected annual salary of the household and not the size of the investment.<sup>12</sup>  $T(m_{j,t})$  is a progressive tax on total income, i.e. labour and capital income of households.  $R_t^b$  and  $R_t^s$  are the realized rates of return on invested bonds and stocks in the last period. The asset market is incomplete. No-short sale constraint in the model prevents households from holding negative assets, i.e.,

$$s_{j,t} \geq 0 \tag{2.3.3}$$

$$b_{j,t} \geq 0 \tag{2.3.4}$$

There is no aggregate technology shock in this economy, and hence, it is assumed that return on stocks is risk-free. No risk in stock return seems unrealistic. However, the focus of this chapter is neither on the transition path of aggregate shock nor on the (intensive margin) portfolio choice per se, but to generate the observed pattern that stocks are disproportionately held by the wealthy households in the steady states.

---

<sup>12</sup>The parameter is calibrated to match the participation rate of 35% in the 1980s, the beginning of my sample period. It also matches well the participation rate of 50% in the 2010s (the end of my sample period) even though it is not targeted. In particular, I calibrate the participation cost as 5.75% of the expected annual salary of a household, i.e.  $F(W_t) = 5.75\%E(z)W_t$ . With the calibrated value, the participation rate from my model hits the target moment well. [Khorunzhina \(2013\)](#) estimates that the magnitude of average stock participation cost is between 4% to 6% of the labour income. My calibrated result is in line with the estimated average stock participation cost in [Khorunzhina \(2013\)](#), but it is higher than the estimated result of [Alan \(2006\)](#), where the stock market participation cost is estimated at around 2% of the permanent component of the annual labour income.

Participation cost serves well for this purpose. Introducing aggregate uncertainty will complicate this model without adding more intuition to the question at hand.

In this model, the households buy stocks in a period when they are highly productive or have already accumulated enough wealth. They buy stocks only if the return on stocks net of the participation cost is greater than the return on bonds. They buy bonds if they end up with low productivity or low saving. Since the stock return is risk-free, within each period, households know for sure that participating in either the stock or bond market maximize their objective functions, and hence, they have no incentive for making (intensive margin) portfolio choice to reduce the risk.<sup>13</sup>

### 2.3.3 The progressive tax system

To understand how a change in income distribution maps into a change in wealth distribution, one needs to capture the progressivity in income tax. Households' income distribution is different. Thus, the distribution of tax burden across households is also different. The progressive tax function is defined such that marginal and average tax rates are increasing with income.<sup>14</sup> Following the works of [Gouveia and Strauss \(1994\)](#) and [Bachmann, Bai, Lee and Zhang \(2017\)](#), I define the progressive income

---

<sup>13</sup>The assets, stocks and bonds, are “fruits” on a tree. Stocks can be “sweater” than bonds as they can deliver higher return but these fruits are higher up the tree and difficult to get to (participation cost). There is no risk that the stock-fruit will turn out to be not sweat (risk free).

<sup>14</sup>This non-linear part of the function is taken from [Gouveia and Strauss \(1994\)](#) which is based on data of 1989. However, over time the tax rate is changing from convex to concave function, and hence, the marginal tax rate is decreasing in income. For more details about changing the tax rate over time, see [Piketty and Saez \(2007\)](#).

tax function as follows:

$$T(m_t) = \begin{cases} \tau_1 \left[ m_t - \left( m_t^{-\tau_2} + \tau_3 \right)^{\frac{1}{-\tau_2}} \right] + \tau_0 m_t & \text{if } m_t > 0 \\ 0 & \text{if } m_t = 0, \end{cases} \quad (2.3.5)$$

where the total taxable income of households,  $m_t$ , is equal to

$$m_t = z_{j,t}W_t + R_t^b b_{j,t} + R_t^s s_{j,t} - F(W_t)I_p, \quad (2.3.6)$$

and  $(\tau_0, \tau_1, \tau_2, \tau_3)$  is a vector of tax coefficients. The first term of Equation 2.3.5 describes the burden of effective federal tax income on US households. [Gouveia and Strauss \(1994\)](#) calibrate  $\tau_1$  and  $\tau_2$  such that to estimate the average tax specification for each year between 1979 and 1989.<sup>15</sup> The second term,  $\tau_0 m_t$ , captures the other tax revenue, including state income taxes, property taxes, and excise taxes.  $\tau_0$  is calibrated based on [Bachmann, Bai, Lee and Zhang \(2017\)](#).  $\tau_3$  is a scale-dependent and model-specific parameter, i.e. it does not come from real data directly. I scaled up  $\tau_3$  such that the model matches real data.

The government in this economy has a passive role. They define the tax system, collect the tax revenue and spend it without any effects on households' utility. In other words, the tax system is not for the purpose of redistribution. The chapter does not claim to reject or underestimate the effect of transfers and tax redistribution on factor shares; instead, it mostly focuses on the impact of gross profit distribution

---

<sup>15</sup>The calibration of  $\tau_1$  and  $\tau_2$  in the literature is usually based on the year 1989 of [Gouveia and Strauss \(1994\)](#)'s paper.

among heterogeneous households when inequality is high.

### 2.3.4 Firms

Firms produce in a standard Dixit-Stiglitz framework. In particular, the monopolistically-competitive firms, which are owned by stockholders, produce intermediate good  $i$  according to the constant return to scale (CRS) technology,  $y_{i,t} = F(k_{i,t}, n_{i,t})$ .  $y_{i,t}$  is the (intermediate) output per unit of efficiency labour;  $k_{i,t}$  is the efficiency units of capital rented from the bondholders at price  $R_{i,t}$ ; and  $n_{i,t}$  is the efficiency units of labour rented from households at price  $W_{i,t}$ . In order to optimize the utility of stockholders, the monopolistic intermediate firms maximize their profit subject to the CES technology:

$$\max_{y_{i,t}, k_{i,t}, n_{i,t}} \Pi_{i,t} = p_{i,t}(y_{i,t})y_{i,t} - R_{i,t}k_{i,t} - W_{i,t}n_{i,t}, \quad (2.3.7)$$

subject to their CES technology

$$y_{i,t} = \left[ \alpha(k_{i,t})^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(n_{i,t})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \quad (2.3.8)$$

The First Order Conditions (F.O.Cs) with respect to inputs and normalizing the output price (i.e.  $p_{it} = 1$ ), imply

$$R_{i,t} = \frac{F_{k_{i,t}}}{\mu_t} = \frac{\alpha \left( \frac{y_{i,t}}{k_{i,t}} \right)^{\frac{1}{\sigma}}}{\mu_t}, \quad (2.3.9)$$

$$W_{i,t} = \frac{F_{n_{i,t}}}{\mu_t} = \frac{(1 - \alpha) \left( \frac{y_{i,t}}{n_{i,t}} \right)^{\frac{1}{\sigma}}}{\mu_t}. \quad (2.3.10)$$

I can obtain the price of both input factors,  $R_{i,t}$  and  $W_{i,t}$ , as a function of efficient capital by normalizing the labour input.

The monopolistically-competitive firms face a downward sloping demand curve and they charge a markup over their marginal costs. The markup,  $\mu_t = \frac{\epsilon_t}{\epsilon_t - 1} \geq 1$ , indicates that the firms set their output price above the marginal cost and generate a profit (for more details, see Appendix A.1). By substituting the optimal price of labour and capital into Equation 2.3.7, I obtain the profit function of the firms:

$$\Pi_{i,t} = \left(1 - \frac{1}{\mu_t}\right) y_{i,t}. \quad (2.3.11)$$

### 2.3.5 Equilibrium

On the firms' side of the economy, there is symmetric equilibrium such that

$$y_{i,t} = Y_t,$$

$$p_{i,t} = P_t^Y = 1,$$

$$k_{i,t} = K_t,$$

$$n_{i,t} = N_t = 1,$$

$$R_{i,t} = R_t, \text{ and}$$

$$W_{i,t} = W_t,$$

where,  $n_{i,t}$  is the firm's labour demand and  $N_t$  is the aggregate demand. Similarly,  $k_{i,t}$  is the demand for capital and  $K_t$  is the aggregate one. Here, general equilibrium is

characterized as the path of firms and households decision  $(s_t, b_t, c_t, k_t)$ , input prices  $(W_t, R_t)$ , profits  $\Pi_t$ , and the distribution of  $\theta_t$  (the joint distribution of  $ds$ ,  $db$ , and  $dz$ ) along with aggregate quantities that are derived from the following conditions:

1. Firms maximize their profit given the exogenous markup,  $\mu_t$ .
2. Households maximize their objective functions given equilibrium prices  $W_t^*$ ,  $R_t^*$  and optimized profit  $\Pi_t^*$ .
3. Bond market clears,

$$B_t = \int b_{j,t} d\theta_t, \quad (2.3.12)$$

where,  $b_{j,t}$  is the total demand of bonds by households and  $B_t$  is the total supply of bonds by firms.

4. Stock market clears,

$$S_t = \int s_{j,t} d\theta_t, \quad (2.3.13)$$

where,  $S_t$  is the total supply of stocks by intermediate firms and  $s_{j,t}$  is the  $j$ 's household demand for stocks. Note that total demand for assets,  $a_{j,t} = b_{j,t} + s_{j,t}$ , is equal to the supply of assets which is equal to the total capital of firms,  $K_t$ .

5. Government budget constraint is balanced,

$$G_t = \int_0^1 T(y_{j,t}) d\theta_t. \quad (2.3.14)$$

6. Goods market clears,

$$Y_t = C_t + I_t + \int_0^1 F(s)I_p d\theta_t + G_t, \quad (2.3.15)$$

where,  $I_t$ ,  $\int_0^1 F(s)I_p d\theta_t$ , and  $G_t$  are aggregate investment, the summation of participation costs for all individuals and government expenditures, respectively.

### 2.3.6 The accounting of assets

In this economy, the total assets of households are a combination of their total bond and stock holdings.

$$A_t = B_t + S_t. \quad (2.3.16)$$

Moreover, I assume that the following model of accounting holds for the firms' capital input in the corporate sector:

$$K_t = B_t + S_t. \quad (2.3.17)$$

The above equation means that the value of capital input on the firms' side is equal to the summation of total bonds and stocks that firms issue. In other words, any change in either bonds or stocks affects the total value of firms' capital. Hence, payment to both bondholders and stockholders is a function of aggregate MPK. To be more precise, the total income of each firm,  $Y_t$ , is distributed among bondholders, stockholders and workers as follows:

- Bondholders receive  $(R_t - \delta)B_t$ , and therefore the corporate bond rate is

$$R^b = (R_t - \delta). \quad (2.3.18)$$

- Workers are paid based on the marginal product of labour,  $W_t n_t$ .
- The remaining income,  $Y_t - W_t n_t - (R_t - \delta)B_t = (R_t - \delta)S_t + \Pi_t$ , is distributed among stockholders.<sup>16</sup> Hence, dividend per unit of stock is equal to

$$d_{i,t+1} = (R_t - \delta) + \frac{\Pi_t}{S_t}. \quad (2.3.19)$$

In the financial market, households demand stocks instead of bonds as long as the return of stocks net of participation cost is greater than the return on bonds:

$$d_{i,t+1} - \frac{F(W_t)}{S_t} \geq (R_t - \delta), \quad (2.3.20)$$

where,  $\frac{F(W_t)}{S_t}$  is the participation cost per unit of stock.

### 2.3.7 The labour share

By combining the F.O.Cs of firms with the definition of the factor shares

$$S_{N,t} = \frac{W_t N_t}{\mu_t (W_t N_t + R_t K_t)}, \quad (2.3.21)$$

---

<sup>16</sup>Here,  $\Pi_t$  is the economic profit which is different from the accounting profit in terms of the opportunity cost of capital. From the accounting point of view, the whole term of  $(R_t - \delta)S_t + \Pi_t$  that will be distributed among stockholders is considered as profit.

$$S_{K,t} = \frac{R_t K_t}{\mu_t (W_t N_t + R_t K_t)}, \quad (2.3.22)$$

$$S_{\pi,t} = \frac{\Pi_t}{Y_t} = 1 - \frac{1}{\mu_t}, \quad (2.3.23)$$

where

$$S_{N,t} + S_{K,t} + S_{\pi,t} = 1, \quad (2.3.24)$$

I obtain the equation of the labour share as follows:<sup>17</sup>

$$S_{N,t} = \left(\frac{1}{\mu_t}\right) \left(1 - \alpha^\sigma \left(\frac{1}{\mu_t R_t}\right)^{\sigma-1}\right). \quad (2.3.25)$$

From there, change in the labour share,  $S_{N,t}$ , can be calculated following an increase in markup,  $\mu_t$ .

### 2.3.8 Computation strategy

There is no aggregate shock in this economy. To determine the cross-sectional distribution of wealth and labour income, I only need to obtain the time distribution of wealth and labour income for one household across a long period of time. I can obtain the distributions by recursion of one household's behaviour, given the exogenous states and prices in the economy. For any exogenous markup, tax policy, and given the initial level of bonds,  $B$ , and stocks,  $S$ , I obtain the prices ( $R$  and  $W$ ), and the profit ( $\Pi$ ). Considering the prices and the profit, households maximize their objective function recursively as follows:

---

<sup>17</sup>Alternatively, I can compute the labour share using the definition of labour share in Equation 2.3.21 because the values of all required variables are generated by the simulation. The same can be done for the capital and profit shares.

Bondholders maximize their objective function according to

$$v(a, z, b) = \max_{a', c} \{u(c) + \beta \max_{\phi'} E[v(a', z', \phi') | a, z, \phi]\} \quad (2.3.26)$$

subject to

$$a' = (1 + R^b)a + zW - T(y) - c \quad (2.3.27)$$

$$a' \geq 0, \quad (2.3.28)$$

where,  $a'$  is the saving amount of the household and  $\phi = b$  determines that they enter the period as a bondholder. They maximize their objective function in the next period by either staying in the bond market or switching to the stock market. In other words, they optimally choose  $\phi' = b'$  or  $s'$ . Similarly, stockholders maximize their objective function such that

$$v(a, z, s) = \max_{a', c} \{u(c) + \beta \max_{\phi'} E[v(a', z', \phi') | a, z, \phi]\} \quad (2.3.29)$$

subject to

$$a' = (1 + R^s)a + zW - T(y) - F(w) - c \quad (2.3.30)$$

$$a' \geq 0. \quad (2.3.31)$$

In this case, the household enters the period as a stockholder. Given the process for their productivity, they decide how much to save and what kind of asset to demand, i.e. they maximize their objective function given the joint distribution of  $\theta(b, s, z)$ .

In the absence of the aggregate shock, I can obtain the cross-sectional distribution of assets for the whole economy by the repetition of assets' demand for one household. Therefore, in order to obtain the aggregate demand of assets, I simulate the Markov chain for the labour income shock that is randomly assigned to the household and obtain a series of 40,000 draws.<sup>18</sup>

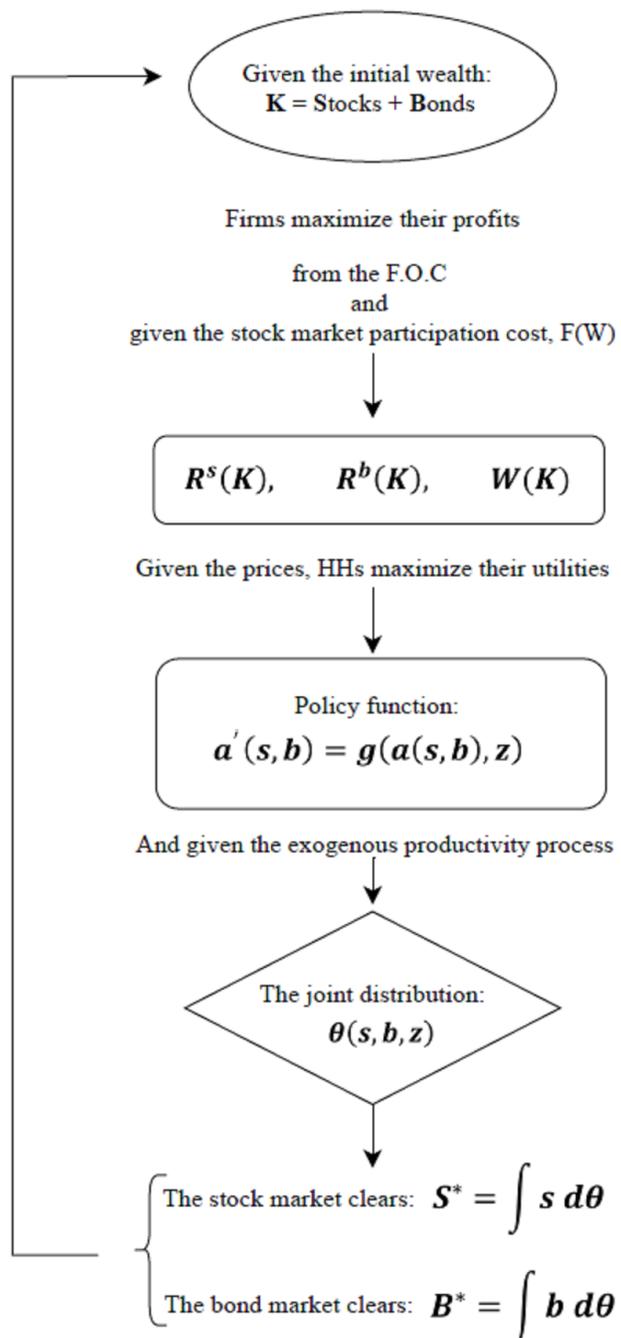
To find the general equilibrium values of aggregate stocks and bonds, I clear the markets and obtain the final values of stocks,  $S^*$ , and bonds,  $B^*$ . If the final values of  $S^*$  and  $B^*$  are different from the initial guesses, I update the guesses and continue the above process until the joint difference between the final assets ( $S^*$  and  $B^*$ ) and the initial guesses ( $B$  and  $S$ ) converges to zero. Mathematically, it is a fixed point iteration method in two dimensions of stock and bond. For a descriptive algorithm, see Figure 2.7.

In terms of the computational strategy, one of the main differences and complications of my model relative to Aiyagari (1994) is that I have to clear not only the bond market but also the stock market simultaneously. To overcome the challenge, I apply a two-dimensional bisection search algorithm. Solving such an algorithm is more complicated, and the sensitivity of the results to any change in the values of variables increases the search burden.

---

<sup>18</sup> Aiyagari (1994) uses a series of 10,000 draws. My results is also robust for as low as 10,000 draws.

Figure 2.7: The algorithm of the model



Notes: The result simulation will be optimum if the joint difference between the initial guesses of assets ( $S, B$ ) and the final ones ( $S^*, B^*$ ) is less than a certain small threshold value. For this purpose, I continue to update the initial guesses of the assets until the joint difference converges to zero.

## 2.4 Results

### 2.4.1 Calibration

Table 2.1 displays the calibration of parameters which I mainly borrow from the literature. Among them, the distribution factor ( $\alpha$ ) is calculated as 0.35, which is the average value of the parameter in the heterogeneous-agent literature.<sup>19</sup> Following the work of Antras (2004) for the Hicks-Neutral technology, the value of the factor elasticity of substitution is considered less than one,  $\sigma = 0.95$ . When the elasticity of substitution is less than unity, labour and capital are complementary. The household income process,  $z$ , is obtained from Davila, Hong, Krusell and Ríos-Rull (2012) and normalized to  $[z_{low} = 0.06, z_{mid} = 0.3, z_{hi} = 2.64]$ .<sup>20</sup> The transition matrix  $\Omega(z'|z)$  is calibrated such that the distribution of households income is right-skewed.<sup>21</sup> I borrow the progressive income tax parameters from the literature. In particular,  $\tau_0 = 0.052$  captures the effects of non-federal-income tax parameters and it is taken from Bachmann, Bai, Lee and Zhang (2017). The parameter values  $\tau_1 = 0.258$  and  $\tau_2 = 0.768$ , which reflect the progressive schedule of federal income tax, are taken from Gouveia and Strauss (1994). Finally, to match the model data with real data, I set the scale parameter  $\tau_3$  to 0.8537.

---

<sup>19</sup>In the literature of heterogeneous households, Aiyagari (1994) and Krusell and Smith (1998) calibrate it as 0.36, while Kaplan, Moll and Violante (2018) define it as equal to 0.33.

<sup>20</sup>By normalization of the income process, I mean that the expected income of a household is equal to one, i.e.  $E(z) = 1$ .

<sup>21</sup>Davila, Hong, Krusell and Ríos-Rull (2012) estimate the stationary distribution as  $z^* = [0.498, 0.443, 0.059]$ , which means the top 5.9% earn 44 times as much as the bottom 50% earn.

Table 2.1: Calibration

Parameter	Value	Description	Source/Target
$\alpha$	0.35	The distribution factor	The average of the heterogeneous-agent literature
$\beta$	0.95	Annual discount factor	<a href="#">Kaplan, Moll and Violante (2018)</a>
$\delta$	0.10	Annual depreciation rate	<a href="#">Krusell and Smith (1998)</a>
$\sigma$	0.95	The elasticity of substitution	<a href="#">Antras (2004)</a> /Estimated for the Hick-Neutral technology
$\tau_0$	0.052	Income tax parameter	<a href="#">Bachmann et al. (2017)</a> /Targeted the non-federal-income tax
$\tau_1$	0.258	Income tax parameter	<a href="#">Gouveia and Strauss (1994)</a> /Targeted the federal-income tax
$\tau_2$	0.768	Income tax parameter	<a href="#">Gouveia and Strauss (1994)</a> /Targeted the federal-income tax
$\tau_3$	0.8537	Income tax parameter	Scaled up to match data
$F(W_t)$	$0.0575E(z)W_t$	Participation cost of stock market	To target the participation rate in data
Income Process			
Labour productivity shock	$z = [0.06, 0.3, 2.64]$		<a href="#">Davila, Hong, Krusell and Ríos-Rull (2012)</a>
Transition matrix	$\Omega(z' z) = \begin{bmatrix} 0.992 & 0.008 & 0 \\ 0.009 & 0.980 & 0.011 \\ 0 & 0.083 & 0.917 \end{bmatrix}$		<a href="#">Davila, Hong, Krusell and Ríos-Rull (2012)</a>

### 2.4.2 The initial distribution of wealth and factor shares

The first row of Table 2.2 illustrates the distribution of wealth, labour share, capital share, the wealth-output ratio and the rate of return on capital for the initial steady state of the model when  $\mu = 1.06$ . The second row shows the values of the corresponding variable in the data which is based on the average of observed data from 1982 to 1991. The estimation of these variables is discussed in Section 2.2. The data on the top 10% and 20% and the wealth Gini are taken from the World Inequality Database. As one can see from the table, the model can predict data well except for the wealth distribution of the top 10%. It has been known that the income inequality alone is not enough to match the top percentiles of the wealth distribution. Thus, [Krusell and Smith \(1998\)](#), for example, use preference heterogeneity at the same time. Although studying inequality is not the primary focus but just a tool to study the stylized facts, my model does a better job compared to the works of [Aiyagari \(1994\)](#) and [Krusell and Smith \(1998\)](#) without preference heterogeneity due to higher return to the wealthy. This model generates a wealth Gini coefficient of 0.77, which is very

close to the observed Gini coefficient of 0.79 in real data, in comparison with 0.32 and 0.25 in their papers, respectively. According to data, the top 10% and the top 20% of households own 61% and 82% of the total wealth, respectively. In my model, they hold 53% and 80% of the total wealth compared with the benchmark model of [Krusell and Smith \(1998\)](#) where the top 10% and the top 20% hold 19% and 35% of the total wealth, respectively.<sup>22</sup>

Regarding the stylized facts, while my model generates a similar labour share and a slightly higher capital share, it generates a higher wealth-output ratio and a lower rate of return on capital compared to data. In the next step, I increase the markup of the imperfectly-competitive firms in order to study the effects of the rising market power on the stylized facts in the two steady states.

Table 2.2: The initial distribution of wealth and the stylized facts in the baseline model and data

Target	The percentage of wealth held by			Gini	Labour share	Capital share	Wealth/output	$R = (R^b + \delta)$
	Top 10%	Top 20%	Top 50%					
Model ( $\mu = 1.06$ )	53%	80%	100%	0.77	65%	29%	250%	11%
Data (1982-1991)	61%	82%	99%	0.79	64%	25%	210%	15%

### 2.4.3 An increase in the markup

The rise of market power and the increasing profit of firms are measured by the level of markup. Although it is widely accepted by scholars that markup has increased in the US since the early 1980s, they have different ideas about the size of the increase in markup. Based on the estimation methods and available data, they find that the

<sup>22</sup>Following the introducing a preference heterogeneity by [Krusell and Smith \(1998\)](#), the wealth share of the top 10% and the top 20% increases to 73% and 88%, respectively, and the Gini coefficient becomes to 0.82.

increase in markup varies from five to around 50 percentage points in the last 35 years.<sup>23</sup> In Chapter 4, I estimate that markups, on average, increased from around 6% in the 1980s to around 13% in the 2010s. So, as a benchmark, I consider a 7 percentage-point increase in markup and explore the effects of this increase on the stylized facts in the two steady states. I particularly increase the markup from 6% ( $\mu_1 = 1.06$ ) in the initial steady state to 13% ( $\mu_2 = 1.13$ ) in the final steady state. To do a robustness check, I increase the markup to 20%, which is the median of the estimated markups in the literature in the recent decade (see Appendix A.2 for the review of markup estimates in the US economy). Some papers, such as [De Loecker and Eeckhout \(2017\)](#), estimate that markups increased up to 67%. Applying such an increase in the markup results in a stock return that is too high to be compatible with stable wealth distribution in the model.

#### 2.4.4 Results

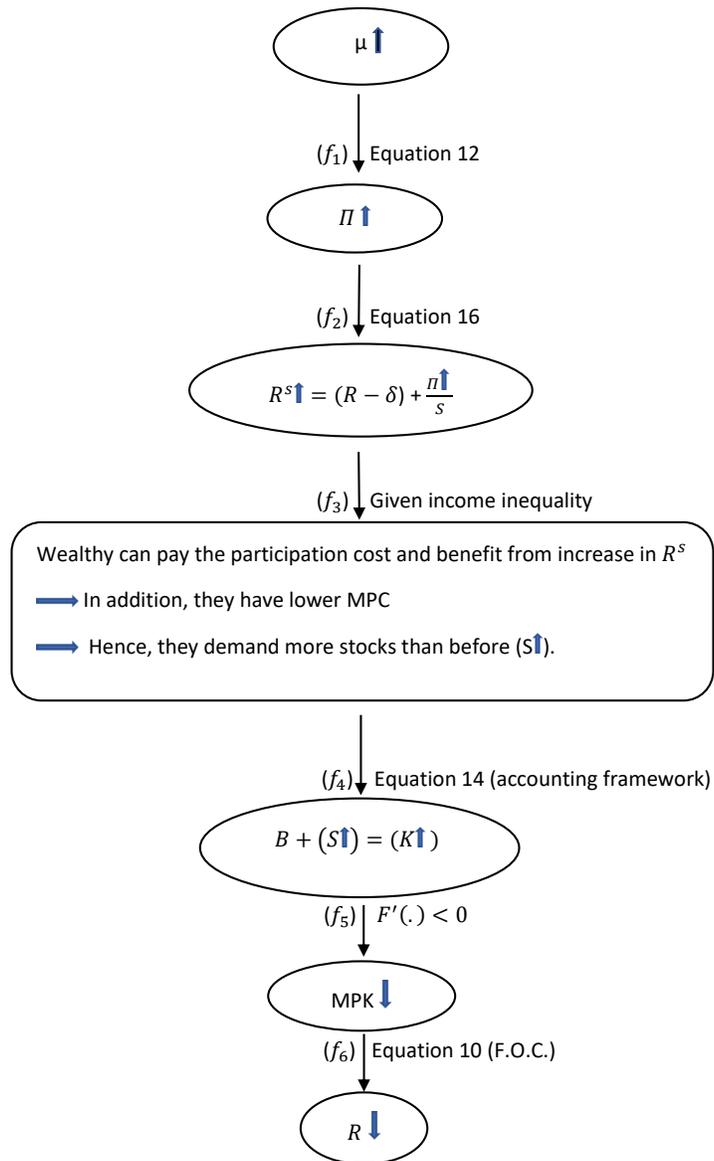
In this model, the increase in markup generates the stylized facts that I documented earlier. Now, I explain the mechanisms that determine these stylized facts. For the purpose of clarity, I display these mechanisms in Figure 2.8 and provide the explanations below.

Following an increase in the markup, the corporate sector becomes more profitable

---

<sup>23</sup>For example, [De Loecker and Eeckhout \(2017\)](#) estimate that markups are increasing from around 18% in the early 1980s to around 67% in recent years. However, some other papers discuss the unfeasibility of a very high increase in the markups. Among them, [Flynn, Gandhi and Traina \(2019\)](#) find that the increase in the markups can be as low as five percentage points, i.e. increasing from 4% in the 1980s to 9% in recent years. [Traina \(2018\)](#) estimates that markups only increase from 9% to around 15%. [Ahmad, Fernald and Khan \(2019\)](#) show that a markup of 12% is high enough to be compatible with the returns to scale and the profit share of the US economy.

Figure 2.8: The flow of logic behind the model



( $f_1$ ). The increase in profit results in a higher rate of return on the stock ( $f_2$ ). Income inequality that is generated exogenously by the model leads to wealth inequality ( $f_3$ ). Wealthy households participate actively in the stock market, and hence, they benefit more from the increased stock return. In addition, the wealthy have lower MPC and consequently, their demand for the stocks increases.<sup>24</sup> According to the accounting framework, higher demand for stocks results in higher demand for capital and, therefore, the wealth-output ratio rises ( $f_4$ ). The higher amount of capital results in lower MPK ( $f_5$ ) that leads to a decline in the real rate of return on capital ( $f_6$ ).

Moreover, the price of capital,  $R_t$ , and the amount of capital,  $K_t$ , are connected through another channel, i.e. the elasticity of factor substitution,  $\sigma$ . In particular, for any value of  $\sigma$ , given the change in factor ratio ( $K/N$ ), the change in price ratio ( $W/R$ ) is determined.<sup>25</sup> The elasticity of substitution plays the role of a multiplier. When elasticity is greater than one, an increase in the price ratio is less than an increase in the factor ratio in percentage. However, when the elasticity is less than one, an increase in the price ratio is greater than an increase in factor ratio in percentage. In other words, when the elasticity is less than unity, the increase in the capital-to-labour ratio results in a proportionally larger decrease in the relative price of capital-to-labour, leading to a larger decline in the capital share relative to the labour share.

---

<sup>24</sup>In a standard life-cycle permanent income model, consumption is proportional to permanent income because households can smooth their consumption by borrowing or lending their resources. Therefore, there is no heterogeneity in the MPC among households. However, in an incomplete market model, such as Aiyagari (1994), where households may face liquidity constraints and income uncertainty, the MPC is heterogeneous across households. In particular, the liquidity-constrained households have a higher MPC because they are hand-to-mouth and cannot smooth their consumption.

<sup>25</sup>Since households supply their labour inelastically, labour input ( $N_t$ ) is not changing. The price of labour,  $W_t$ , declines slightly from 1.0 to around 0.98 when the markup increases by 14 percentage points and  $\sigma$  is calibrated 0.95.

Table 2.3 illustrates the results of the model that correspond to the stylized facts of the US economy. The results show that the model can generate the four stylized facts in the economy only from the considered increase in the markup. Following 7 and 14 percentage-point increase in the markup, the model generates macroeconomic patterns that are similar to those in the data. However, the magnitude of the two stylized facts in the model is distinctly different from that in the observed data. While the changes of the labour share (**S1**) and the rate of return on capital (**S4**) in the model are closer to the changes observed in data, there are notable differences between my model and the data regarding the changes in the capital share (**S2**) and the wealth-output ratio (**S3**).

In the benchmark model, when I increase the markup by 7 percentage points, my model can explain 67% of the observed labour share change (**S1**) and 45% of the observed change in the rate of return on capital (**S4**). However, it can only explain 28% of the observed change in the wealth-output ratio (**S3**) and 22% of the observed change in the capital share (**S2**). When I increase the markup by 14 percentage points (from 6% to 20%) for the robustness check, my model explains 118% of the observed labour share change (**S1**), i.e. it overpredicts the decline in the labour share by around 18%. At the same time, while my model explains 70% of the observed change in the rate of return on capital (**S4**), it can only explain 46% of the observed change in the wealth-output ratio (**S3**) and 31% of the observed change in the capital share (**S2**).

The results indicate that a 7 percentage increase in markups underpredicts the observed decline in the labour share by 33%, but a 14 percentage increase in markups

overpredicts the observed decline in the labour share by 18%. This finding is consistent with the statement of [Basu \(2019\)](#) that a mild increase in markups is enough to explain the labour share decline in the US. [De Loecker and Eeckhout \(2017\)](#) estimate that markups have increased by around 50 percentage points since the early 1980s. Applying such an increase in markups results in a stock return that is too high to be compatible with stable wealth distribution in the model.

One plausible reason for the notable difference in the capital share decline is the possibility of mismeasurement of corporate capital in databases such as BEA that is discussed and reviewed by [Karabarbounis and Neiman \(2019\)](#). They explain that the estimate of capital stock is sensitive to mismeasurement from initial capital, depreciation rate, organizational and intangible capital. Similarly, [Gu and Macdonald \(2020\)](#) highlight the effects of including intangible capital in the estimates of gross fixed capital formation. Moreover, [Ewens, Peters and Wang \(2020\)](#) explain that the current accounting standards result in an underestimate of reported assets because the internally created intangible capital is not reported in firms' balance sheets.

In this paper, I study how much of the new stylized facts of the US economy can be explained by the rise of market power alone. I acknowledge that my model does not capture all factors that may be responsible for the stylized facts, such as automation ([Martinez \(2018\)](#) and [Acemoglu and Restrepo \(2018\)](#)), globalization ([Young and Tackett \(2018\)](#)), government policies such as corporate tax changes ([Kaymak and Schott \(2018\)](#)), increase in income inequality and income effective tax rates ([Saez and Zucman \(2016, 2020\)](#)).

Table 2.3: The stylized facts in data and in the model

Variable	Labour share (S1)	Capital share (S2)	Wealth/output (S3)	$R = (R^b + \delta)$ (S4)
Model ( $\mu = 1.06$ )	65%	29%	255%	13%
Model ( $\mu = 1.13$ )	61%	27%	285%	11%
% $\Delta$	-6%	-7%	+12%	-15%
Model ( $\mu = 1.06$ )	65%	29%	255%	13%
Model ( $\mu = 1.20$ )	58%	26%	307%	10%
% $\Delta$	-11%	-10%	+20%	-23%
Data (1982-1991)	64%	25%	210%	15%
Data (2007-2016)	58%	17%	300%	10%
% $\Delta$	-9%	-32%	+43%	-33%

## 2.5 Conclusion

To study the impacts of the rise of market power on the economy, it is not only the increase of profit but also its distribution among heterogeneous households that matter. When income inequality is high, the redistribution of profit among households is very uneven. This feature is due to heterogeneity of households in terms of income, and hence in terms of wealth and MPC. The rise of the profit share and its unequal distribution among heterogeneous households lead to a decline in the factor shares, an increase in the wealth-output ratio and a decrease in the rate of return on capital. When markup and, consequently, profit increase, the poor households cannot benefit from it because they cannot participate in the stock market actively. Moreover, wealthy households have a lower MPC and a higher return on their assets which leads to the above patterns that are known as the new stylized facts of the US economy over the last 35 years.

To alleviate the consequences of market power, the government should implement policies that promote more competition among firms, such as removing entry barriers and regulating merging and acquisitions (M&As). The market power may increase

if M&As are not regulated properly. For example, [Blonigen and Pierce \(2016\)](#) argue that M&As are associated with increases in markups or [Grullon, Larkin and Michaely \(2019\)](#) find a positive relationship between merging, profitability and market concentration.

This chapter shows that an increase in market power can aggravate wealth inequality among households. If this is a concern, an increase in the capital income tax is a way to counteract it. In the US, the capital income tax rate is lower than the labour income tax rate. In the literature, taxing capital income is very controversial.<sup>26</sup> However, some recent studies find that the optimal tax rate on capital income is positive and significant ([Straub and Werning \(2020\)](#) and [Scheuer and Slemrod \(2020\)](#)). Given the rise of wealth and income inequality, it is worth examining a higher capital income tax rate for the purpose of redistribution. In addition to taxing capital income, the focus can be on taxing the corporate profit and the high productive labour.<sup>27</sup> [Guo and Lansing \(1999\)](#) find that the optimal tax on the profit of monopolistic firms depends on the profit share and can be positive. Regarding the labour income tax, [Piketty and Saez \(2013\)](#) demonstrate that the marginal rate of optimal labour tax for top income percentile can be as high as 80%.

---

<sup>26</sup>For example, [Mirrlees \(1971\)](#), [Judd \(1985\)](#), and [Chamley \(1986\)](#) argue that no tax on capital income is optimal, while [Conesa, Kitao and Krueger \(2009\)](#), [Straub and Werning \(2020\)](#), and [Scheuer and Slemrod \(2020\)](#) argue for a positive optimal capital income tax. Recently, [Straub and Werning \(2020\)](#), dispute that [Judd \(1985\)](#)'s finding of optimality of zero capital income tax is not valid if the intertemporal elasticity of substitution (IES) is less than unity.

<sup>27</sup>Profit from the accounting point of view which is excluded the capital income of shareholders.

# Chapter 3

## The Estimates of Firm-Level

## Markups in Canada Using a Novel

## Data to Overcome Measurement

## Issues

### 3.1 Introduction

An increase in market power is associated with lower welfare and increased inefficiency.

Recent studies find that markups, as a measure of market power, have increased for some countries in the past decades (for the US, see [Barkai \(2017\)](#), [Hall \(2018\)](#) and [De Loecker, Eeckhout and Unger \(2020\)](#); for France, see [Hong \(2018\)](#), and for the global markups, see [Diez, Leigh and Tambunlertchai \(2018\)](#) and [De Loecker](#)

and Eeckhout (2018a)).<sup>1</sup> These studies also argue that increasing markups are one of the main drivers of some macroeconomic trends, such as labour share decline (see Eggertsson, Robbins and Wold (2018)). It is, therefore, important to know the dynamics of markups over time, as well as their estimates at a point in time. However, there are few papers that estimate markups in Canada. Also, those papers mainly use industry-level data and do not cover recent years of post Great-Recession.

Using firm-level data instead of industry-level data is a key advantage in estimating markups. On the one hand, using firm-level data allows for precise and rigorous estimates when the “production function” approach is applied (see Berry, Gaynor and Scott Morton (2019)). On the other hand, it enables researchers to capture the heterogeneity across firms by using firm-level data rather than industry-level data. De Loecker and Eeckhout (2018b) emphasize that the dispersion of markups across firms and their evolution over time are significantly different. Despite these advantages, the current literature of the firm-level markup estimation suffers from two problems. First, it does not use a reliable measure of a variable input cost (hereafter variable cost) for production functions. The firm-level markup using the “production function” approach can be derived as  $\mu = \beta_v \frac{\text{output}}{\text{variable cost}}$ , where  $\beta_v$  is the output elasticity of a variable input. The estimated markup using the “production function” approach depends highly on measuring the *variable cost* in the denominator of the markup definition. So, the main challenge is how to measure a variable cost from accounting data at the firm level. The literature considers the *cost of goods sold* (*COGS*) and the *operating expenses* as the measures of variable costs, although they

---

<sup>1</sup>Markup is defined as the ratio of price over marginal cost.

may not be reliable measures (see [Basu \(2019\)](#) and [Syverson \(2019\)](#)). Second, the literature focuses only on publicly-traded firms due to data availability constraints (see [De Loecker and Eeckhout \(2018a\)](#), [Traina \(2018\)](#), [Karabarbounis and Neiman \(2018\)](#), [Flynn, Gandhi and Traina \(2019\)](#), and [De Loecker, Eeckhout and Unger \(2020\)](#)). Excluding non-publicly-traded firms causes a sample selection bias, and therefore, the implications of macroeconomic trends using biased estimated markups can be misleading (see [Van Reenen \(2018\)](#)).

In this chapter, I contribute to the literature of the firm-level markup estimation by addressing these issues using the novel database of Canadian firms. Recently, Statistics Canada has made the T2-Longitudinal Employment Analysis Program (T2-LEAP) database available for researchers (see [Section 3.3](#) for more details about the T2-LEAP database). T2-LEAP has two notable advantages over databases that are widely used in the literature, such as Compustat. First, T2-LEAP provides a separate report for the wage bill of all firms. The wage bill is a more reliable measure of variable costs than some alternative measures used in the literature. Second, T2-LEAP contains the information of financial statements on the universe of firms, i.e. both publicly-traded and non-publicly-traded firms. Considering these advantages, I contribute to the literature in the following three dimensions.

First, I use a more reliable measure of the variable cost, i.e. wage bill or labour input cost. The convention in the literature is to measure variable costs either by the *COGS* ([De Loecker and Eeckhout \(2018a\)](#) and [De Loecker, Eeckhout and Unger \(2020\)](#)) or by the *operating expenses* ([Traina \(2018\)](#), [Karabarbounis and Neiman](#)

(2018) and Flynn, Gandhi and Traina (2019)).<sup>2</sup> The *COGS* undermeasures variable costs because it excludes some variable costs such as salaried workers (see Basu (2019)). On the other hand, the *operating expenses* overmeasures variable costs since it contains some semi-fixed costs, such as marketing costs. Therefore, using the *COGS* leads to overestimating the markups, while using the *operating expenses* leads to underestimating them. In addition, both *COGS* and *operating expenses* can suffer from the measurement error of capital depreciation (see Section 3.4). Using only the wage bill as a reliable measure of the variable cost alleviates these issues.

Second, I consider both publicly-traded and non-publicly-traded firms to estimate the average of firm-level markups. Some recent papers use only publicly-traded firms to explain the macroeconomic stylized facts (Barkai (2017) and De Loecker, Eeckhout and Unger (2020)). However, Van Reenen (2018) argues that generalization from publicly-traded firms can suffer from the sample selection bias, and thus, their macroeconomic implications may be misleading. The distribution of non-publicly-traded firms across industries is notably different from the distribution of publicly-traded ones. Including non-publicly-traded firms will remarkably change the represented shares of industries in the economy. The selection bias due to using only publicly-traded firms that usually have a larger size and potentially higher market power tends to overestimate the average of firm-level markups (see also Traina (2018)). Considering both publicly-traded and non-publicly-traded firms resolves the problem of sample selection bias and provides more precise average markups.

---

<sup>2</sup>*COGS* refers to the costs that are directly used in the production process, such as costs of labours, production managers and materials, while *operating expenses* contain the overall cost of sales which is the combination of *COGS* and the overhead costs. See Section 3.3 for more details.

Third, I am the first to estimate markups in the Canadian economy using the firm-level data, which allows me to capture the heterogeneity of firms and overcome the most important issues that have arisen in the previous literature. There are only few papers that estimate markups in Canada. They mostly use industry-level data and do not cover the post Great-Recession period.<sup>3</sup> One exception is the recent work of [De Loecker and Eeckhout \(2018a\)](#), where they estimate the global markups, including Canada. However, they use a very limited firm-level data and consider the *COGS* as the measure of variable costs.<sup>4</sup>

Figure [3.1a](#) displays the result of value-added markups for the baseline model, where I use the wage bill as the measure of variable costs. It indicates that the average of value-added markups for Canadian firms increased from 24% in 2002 to 29% in 2015. Three points are noticeable about this result. First, from 2002 to 2015, the value-added firm-level markups increased mildly by five percentage points. Second, during the recession from 2007 to 2009, markups stayed roughly constant at 27%. Third, the overall average of the value-added markups over the study period was around 26%.

Acknowledging that the value-added markups overestimate the gross output markups ([Hall, Blanchard and Hubbard \(1986\)](#), [Domowitz, Hubbard and Petersen \(1986\)](#), [Shapiro \(1987\)](#), [Hall \(1988\)](#), [Rotemberg and Woodford \(1993\)](#), [Basu \(1996\)](#), [Basu and Fernald \(1997\)](#), [Basu and Fernald \(2002\)](#), and [Basu \(2019\)](#)), I calculate the alternative gross output markups (hereafter, gross markups) from the estimated value-added

---

<sup>3</sup>For example, see [Martins, Scarpetta and Pilat \(1996\)](#), [Leung \(2008\)](#), and [Khan and Kim \(2013\)](#).

<sup>4</sup>Their dataset only has a few hundred firms for Canada. In comparison, T2-LEAP provides financial statements of around one million Canadian firms.

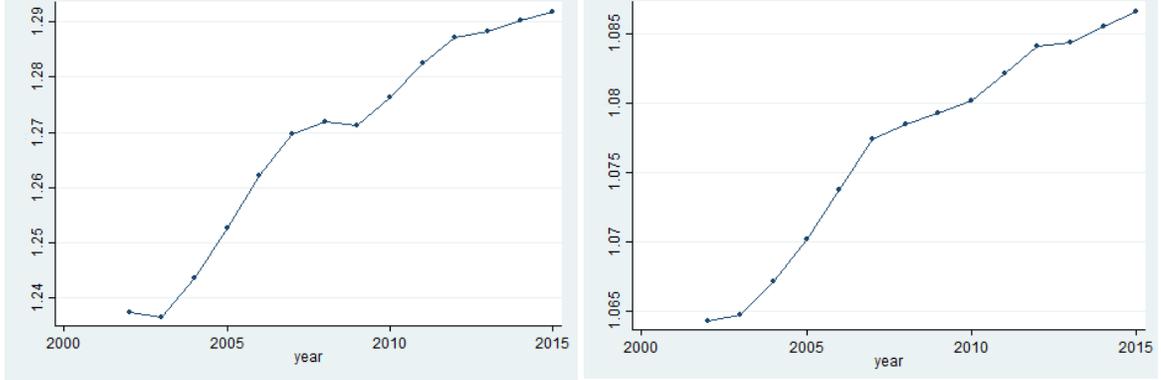
markups. Figure 3.1b displays the evolution of gross markups using the wage bill for the baseline model. The gross markups increased from 6% in 2002 to 9% in 2015. The results are distinctively lower than the finding of De Loecker and Eeckhout (2018a) for Canada. They consider the *COGS* as the measure of variable costs and find that the gross markups in Canada increased from a negative value in the early 1980s to more than 50% in recent years.

To compare the benchmark with the literature, I also study the following specifications. First, to see the impacts of using alternative measures of variable costs, I estimate the gross markups using the *operating expenses* and the *COGS* as the measures of variable costs.<sup>5</sup> Second, to see the impacts of the sampling bias due to excluding non-publicly-traded firms, I estimate the gross markups for only publicly-traded firms using the Compustat dataset. The result shows that when using the *operating expenses* as the measure of variable costs, the estimated gross markups for both categories of firms are relatively close to those of the benchmark. These findings confirm the argument of Basu (2019) that “it is safer to use a more comprehensive input measure”(p. 18). In other words, the *operating expenses* as the measure of variable costs would be a better choice than the *COGS* when the wage bill is not available.

The Compustat dataset for the US reports the wage bill for a subset of firms that contains around 20% of total firms. Using the limited firm-level data on the wage bill, I estimate firm-level markups in the US and get the same results those as in

---

<sup>5</sup>For the case of *COGS*, I have to rely on Compustat because T2-LEAP does not report the *COGS*.



(a) The average value-added markups

(b) The average gross markups

Figure 3.1: The average markups using both publicly-traded and non-publicly-traded firms in Canada.

Notes: The figure displays the estimates of markups using the wage bill as the measure of the variable cost. Panel (a), which is scaled from 1.24 to 1.29, shows the evolution of average value-added markups and Panel (b), which is scaled from 1.06 to 1.09, exhibits the evolution of gross markups. The average markups are weighted by the value-added output of firms.

Canada (see Appendix B.1). I find that when using the wage bill as the measure of the variable cost, the average markups for publicly-traded firms in the US increased from 8% in the early 1980s to 17% in 2018. For the same number of firms, the average markups increased from 4% (10%) in the early 1980s to 10% (40%) in 2018 when the *operating expenses* (*COGS*) is used as the measure of variable costs.

The rest of this chapter is organized as follows. Section 3.2 describes the methodology of the baseline model, where the wage bill is considered as the measure of the variable cost. Section 3.3 introduces the T2-LEAP database and describes its advantages over some other databases used in the literature. In Section 3.4, I explain the challenges of finding a reliable measure of variable cost from the financial statements of firms. Section 3.5 discusses the results of this chapter that includes five subsections. Subsection 3.5.1 illustrates the evolution of value-added and gross markups using the wage bill. In Subsection 3.5.2, I change the measures of variable costs from

the wage bill to the *operating expenses* and the *COGS*, and compare their results with the benchmark markups. In Subsection 3.5.3, I estimate the average markups for only publicly-traded firms and compare their results with the benchmark ones which are based on both publicly-traded and non-publicly-traded firms. Subsection 3.5.4 illustrates the dispersion of the average markups, and the measures of industry-level markups. Subsection 3.5.5 discusses the discrepancies between the micro-data (firm-level) markups and the macro-data markups that are obtained from the labour share of income. Finally, Section 3.6 presents the conclusion of this chapter.

## 3.2 Methodology

In terms of methodology, I closely follow the work of De Loecker and Warzynski (2012), which is based on the spirit of Hall (1988). Particularly, in each period, firm  $i$  of industry  $j$ , minimizes its costs

$$\min_{L_{it}, K_{it}} P_{it}^l L_{it} + P_{it}^k K_{it} \quad (3.2.1)$$

subject to

$$Q_{it}(L_{it}, K_{it}, \Omega_{it}) = \Omega_{it} L_{it}^{\beta_l} K_{it}^{\beta_k} \quad (3.2.2)$$

where,  $P_{it}^l L_{it}$ ,  $P_{it}^k K_{it}$ ,  $Q_{it}(\cdot)$ , and  $\Omega_{it}$  are the cost of the variable input (labour), the cost of the fixed input (capital), output technology, and Hicks-neutral productivity, respectively. By rearranging the F.O.Cs of the optimization problem, one can obtain the following equation for the markup:

$$\mu_{it} = \beta_l \frac{P_{it}Q_{it}}{P_{it}^l L_{it}}, \quad (3.2.3)$$

where,  $P_{it}Q_{it}$  and  $P_{it}^l L_{it}$  are the value-added output and the cost of the variable input (labour), respectively. Both of them can be obtained from the T2-LEAP database. The output elasticity of the variable input,  $\beta_l$ , however, needs to be estimated. Even though firms in the same industry have different productivity, they have access to similar technology. Therefore, I assume that firms inside each industry have the same output elasticity of input ( $\beta_l$ ) that is not firm-specific and not varying over time. Although the assumption of constant elasticity over time and across firms within the same industry is typically used in the literature, it can be a strong assumption.

In order to obtain the output elasticity, following the work of [Akerberg, Caves and Frazer \(2015\)](#), I estimate the following value-added production function for each industry:

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}, \quad (3.2.4)$$

where  $q_{it}$ ,  $l_{it}$ ,  $k_{it}$ , and  $\omega_{it}$  are the logs of deflated firm-level value-added, variable labour input, capital, and productivity, respectively; and  $\epsilon_{it}$  is an *i.i.d.* measurement error.<sup>6</sup> Since the above equation is in log-log form, the estimated coefficients will be interpreted as elasticity. In particular, the estimated coefficient of the variable input,  $\hat{\beta}_l$ , is the output elasticity of the variable input.

Productivity ( $\omega_{it}$ ) captures factors such as managerial ability, expected defect

---

<sup>6</sup>It should be noted that the quantity of variable inputs (labour) is not available at the firm level; so the literature considers the variable input cost (wage bill) as a proxy for the measure of the variable input to estimate the production function. In this chapter, both the measure of the variable input cost and the measure of variable input refer to the same value.

rates, or soil quality, which are unobservable by econometricians but observable by firms when they take input decisions. Therefore,  $\omega_{it}$  is likely to be correlated with inputs, so omitting it will result in biased estimates of the elasticities.

To tackle this problem, similar to [De Loecker and Warzynski \(2012\)](#), I rely on the two-stage estimation method suggested by [Olley and Pakes \(1992\)](#) and [Levinsohn and Petrin \(2003\)](#) and developed by [Akerberg, Caves and Frazer \(2015\)](#).<sup>7</sup>

For the problem of unobservable productivity, I rely on an intermediate input bundle to proxy for productivity. In particular, I assume that the firm's intermediate demand is given by

$$m_{it} = f_t(k_{it}, l_{it}, \omega_{it}). \quad (3.2.5)$$

By the assumption of monotonicity, productivity can be defined as follows:

$$w_{it} = f_t^{-1}(k_{it}, l_{it}, m_{it}). \quad (3.2.6)$$

Given the function of productivity, one can estimate Equation 3.2.4 using a two-stage process. In the first stage, I estimate the following specification:

$$q_{it} = \phi_{it}(k_{it}, l_{it}, m_{it}) + \epsilon_{it}, \quad (3.2.7)$$

---

<sup>7</sup>[De Loecker and Eeckhout \(2017\)](#), [De Loecker, Eeckhout and Unger \(2020\)](#), [De Loecker and Eeckhout \(2018a\)](#), [Traina \(2018\)](#) and [Karabarbounis and Neiman \(2018\)](#) also refer to a modified version of this method.

where the estimated value-added output ( $\hat{\phi}$ ) is given by

$$\phi_{it} = \beta_l l_{it} + \beta_k k_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it}).^8 \quad (3.2.8)$$

The functional form of  $f_t^{-1}(\cdot)$  is unknown, and therefore, I rely on a non-parametric regression (higher-order polynomial) to obtain  $\hat{\phi}$  in Equation 3.2.7. Then, productivity can be obtained for any value of  $\boldsymbol{\beta} = (\beta_l, \beta_k)$  as follows:

$$\omega_{it}(\boldsymbol{\beta}) = \widehat{\phi}_{it} - \beta_l l_{it} - \beta_k k_{it}. \quad (3.2.9)$$

In the second stage, I assume that productivity follows an AR(1) process,  $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$ , where  $\rho$  is the AR parameter and  $\xi_{it}$  is an innovation term. The key idea behind constructing the moment conditions is that the shock to the productivity should be orthogonal to the input decisions made in the previous period ( $k_{it}$  and  $l_{i,t-1}$ ). In other words, the key parameters that determine the value-added elasticity with respect to labour and capital are estimated to satisfy:

$$E\left(\xi_{it}(\boldsymbol{\beta}) \begin{pmatrix} k_{it} \\ l_{i,t-1} \end{pmatrix}\right) = 0, \quad (3.2.10)$$

where  $\boldsymbol{\beta} = (\beta_l, \beta_k)$ . The standard errors of the estimates can be calculated based on block bootstrapping.

---

<sup>8</sup>It should be noted that the intermediate input ( $m_{it}$ ) is not an explicit argument in Equation 3.2.4. The value of the intermediate inputs simply transfers to the value of the gross output and hence is not part of value-added. However, I use the intermediate input as a proxy for productivity in  $f_t^{-1}(\cdot)$  function of Equation 3.2.8, based on the idea that firms' demand for the intermediate input is higher when productivity is higher (see [Akerberg, Caves and Frazer \(2015\)](#)).

Finally, after the estimation of the output elasticity of labour,  $\widehat{\beta}_l$ , I measure firm-level markups using the following:

$$\mu_{it} = \widehat{\beta}_l \frac{S_{it}}{C_{it}}, \quad (3.2.11)$$

where,  $S_{it}$  and  $C_{it}$  are the value-added output of firms and the cost of labour, respectively.<sup>9</sup> Though the elasticity ( $\beta_l$ ) is identical for all firms within each industry, the level of markup ( $\mu_{it}$ ) varies within each industry based on the ratio of value-added over the variable costs of firms.

### 3.3 The T2-LEAP database

Statistics Canada has recently made the T2-Longitudinal Employment Analysis Program (T2-LEAP) database available for researchers through Research Data Centres (RDCs) as a pilot project. T2-LEAP is a combination of two datasets: The Corporate Income Tax File (T2) dataset and the Longitudinal Employment Analysis Program (LEAP) dataset. The T2 files are used by firms when they file their tax return with the Canada Revenue Agency (CRA) and thus, the T2 files contain income statements for all incorporated firms in Canada. The T2 files must be filled in based on the standard codes of the General Index of Financial Information (GIFI). The variables of T2 dataset come with a GIFI code and researchers can refer to the code to see what is captured in each variable. The T2 files of T2-LEAP contain the following variables:

---

<sup>9</sup>To adjust for the revenue's measurement errors, I deduct the  $\exp(\epsilon_{it})$  from  $S_{it}$ , where the measurement error ( $\epsilon_{it}$ ) can be obtained from Equation 3.2.8. For more details, see [De Loecker and Warzynski \(2012\)](#).

assets, total expenses, gross profit, net income, revenue, shareholder's equity, tangible assets (PPE), total liabilities, and working capital.<sup>10</sup> The LEAP dataset contains employment information for the universe of firms in Canada. The information in LEAP is obtained from the annual statements of remuneration paid (T4 slips) that Canadian firms issue to their employees for tax purposes. The number of employees, wage bill (payroll), first year, and last year of firms are the variables that are reported in the LEAP dataset. T2-LEAP classifies firms into industries based on the two-digit level of the North American Industry Classification System (NAICS) code, excluding the public sectors of education, health and public administration which are classified as NAICS 61, 62, and 91, respectively. The frequency of data is annual, covering from 2001 to 2015. T2-LEAP has two notable advantages over databases that are widely used in the literature, such as Compustat. First, T2-LEAP provides a separate report for the wage bill of all firms. Second, T2-LEAP contains the information of financial statements on the universe of firms, i.e. both publicly-traded and non-publicly-traded firms. One disadvantage, however, is that the sample period in T2-LEAP is considerably shorter than that in Compustat.

Regarding the required variables for the model, *revenue*, *total expenses*, *tangible capital*, wage bill and the *depreciation* of capital are directly reported in T2-LEAP. To obtain the *value-added output* and *intermediate input*, I follow the work of Keller and Yeaple (2009). I deduct the adjusted wage bill and the *depreciation* of capital from the *total expenses* of firms to obtain the value of *intermediate input*. I subtract the

---

<sup>10</sup>For more information about the variables included in T2-LEAP and their definitions, see the website of the [Canadian Research Data Centre Network](#).

calculated *intermediate input* from the *revenue* of firms to generate the *value-added* of the firms. All of these variables are nominal, so I deflate them using the consumer price index reported by Statistics Canada (Table: 18-10-0005-01, 2002=100).

I categorize the firms in 16 sectors (2-digits NAICS industry level), excluding the public and finance sectors. I remove the finance sector (NAICS 52-53) because their balance sheets are different from those of the rest of the firms, and a standard Cobb-Douglas function may not be a good representative for the production function of financial firms. I drop any observation with missing or negative values of revenues, tangible assets and total expenses. To eliminate the distortions due to merger and acquisition, I remove any observation if the growth of assets or sales is more than +100% or less than -100%. To mitigate the effect of outliers, I winsorize the top and bottom 2% of the variables used in the regression (i.e., revenue, wage bill, and capital) and the top and bottom 5% of the value-added over variable cost ratio.<sup>11</sup> I also take out the firms with less than five employees. I acknowledge that very small firms are likely to have smaller markups and removing them can be a potential source of biased average markups. I decide to remove them because assuming the same production function for large and very small firms is unrealistic; and the labour input might not be that “variable” for the very small firms, especially if they are family-owned. Furthermore, the contribution of the very small firms to GDP is very limited. Given that the small firms with less than 100 employees only contribute to 12% of the Canadian GDP, the contribution of very small firms with less than five employees will

---

<sup>11</sup>The outliers are more prevalent in the ratio variable (the value-added over variable cost) than the other regressors that are in levels. So, I consider a more strict cutoff for the former compared to the latter. The overall result is also robust to using the 2% cutoff for the ratio variable.

be significantly smaller (see [Innovation, Science and Economic Development Canada \(2019\)](#)). Therefore, while removing the very small firms alleviates potential identification problems, it may not have serious impacts on the average markups that are weighted based on the sales of firms. The above cleaning process reduces the number of observations from around 17 million to around 1.8 million. Most of this reduction is due to removed firms with less than five employees and missing observations.

To construct the measure of capital, I apply the perpetual inventory method. To generate the initial level of capital, I drop the accumulated depreciation from the firms' first-year stock of capital and add the flow of depreciation for that specific year. After the first year, capital accumulates as  $k_{it} = k_{it-1} + i_{it} - \delta k_{it-1}$ . The reported amount of capital in the balance sheet is equal to the stock of capital and the current investment. I deduct the last year's accumulated depreciation from the stock of capital to obtain the current level of capital,  $k_{it}$ . Moreover, in constructing the labour cost, I considered 15% of the wage bill as the employee benefits such as pension and insurance which are not reported in T4 slips.<sup>12</sup>

---

<sup>12</sup>According to the survey of the Conference Board of Canada that is reported by [Stewart \(2015\)](#), the total cost of employee benefits for mid- to large-sized firms is on average \$8,330 per full-time equivalent. This makes around 18% of the average income of Canadians in 2015, which is \$46600 according to Statistics Canada (Table:11 – 10 – 0239 – 01). In this chapter, I consider the benefit ratio as 15% because my model also contains the small-sized firms which usually have lower benefit ratio. Moreover, some firms may also report their cash bonus as employee benefits, which is already reported in the T4 slips as wage bill. My overall result is robust for any employee benefit ratio between 12% and 20%. At the aggregate level, the total employee benefits are between 14% to 16% of the total wages and salaries over the study period (Statistics Canada, Table: 36-10-0221-01). The distribution of employee benefits is not available at the firm level. I acknowledge that there might be some variation of employee benefits across firms and industries, which might impact the analysis.

### 3.4 The challenges in measuring a variable input

As discussed earlier, the estimated markup using the “production function” approach depends highly on the defined measure of the variable cost in the denominator of the markup definition. A common challenge in the literature on firm-level markups is to correctly measure the variable costs using information from financial statements of firms. In particular, one needs to derive measures of the variable and the fixed costs of production from the financial statements. The categories used in the financial statements, however, do not necessarily correspond to the inputs in the production function. The inputs in the financial statements are not defined as fixed vs. variable costs; instead, they are categorized as direct and indirect costs.

In general, the total costs of production and sales of firms will be recorded in the financial statements as *operating expenses*, which is divided into two categories: the *cost of goods sold (COGS)* and the *selling, general and administrative (SG&A)* expenses. Theoretically, if expenses have been used directly in the production process, such as materials or salaries of technicians in the production line, they should be recorded as *COGS* or direct costs. Other expenses that are used for selling and administrative purposes should be recorded as *SG&A* or indirect costs. For instance, the purchased price of office supplies for a marketing department or the salaries of salespersons should be recorded as *SG&A* expenses. See Table 3.1 for more details.

However, if the firm purchases a fixed capital, such as a building, its depreciation can be recorded in both forms of costs, depending on the purpose of purchase. In the balance sheet of firms, the purchased building will be recorded in the plant, property

and equipment (*PPE*) account and the annual depreciation of this capitalized asset will be reported in the financial statement of the firm either as *COGS* or *SG&A* expenses, depending on the purpose of the purchase. If the building is used for the production department, its depreciation will be recorded as the *COGS*, while it will be part of the *SG&A* if the building is used for the marketing and administrative department. Purchases of intangible capital, such as software, will be treated in the same way. The amortization of intangible capital will be recorded either as *COGS* or *operating expenses*.<sup>13</sup>

Table 3.1: A sample income statement

<b>Revenues</b>		<b>XXX</b>
<b>Operating Exp.</b>		
	<b>COGS</b>	
	direct labour salaries	XXX
	direct materials	XXX
	depreciation of direct capital	XXX
	amortization of direct intangible capital	XXX
	<b>SG&amp;A expenses</b>	
	salespersons salaries	XXX
	admin staff salaries	XXX
	marketing expenses	XXX
	depreciation of sales' building	XXX
	amortization of patent, goodwill, etc.	XXX
		<b>(XXX)</b>
<b>Net income</b>		<b>XXX</b>

Notes: The *COGS* undervalues the variable costs because it excludes some part of salaried workers. The *operating expenses*, on the other hand, overvalues the variable costs because it includes some fixed costs such as marketing expenses. Moreover, both the *COGS* and *operating expenses* can suffer from the measurement error of capital depreciation and amortization.

The literature is split in terms of what should be used as a measure of variable costs. One strand uses *COGS* as the measure of variable costs, while the other defines *operating expenses* ( $COGS + SG\&A$ ) as the measure of variable costs. [De Loecker](#)

<sup>13</sup>In accounting terminology, “amortization” is used to refer to the depreciation in intangible capital.

and Eeckhout (2017, 2018a) use *COGS* as the measure of the variable costs of production and find that markups increased dramatically for the US. On the other hand, Flynn, Gandhi and Traina (2019), Karabarbounis and Neiman (2018) and Traina (2018) use the *operating expenses* as the measure of the variable costs and find that markups increased only slightly over time for the US. They all use the firm-level Compustat data.

Traina (2018) and Karabarbounis and Neiman (2018) criticize De Loecker and Eeckhout (2017) for the use of *COGS* as the measure of variable costs. They claim that the true measure of variable costs is the *operating expenses* and explain that the missing of *SG&A* expenses understates the variable costs and causes a dramatic increase in markups. In response to their criticism, De Loecker, Eeckhout and Unger (2020) argue that *SG&A* expenses cannot be a variable cost. They add *SG&A* expenses as a fixed input in their production function and still find that markups increased dramatically for the US.

However, the bundle of *SG&A* expenses is neither purely fixed nor purely variable cost but a combination of both. For example, payments to salespersons or admin staff (in *SG&A*) can be even more variable than the salaries of production line technicians (in *COGS*). On the other hand, marketing costs (in *SG&A*) are more likely to be fixed than a variable cost. Hence, considering the *SG&A* either as purely fixed costs or purely variable costs results in a measurement error.

Basu (2019) advocates the use of a more comprehensive measure of variable costs such as *operating expenses* ( $SG\&A + COGS$ ) instead of *COGS*. He highlights that “the underlying theory does not require that all of the input on the examined margin

be variable. It requires only that there be some variable inputs in the input bundle under consideration, and that the bundle be defined consistently over time” (p. 18). Further, he indicates that the possibility of measurement errors is higher when the *COGS* is used as a measure of variable costs. For example, he says that payment to salaried workers is classified as *SG&A*, while payment to hourly workers is recorded as *COGS* by convention, no matter if they work in the production line or administrative and selling departments.

Although *operating expenses* may have some advantages over *COGS*, both of them will suffer from the measurement error caused by including accounting depreciation and amortization of capital. If depreciation of fixed inputs is captured in the measure of variable costs, the depreciation is double-counted among the regressors, once as a part of variable costs and again as a part of fixed costs. Furthermore, depreciation calculation based on accounting rules will be affected by firms’ accounting systems and government’s tax policies, and hence, they are different from the economic depreciation and amortization. Using the cost of labour, or simply the wage bill, allows me to avoid both issues.

## 3.5 Results

### 3.5.1 The evolution of firm-level markups

#### The value-added markups

Figure 3.1a depicts the evolution of the value-added markups when the wage bill is considered as the variable cost of production. The average of the value-added markups is calculated with the value added of firms as the weight. The average markups increased mildly by five percentage points from 24% in 2002 to 29% in 2015. The markups stayed roughly constant at 27% over the recent financial crisis (2007–2009) and gradually increased after 2009.

As shown in Equation 3.2.11, the average markups at any point in time will be determined by both the value-added elasticity of the variable input ( $\beta_l$ ) and the ratio of value added over the variable costs ( $\frac{S_{it}}{C_{it}}$ ). The value-added elasticity of wage bill, on average, is around 0.96 and it is constant over time.<sup>14</sup> The averages of elasticities change only slightly when I consider different measures of variable inputs, and hence, the estimated elasticities cannot explain the differences in the markups of various specifications (for the measures of elasticities, see Appendix B.2). So, the estimated markups and their evolution over time are mainly driven by the ratio of value added (or revenue) over the variable costs. For this reason, it is vital to use a reliable measure of variable costs that can generate a true ratio of value-added (or revenue)

---

<sup>14</sup>De Loecker and Eeckhout (2017) specify a model with constant elasticity of variable inputs, but De Loecker, Eeckhout and Unger (2020) update the paper to time-varying elasticity by estimating production functions over every 5-year of data. They claim that the elasticity of variable inputs declines over time. However, Bond, Hashemi, Kaplan and Zoch (2020) indicate that defining a very short period of time in panel data may result in a biased estimate of the elasticity.

over the variable costs. To highlight the importance of the variable cost measurement, I discuss the consequences of changing the measure of variable costs from the wage bill to the *operating expenses* and the *COGS* in Subsections 3.5.2 and 3.5.2. The last two measures are widely used in the literature and there are significant discrepancies in their estimated markups. The estimated markups in the literature are based on the gross markups, so I compute the gross markups and compare the results to the literature.

### **The gross markups**

The literature of firm-level markups is mainly focused on gross markups. To compare my results with the literature, I compute gross markups based on the estimated value-added markups. The value-added output is defined as the gross output net of intermediate inputs. In the process of the production function estimation, intermediate inputs are used as a proxy for the unobservable productivity (see Equation 3.2.5). Naturally, I choose the value-added output instead of the gross output to estimate the production function.

Hall, Blanchard and Hubbard (1986) state that the exclusion of intermediate inputs from gross output may lead to an overestimate of the markups (see Domowitz, Hubbard and Petersen (1986), Shapiro (1987), Rotemberg and Woodford (1993), Basu (1996), Basu and Fernald (1997), and Basu and Fernald (2002)).<sup>15</sup> To compute gross markups, I follow the above literature and assume that gross production function

---

<sup>15</sup>Basu (2019) also explains how “double-marginalization” may cause the problem of overestimation. The final firms charge a markup on the intermediate inputs while the intermediate firms have already priced those materials with a markup.

is Leontief in intermediate inputs, i.e. the production function is proportional to intermediate inputs as the following:

$$Q_t^g = \min\left\{\frac{M_t}{\gamma_t}, \frac{F(K_t, L_t)}{1 - \gamma_t}\right\}, \quad (3.5.1)$$

where  $Q_t^g$  is gross output,  $M_t$  is the amount of intermediate inputs,  $F(\cdot)$  is the value-added production function with labour ( $L_t$ ) and capital ( $K_t$ ) as the inputs, and  $\gamma_t$  is the weight of intermediate input in one unit of output. Given the definition of the marginal cost of production in gross and value-added output, the following relationship can be obtained between the gross markups and the estimated value-added markups.<sup>16</sup>

$$\mu^* = \frac{\mu}{1 + (\mu - 1)S^m} \quad (3.5.2)$$

where  $S^m$  is the ratio of intermediate input costs over total revenue, and  $\mu^*$  is the gross markup that determines the price of a product over its marginal cost. Computing gross markups using Equation 3.5.2 is based on the assumption that firms are price-taker in intermediate input markets. [Hall, Blanchard and Hubbard \(1986\)](#) discuss that the calculated  $\mu^*$  may underestimate the gross markups when the markets of intermediate inputs are not competitive.

Figure 3.2a illustrates the measure of gross markups using Equation 3.5.2. The markups increased by about three percentage points from 6% in 2002 to 9% in 2015. The average gross markups are considerably lower than the finding of [De Loecker and Eeckhout \(2018a\)](#) for Canada. They use the *COGS* (instead of the wage bill) as

---

<sup>16</sup>For complete derivation steps, see [Khan and Kim \(2013\)](#).

the measure of variable costs and find the gross markups increased from a negative value in the early 1980s to more than 50% in recent years.<sup>17</sup> As stated earlier, the discrepancies in the literature are mainly due to different measures of variable costs rather than differences in the estimated elasticities. The estimated elasticities change only slightly when I consider various specifications (see Appendix B.2). Therefore, the estimated markups and their evolution over time are mainly driven by the measure of variable costs in the ratio of output over variable costs. For this reason, it is crucial to use a reliable measure of variable costs in estimates of firm-level markups using the “single-input” approach. To highlight the importance of the variable cost measurement, I discuss the consequences of changing the measure of variable costs from the wage bill to the *operating expenses* and the *COGS* in the next subsections.

De Loecker, Eeckhout and Unger (2020) use the wage bill from the US Compustat dataset to estimate markups as a robustness check. The US Compustat dataset provides the wage bill for roughly 20% of publicly-traded firms. They claim that using the wage bill, similar to the *COGS*, will lead to substantially high gross markups for the US. In particular, they find that markups increased by around 35 percentage points. Having examined their work, I find that they undermeasure the intermediate inputs and overmeasure the value-added outputs. When addressing the issues, I find that the average markups for the US increased only by five percentage points from 10% in the early 1980s to 15% in 2018. For more details on the estimating method

---

<sup>17</sup>Comparing to the literature of the industry-level data, my result is close to the implied markups obtained by Khan and Kim (2013) and smaller than the estimated markups by Leung (2008) and Martins, Scarpetta and Pilat (1996). Martins, Scarpetta and Pilat (1996) find the average markups of around 25%. Leung (2008) estimates that the business sector markups fluctuate around 30% between the 1980s to 2003. Khan and Kim (2013) find much lower (implied) markups. They obtain that the gross output markups increased from 2% to 6% from the mid-1990s to 2009.

and the results, see Appendix B.1 and Figures B.1 and B.2.

### 3.5.2 Changing the measure of variable inputs

To compare the gross benchmark results with the literature, I re-estimate the gross markups using the more conventional measure of variable costs, i.e. the *operating expenses* and the *COGS*.<sup>18</sup> In terms of methodology, I closely follow the baseline model of De Loecker, Eeckhout and Unger (2020). In particular, I estimate the following gross production function:

$$q_{it}^g = \beta_v v_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}, \quad (3.5.3)$$

where,  $q_{it}^g$ ,  $v_{it}$ ,  $k_{it}$ , and  $\omega_{it}$  are the logs of deflated gross output (revenue), variable inputs, capital, and productivity, respectively. The variable inputs are measured either by the *COGS* or the *operating expenses*. The productivity,  $\omega_{it}$ , is defined as a function of capital and a control variable (either the *COGS* or the *operating expenses* in this case). The gross markup is obtained directly from the cost minimization of the firms.

$$\mu_{it}^* = \hat{\beta}_v \frac{S_{it}^*}{C_{it}^*}, \quad (3.5.4)$$

where  $\mu_{it}^*$ ,  $S_{it}^*$ ,  $C_{it}^*$  are gross markup, revenue, and variable costs, respectively.  $\hat{\beta}_v$  is the estimated output elasticity of variable input.

The average markups using the *operating expenses* increased roughly by two per-

---

<sup>18</sup>For the case of the *COGS*, I use Compustat data since T2-LEAP does not report this variable. Compustat only reports the financial statements of publicly-traded firms.

centage points from around 5% in 2002 to around 7% in 2015 (see Figure 3.2b). The estimated average markups are calculated using the revenue of firms and industries as the weight. The estimated markups are close to the gross markups with the wage bill as the measure of the variable cost (see Figure 3.2a). Similarly for the US, Traina (2018), Karabarbounis and Neiman (2018), and Flynn, Gandhi and Traina (2019) use the *operating expenses* as the measure of variable costs and find a very low level and small slope for the evolution of markups. The average of their estimated markups increased by only five percentage points from about 10% in the early 1980s to about 15% in 2018.

T2-LEAP does not report the *COGS*, so I rely on Compustat and estimate the evolution of markups for the Canadian publicly-traded firms using the *COGS* as the measure of variable costs. I find an incredibly higher value for markups using this new measure of variable costs. As Figure 3.3b illustrates, markups increased from around 32% in 2002 to around 36% in 2015. This result is in line with the finding of De Loecker and Eeckhout (2018a) for Canada, where they also use the *COGS* and find markups increased from a negative value in 1980 to around 50% in recent years. In sum, using the *COGS* results in much higher levels and steeper slopes for the evolution of estimated markups.

The difference between the estimated markups using the wage bill, the *operating expenses*, and the *COGS* can be explained by Equations 3.5.4 and 3.2.11. The estimated markups are determined by both the elasticity of variable inputs  $\beta_v$  ( $\beta_l$ ) and the ratio of revenue (value-added) over the variable costs. The elasticities of variable inputs under various specifications are not too different to explain the discrepancies

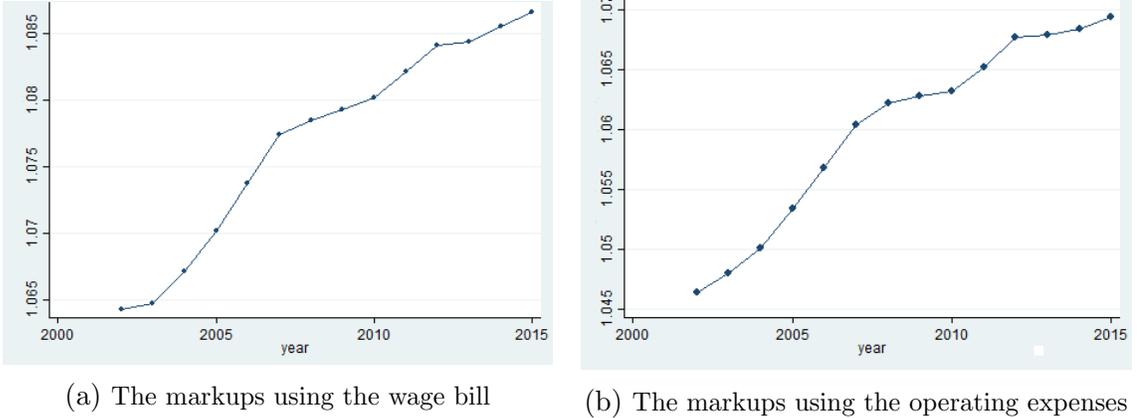


Figure 3.2: The average gross markups, considering both publicly-traded and non-publicly traded firms (source: T2-LEAP)

Notes: The figure shows the estimates of gross markups for the universe of Canadian firms using two different measures of variable costs. Panel (a), which is scaled from 1.06 to 1.09, displays the average gross markups using the wage bill (benchmark model) and Panel (b), which is scaled from 1.04 to 1.07, exhibits the average gross markups using the operating expenses. The average markups are weighted by the value-added output and the revenue of firms.

in estimated markups (see Tables B.1 and B.2 in Appendix B.2). Therefore, the discrepancies in estimated markups are mainly coming from the ratio of revenue over variable costs. As explained in Section 3.4, compared to the wage bill, the *COGS* undervalues the variable costs and so overestimates the markups. On the other hand, the *operating expenses* overvalues the variable costs and hence, slightly underestimates the gross markups. Basu (2019) explains that, theoretically, estimating markups using the “production function” approach does not suffer from an inclusive measure of variable costs such as *operating expenses*. In contrast, it suffers from an exclusive measure of variable costs such as the *COGS*. This can explain as to why the markups using the *operating expense* are similar to the markups when the wage bill is used.

The above results also hold when I estimate firm-level markups for the US. Figure B.2 in Appendix B.1 demonstrates that the US markups using the *operating expenses* and the wage bill are very close to each other, especially after the 1980s. The similarity

between the results of the *operating expenses* and the wage bill confirms [Basu \(2019\)](#)'s argument that it is safer to use a more inclusive bundle. In other words, in the absence of the wage bill, the *operating expenses* is a better measure of variable costs than the *COGS*.

### 3.5.3 Excluding non-publicly-traded firms

Some recent papers such as [Barkai \(2017\)](#), [De Loecker and Eeckhout \(2018a\)](#) and [De Loecker, Eeckhout and Unger \(2020\)](#) use the dynamics of the markups estimated from data for only publicly-traded firms to explain the macroeconomics stylized facts, including the labour share decline. Markups estimated from only publicly-traded firms may not be a good proxy for the aggregate markups of the economy. The distribution of non-publicly-traded firms across industries is notably different from the distribution of public ones. Adding non-publicly-traded firms will remarkably change the weight of each industry in the economy. For instance, when only publicly-traded firms are considered, the industries of Mining, Oil and Gas are over-represented, while Construction, Wholesale Trade, Retail Trade, and Accommodation and Food Services are under-represented. On this account, the average markups, which are weighted based on the sales of firms, will be biased when non-publicly-traded firms are excluded. [Van Reenen \(2018\)](#) also discusses that generalization from publicly-traded firms can suffer from the sample selection bias, and thus, the findings related to the macroeconomic stylized facts can be misleading.

To assess how misleading the outcomes would be when non-publicly-traded firms

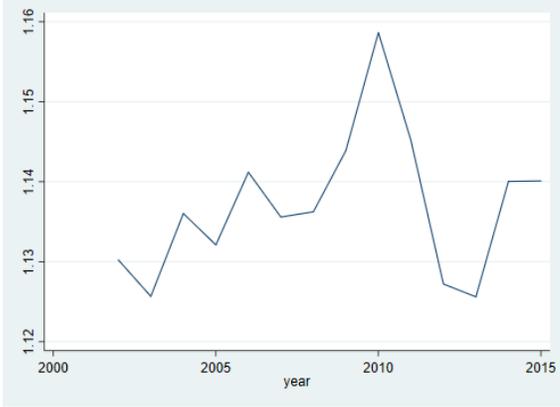
are excluded, I estimate average markups from only publicly-traded firms. To do so, I rely on Compustat since T2-LEAP does not report whether the firms are publicly traded or not. Canadian Compustat does not have enough data on the wage bill, so I can only compare the markups using the *operating expenses* and the *COGS* as the measures of variable costs.

Figure 3.3 illustrates the average markups for publicly-traded firms. The markups, when using *operating expenses* as the measure of the variable costs, increased from 13% in 2002 to 14% in 2015. The estimated markups are higher than the results obtained from T2-LEAP using the *operating expenses*. The higher markups can be because of the selection bias due to using only publicly-traded firms that usually have a larger size and potentially higher market power (see Traina (2018)). When considering the *COGS* as the measure of variable costs, markups for publicly-traded firms increased from 32% in 2002 to 40% in 2015.<sup>19</sup> This finding is in line with the literature that *COGS* results in much higher estimated markups, and it also confirms the Basu (2019)'s argument that using the *operating expenses* as the measure of variable costs would be a better choice than the *COGS* when the wage bill is not available.

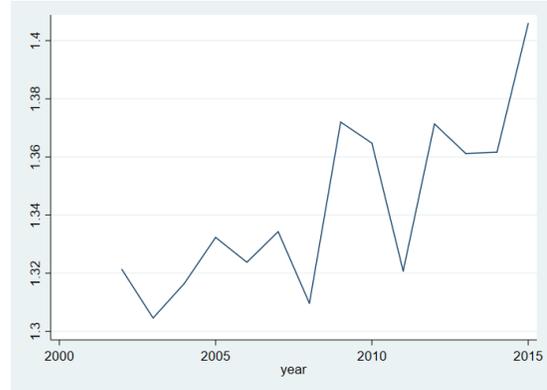
In comparing results from the universe of firms (T2-LEAP) and those from the publicly-traded firms (Compustat), one should bear in mind that these outcomes are obtained from two different data sources where neither of them is perfect. For instance, the *operating expenses* in Compustat is net of the depreciation of tangible

---

<sup>19</sup>Using the *COGS*, I cannot compare the estimated markups of only publicly-traded firms with the results of the universe of firms because T2-LEAP does not report the *COGS*.



(a) The markups using the operating expenses



(b) The markups using the COGS

Figure 3.3: The average gross markups for publicly-traded firms (source: Compustat)

Notes: The figure displays the evolution of average gross markups for publicly-traded firms using the operating expenses and the COGS as the measures of variable costs. Panel (a), which is scaled from 1.12 to 1.16, exhibits the average markups using the operating expenses and Figure (b), which is scaled from 1.30 to 1.40, shows the average markups using the COGS. The average markups are weighted by the revenue of firms.

capital and partially net of amortization of intangible capital, while T2-LEAP only reports the depreciation of tangible capital. By excluding the amortization of the intangible capital from the *operating expenses* in T2-LEAP, the amount of the net *operating expenses* would decline, and the resulted markups would slightly increase (see Equation 3.2.11). On the other hand, a key challenge in using Compustat is the missing observations, particularly for the *COGS*. Almost one-third of the firms use the value of *operating expenses* to report the *COGS* in their financial statements. After removing those firms along with other missing observations, I lose a considerable amount of data.<sup>20</sup> The missing observations make the results of publicly-traded firms unreliable, especially the markups estimated based on the *COGS*.

<sup>20</sup>The Compustat dataset has 57K observations where 23% of the observations, mainly non-manufacturing sectors, do not report the true *COGS* and around 20% of the observations are missing. I restore around 1% of missing ones by interpolation. For the case of *operating expenses*, only 19% of the observations are missing.

### 3.5.4 The dispersion of markups and the industry-level markups

#### The dispersion of value-added markups

Having the firm-level markups allows for tracking the distribution of markups over time. To study the dispersion of markups across firms, I sort the firms according to their markups and find the 50, 75, and 90 percentiles of markups.<sup>21</sup> To obtain the percentiles, I weigh each firm by its share of value added. The result is exhibited in Figure 3.4. The value-added markups show a remarkably large distribution across firms. The median of markups is slightly below the average and increased by five percentage points from around 22% to 27% over the study period. In comparison, markups for the 75th percentiles increased by six percentage points, from 32% to 38%, and those for the 90 percentiles increased by 10 percentage points from 45% to 55%. The dispersion of markups indicates that first, there is significant heterogeneity among firms at a point in time, and second, the tendency over time is almost similar across the distribution.

#### The industry-level markups

The average of the value-added markups at the industry level is illustrated in Appendix B.3. The average of markups in each industry is calculated using the value-added of the firms as the weight. The industries are categorized based on the two-digit level North American Industry Classification System (NAICS) codes, i.e. at the sec-

---

<sup>21</sup>It should be noted that over time firms may change their positions across the distribution of markups. The distributions are simply calculated period by period, and I do not assume any persistence in firms' positions in the distribution. The fat right tail in recent years can be driven by the new superstar firms that might not even have existed in the past.

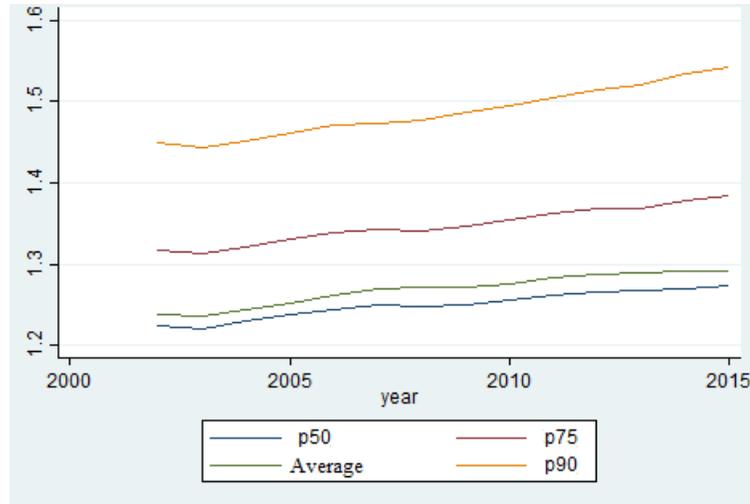


Figure 3.4: The evolution of the distribution (percentiles) of markups

Notes: In this figure, p50, p75, p90, and Average refer to the median, 75 percentile, 90 percentile, and the weighted average of markups, respectively. To study the dispersion of markups across firms, I sort the firms according to their markups and find the 50, 75, and 90 percentiles of markups. The sorted firms are weighted by their share of value added.

tor level. I briefly explain the evolution of markups for the following four major sectors: the Construction sector (NAICS 23), the Manufacturing sector (NAICS 31, 32, and 33), the Wholesale Trade (NAICS 42), and the Retail Trade (NAICS 43). The value-added markups for the Construction sector increased from around 24% in 2002 and reached 30% in 2007 and stayed constant for a couple of years. The markups started to decline from 2012 to 29% in 2015. For the Manufacturing sectors of NAICS 31, 32, and 33, the markups gradually increased from 35%, 32% and 25% in 2002 to 48%, 39%, and 35% in 2015, respectively. For Wholesale Trade, the markups steadily increased from 32% to 45% during the study period. Finally, the markups for Retail Trade increased from 18% in 2002 to 22% in 2015. The level of the value-added markups stayed almost constant for these sectors during the financial crisis and then increased except for the Construction sector, whose markups slightly declined even

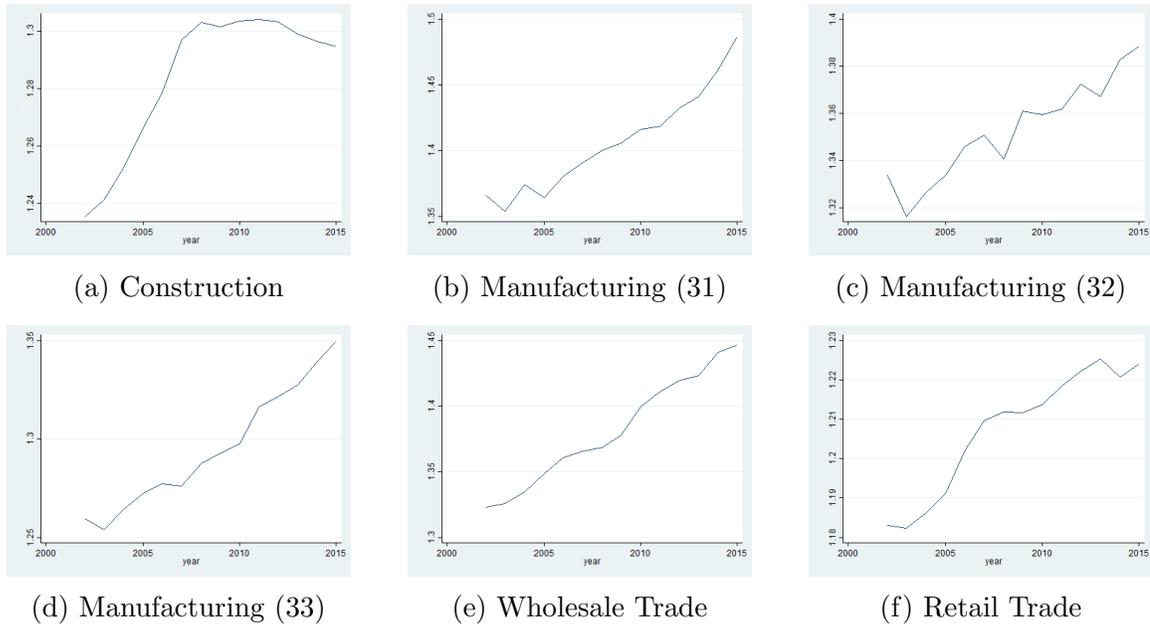


Figure 3.5: The average of the value-added markups for the major sectors.

after the crisis was over. The tightened regulations implemented by the government to prevent another bubble in the housing market can be one of the reasons for the different trend of markups in the Construction sector. Some examples of tightened regulations include banning subprime mortgages, reducing the amortization period, and introducing stress tests (see [Curry \(2016\)](#) and [Mishkin and Serletis \(2018\)](#)).

I also compute the weighted average of the gross markups at the industry level in Canada. The computed markups are smaller than and roughly parallel to the value-added markups for most of the industries. For more details, see [Appendix B.4](#).

### 3.5.5 Micro-data vs. macro-data markups

In the macroeconomic literature, it is common to use the inverse of the aggregate labour share of GDP as a proxy for aggregate markups. The aggregate labour share

is defined as the share of value added that paid out to the workers.<sup>22</sup> Given the labour share, the aggregate markup is defined as  $\mu_t = \frac{\alpha_l}{S_l}$  where,  $S_l$  and  $\alpha_l$  are the labour share of GDP and the elasticity of labour, respectively (for more details, see Rotemberg and Woodford (1991), Rotemberg and Woodford (1999), Khan and Kim (2013), Nekarda and Ramey (2013), and Nekarda and Ramey (2020)). To obtain the aggregate value-added markups, I use the aggregate labour share of GDP from the Federal Reserve Economic Data (FRED) and, I follow the work of Rotemberg and Woodford (1999) to calibrate the elasticity of labour as  $\alpha_l = 0.70$ . To compare the result of markups with the literature, I transform the aggregate value-added markups into the aggregate gross markups using Equation 3.5.2. To do the transformation, I consider the average share of intermediate input in gross output,  $S^m$ , equal to 0.5 (see Statistics Canada, Table: 36-10-0217-01). Figures 3.6 and 3.7 illustrate the evolution of the labour share and aggregate gross markups for Canada.<sup>23</sup> The result indicates that the average of aggregate gross markups is around 4% over the study period. The average of aggregate markups is slightly smaller than the overall average of gross markups that I estimate at the firm level. Using wage bill as the measure of the variable cost, I find that the overall average gross markups is around 7%. However, the average of gross aggregate markups is remarkably smaller than the overall average

---

<sup>22</sup>Mathematically, the aggregate labour share,  $S_l$ , is equal to  $S_l = \frac{W_t L_t}{Y_t} = \left(\frac{1}{\mu_t}\right) \left(\frac{W_t L_t}{W_t L_t + R_t K_t}\right)$  where  $W_t$  is wage rate,  $L_t$  is total employment,  $R_t$  is the rental rate of capital,  $K_t$  is the value of capital,  $Y_t$  is total GDP, and  $\mu_t$  is markup.

<sup>23</sup>Contrary to the literature arguing for a decrease in the labour share, especially in the US, the aggregate labour share of GDP in Canada slightly increased from around 65% in 2002 to around 66% in 2015. Consequently, the gross aggregate markups have a small downward slope. It should be noted that when we consider a longer period of time, the labour share demonstrates a decreasing trend. Khan and Kim (2013) find a significant negative correlation between markups and the labour share by studying a longer period of time from 1980 to 2009.

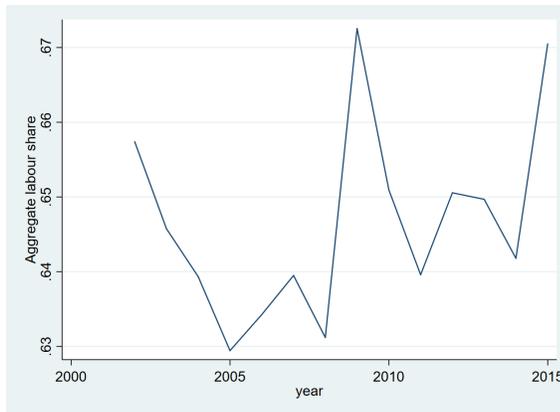


Figure 3.6: The aggregate labour share of GDP for Canada

of gross markups obtained by [De Loecker and Eeckhout \(2018a\)](#) using the *COGS* as a measure of variable costs. According to their estimates, the overall average of gross markups is more than 40% over the same study period, from 2002 to 2015.

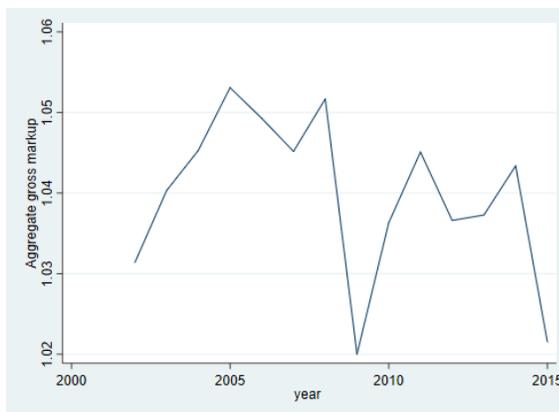


Figure 3.7: The evolution of aggregate markups for Canada

Notes: The figure illustrates the evolution of the gross aggregate markups for Canada. The aggregate gross markups are obtained based on the aggregate value-added using Equation 3.5.2. The aggregate value-added markups are defined as the inverse of the aggregate labour share of GDP multiplied by the elasticity of labour. Data on the labour share of GDP is obtained from the FRED database, and the elasticity of labour is calibrated as  $\alpha_l = 0.70$ .

## 3.6 Conclusion

An increase in market power is associated with lower welfare and higher inefficiency. Therefore, it is important to know the dynamics of markups over time, as well as their estimates at a point in time.

Although the literature of firm-level markups has some advantages, it presents some challenges too. The sample selection bias due to using only publicly-traded firms and the measurement errors in the commonly used measures of variable costs are the two main challenges. In this chapter, I address these challenges by using the novel and rich database of T2-LEAP. In particular, I consider both publicly-traded and non-publicly-traded firms to address the selection bias of using only publicly-traded firms. I also define the wage bill as the measure of the variable cost in the production function. The wage bill is a clearer measure of variable cost than the

*COGS* and the *operating expenses* that are widely used in the literature. In general, the *COGS* undervalues the variable costs, while the *operating expenses* overvalues them. Also, both of them may suffer from the depreciation and amortizations of fixed inputs, i.e. the stocks of tangible and intangible capital. As a result, I find that from 2002 to 2015, the average of the gross markups increased mildly from 6% to 9%. My results are considerably lower than the findings of [De Loecker and Eeckhout \(2018a\)](#) for Canada. They use the *COGS* as a measure of variable costs, and claim that the gross markups in Canada increased from a negative value in the early 1980s to more than 50% in recent years.

Moreover, I show that contrary to the *COGS*, the markups using the *operating expenses* are not too different from the markups using the wage bill. In other words, the *operating expenses* can be a more reliable measure of the variable costs than the *COGS* when the wage bill data is not available. I also argue that excluding non-publicly-traded firms may result in overestimates of markups, and hence its macroeconomic implications can be misleading.

Finally, similar to the macroeconomic literature, I define the aggregate markup as the inverse of labour share and obtain the average of aggregate gross markup equal to 4%. The average of aggregate markups is slightly smaller than the overall average of gross markups at the firm level. Using the wage bill as the measure of the variable cost, I find that the overall average of gross markups is around 7%. However, the average of gross aggregate markups is remarkably smaller than the overall average of gross markups that is obtained by [De Loecker and Eeckhout \(2018a\)](#) using the *COGS* as the measure of variable costs. According to their estimates, the overall average of

gross markups is more than 40% over the same study period, from 2002 to 2015.

# Chapter 4

## Firm-Level Markups and Accounting Data: A New Approach

### 4.1 Introduction

Recently, a wave of “production function” approach has been used to estimate markups using accounting data at the firm level. In this approach, the markup ( $\mu$ ) is defined as the product of two terms, the output elasticity of a variable input and the ratio between revenue and the cost of that variable input (hereafter variable cost). Since markups depend on the measure of a single variable input, this approach is also called a “single-input” approach ([Basu \(2019\)](#)).

A common challenge in estimating markups using “single-input” approach is to accurately measure the variable costs from accounting data. An inappropriate mea-

sure of a variable cost not only results in a biased estimate of output elasticity of variable input but also causes an inaccurate ratio of sales (revenues) over variable costs. Variable costs at the firm level are not directly available and researchers need to define some measures of variable costs from accounting data and income statements of firms. Accounting systems categorize the expenses of firms as direct and indirect costs. These costs do not necessarily correspond to the costs of variable and fixed inputs that are required for estimating markups in the “production function” approach. To measure the cost of variable inputs, researchers use some assumptions that affect the results of estimated markups significantly.

Total production expenses of firms are recorded in income statements as the *operating expenses*, which is divided into two categories: the *cost of goods sold (COGS)* and the *selling, general and administrative (SG&A)* expenses. Theoretically, if expenses are made during the production process, such as materials or salaries of technicians, they should be recorded as *COGS* (i.e. direct costs). Other expenses, such as salaries of human resource staff or advertising costs that are used for selling and administrative purposes should be recorded as *SG&A* (i.e. indirect costs). The convention in the literature is to measure variable costs either by the *COGS* (De Loecker and Eeckhout (2018a), Diez, Leigh and Tambunlertchai (2018) and De Loecker, Eeckhout and Unger (2020)) or by the *operating expenses* (Traina (2018), Karabarbounis and Neiman (2018), and Flynn, Gandhi and Traina (2019)). Using the *operating expenses* as a proxy for variable costs results in a mild increase of markups from around 10% in 1980 to around 15% in recent years, while using the *COGS* leads to a very large increase in markups from around 20% in 1980 to more than 60% in recent years.

The large markups, estimated with the *COGS* as the measure of variable costs, have been questioned by some researchers. The studies cast doubt either on the reliability of the *COGS* as a measure of variable costs or consistency of large markups with the observed patterns in some key aggregate variables. [Basu \(2019\)](#) argues that firms in practice may not include all direct costs in the *COGS*. He states that firms usually exclude salaried workers from the *COGS*. [Traina \(2018\)](#) explains that *COGS* undermeasures the variable costs. Moreover, depending on the accounting system and the nature of production, some firms may not report *COGS*. For example, in the Compustat database, almost 25% of firms do not report the *COGS*, and Compustat reports the *operating expenses* for the measure of missing *COGS* in those firms. The studies also question the consistency of large markups with the observed patterns in some key aggregate variables such as profit rates, returns to scale, labour share decline, and productivity. For example, using aggregate data, [Ahmad, Fernald and Khan \(2019\)](#) estimate that the US economy exhibits a constant returns to scale and state that the low observed profit rate is not consistent with large markups. [Basu \(2019\)](#) also indicates that, in the US, a small increase in markups is more consistent with the observed profit rate and labour share decline. [Syverson \(2019\)](#) indicates that a large increase in markups should be followed by either a significant increase in prices or a decrease in unit costs of productions, but neither of them happened. [De Loecker, Eeckhout and Unger \(2020\)](#), however, claim that the heterogeneity of firms plays an important role in the measure of average firm-level markups, and they dispute any results that rely on aggregate data.

In this chapter, I introduce a refined “production function” approach to estimate

markups. In this approach, I estimate markups by combining the two methodologies that are applied in the literature of estimating markups, i.e. the methodology that uses aggregate data and the one that uses firm-level data. In this approach, the markup is defined as a function of economic profit rates and returns to scale of firms. I attain the economic profit rates from income statements and balance sheets of firms, and I obtain returns to scale by estimating the production function of firms using accounting data. I call this method a “total-input” approach because I use total inputs of firms from accounting data instead of relying on a single choice of variable input of production. Therefore, the result is less sensitive to the choice of measure of variable inputs.

Figure 4.1 illustrates the result of the average markups using the new approach. The average markups increased by six percentage points, from 7% in 1980 to 13% in 2018. Over the same period, production technology across industries exhibits constant returns to scale. Therefore, the economic profit rates of firms are the only drivers of time-varying markups. The average economic profit rates increased by five percentage points, from 4% in 1980 to 9% in 2018. The mild increase in the markups estimated in this approach suggests that the discrepancy between the findings in the literature of firm vs. aggregate data are more likely due to the use of *COGS* as the measure of variable costs, instead of heterogeneities across firms

The rest of this chapter is organized as follows. Section 4.2 introduces the methodology of the refined “production function” approach, i.e. the “total input” approach. Section 4.3 describes how to generate firm-level economic profit from accounting data and financial statements of firms, and introduces the data sources. Section 4.4 demon-

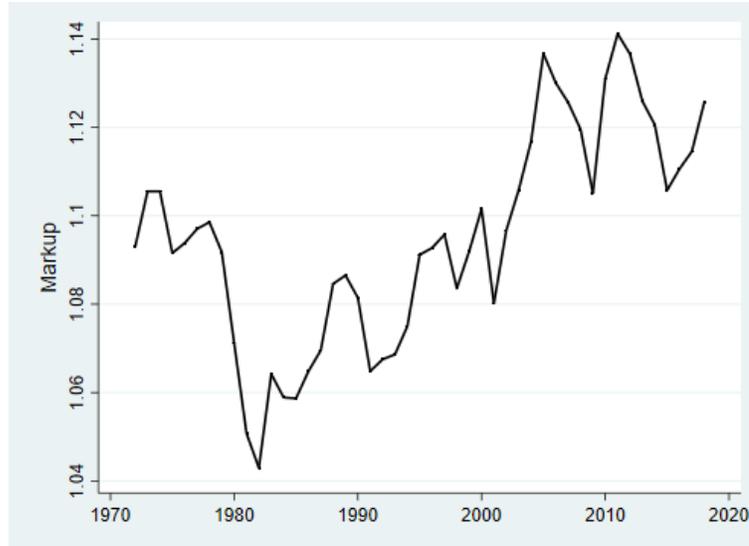


Figure 4.1: The average markups for the US

Notes: The figure displays the evolution of average markups using the “total-input” approach for publicly-traded firms in the US . In estimates of markups, I use the total inputs of firms, i.e. capital (PPE), COGS, and SG&A, from accounting data. The markups are weighted by the sales of firms.

strates the results of this chapter. In Section 4.5, I compare the refined approach with the existing method of firm-level markup estimation. Section 4.6 concludes.

## 4.2 Methodology

To estimate markups, I combine the two methodologies that are applied in the literature of estimating markups using aggregate data and firm-level data. First, I obtain an equation for the markup based on the methodology that is used in estimating markups using aggregate data, and then, I estimate the relevant production function of the defined markup based on the methodology that is used for firm-level data.

To derive an equation for markup, similar to [Ahmad, Fernald and Khan \(2019\)](#), I follow the work of [Basu and Fernald \(1994\)](#), which is based on the spirit of [Hall, Blanchard and Hubbard \(1986\)](#) and [Hall \(1988\)](#). In particular, at time  $t$ , firm  $i$  of

each industry minimizes its costs as follows:

$$\min_{V_{it}, X_{it}, K_{it}} P_{it}^v V_{it} + P_{it}^x X_{it} + P_{it}^k K_{it} \quad (4.2.1)$$

subject to

$$Y_{it} = F_{it}(V_{it}, X_{it}, K_{it}, \Omega_{it}), \quad (4.2.2)$$

where,  $P_{it}^k K_{it}$ ,  $P_{it}^v V_{it}$ , and  $P_{it}^x X_{it}$  are the costs of production inputs. While  $P_{it}^k K_{it}$  is the cost of fixed input (capital),  $P_{it}^v V_{it}$  and  $P_{it}^x X_{it}$  can be either the costs of variable, semi-variable, or fixed inputs.  $Y_{it}$ ,  $F_{it}(\cdot)$ , and  $\Omega_{it}$  are the total output, production technology, and Hicks-neutral productivity, respectively. The Lagrangian function and its first-order conditions are

$$L(V_{it}, X_{it}, K_{it}, \lambda_{it}) = P_{it}^v V_{it} + P_{it}^x X_{it} + P_{it}^k K_{it} + \lambda_{it}[Y_{it} - F_{it}(\cdot)], \quad (4.2.3)$$

$$\frac{\partial F_{it}(\cdot)}{\partial V_{it}} = \frac{P_{it}^v}{\lambda_{it}} \implies \frac{\partial F_{it}(\cdot)}{\partial V_{it}} \frac{V_{it}}{Y_{it}} = \frac{P_{it}^v}{\lambda_{it}} \frac{V_{it}}{Y_{it}}, \quad (4.2.4)$$

$$\frac{\partial F_{it}(\cdot)}{\partial X_{it}} = \frac{P_{it}^x}{\lambda_{it}} \implies \frac{\partial F_{it}(\cdot)}{\partial X_{it}} \frac{X_{it}}{Y_{it}} = \frac{P_{it}^x}{\lambda_{it}} \frac{X_{it}}{Y_{it}}, \quad (4.2.5)$$

$$\frac{\partial F_{it}(\cdot)}{\partial K_{it}} = \frac{P_{it}^k}{\lambda_{it}} \implies \frac{\partial F_{it}(\cdot)}{\partial K_{it}} \frac{K_{it}}{Y_{it}} = \frac{P_{it}^k}{\lambda_{it}} \frac{K_{it}}{Y_{it}}. \quad (4.2.6)$$

The left-hand side of Equations 4.2.4, 4.2.5, and 4.2.6 are equal to the output elasticity

of inputs ( $\beta_v$ ,  $\beta_x$ , and  $\beta_k$ , respectively). On the right-hand side, the parameter  $\lambda_{it}$  determines the marginal cost of production at a given level of output. The markup can then be defined as  $\mu_{it} = \frac{P_{it}}{\lambda_{it}}$ , where  $P_{it}$  is the unit price of an output. Therefore, the equations can be written as the following:

$$\beta_v = \mu_{it} \frac{P_{it}^v V_{it}}{P_{it} Y_{it}}, \quad (4.2.7)$$

$$\beta_x = \mu_{it} \frac{P_{it}^x X_{it}}{P_{it} Y_{it}}, \quad (4.2.8)$$

$$\beta_k = \mu_{it} \frac{P_{it}^k K_{it}}{P_{it} Y_{it}}. \quad (4.2.9)$$

By summing up Equations 4.2.7, 4.2.8, and 4.2.9, one can get the following relationships:

$$\gamma_{it} = \mu_{it}(s_{v,it} + s_{x,it} + s_{k,it}), \quad (4.2.10)$$

where  $s_{j,it}$  ( $j = v, x, k$ ) and  $\gamma = \beta_v + \beta_x + \beta_k$  represent the share of input costs in revenue and the returns to scale of a firm, respectively. The economic profit is defined as total revenue (or total sales) minus the cost of three inputs of production, i.e.  $\Pi = TR - (P^v V + P^x X + P^k K)$ . The economic profit share can be obtained as  $\pi_{it} = 1 - (s_{v,it} + s_{x,it} + s_{k,it})$  when dividing both sides by total revenue. Therefore, one can rewrite Equation 4.2.10 as

$$\mu_{it} = \gamma \left( \frac{1}{1 - \pi_{it}} \right), \quad (4.2.11)$$

where,  $\pi_{it}$  is the share of economic profit in revenue of firm  $i$  at period  $t$ . Given Equation 4.2.11, to estimate markup ( $\mu_{it}$ ), I generate the share of economic profit ( $\pi_{it}$ ) from accounting data and financial statements of firms and obtain the returns to scale of firms ( $\gamma$ ) by estimating the production function of firms in each industry.

To estimate the firm-level production function, I assume firms have a Cobb-Douglas production technology as follows:

$$q_{it} = \beta_v v_{it} + \beta_x x_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}, \quad (4.2.12)$$

where,  $v_{it}$ ,  $x_{it}$ ,  $k_{it}$  are the logs of deflated firm-level inputs obtained from accounting data and income statements of a firm, i.e. variable inputs, semi-variable inputs, and fixed inputs, respectively. In particular, I define *COGS* as the measure of variable inputs<sup>1</sup>, *SG&A* as the measure of semi-variable inputs, and *PPE* as the measure of fixed inputs. Also,  $q_{it}$  and  $\omega_{it}$  are the logs of deflated firm-level gross output and productivity, respectively, and  $\epsilon_{it}$  is an *i.i.d.* measurement error. Since the above equation is in log-log form, the estimated coefficients are interpreted as the output elasticity of inputs, and hence, the returns to scale of the production is  $\gamma = \beta_v + \beta_x + \beta_k$ .

Following the literature, I assume that productivity follows an AR(1) process, i.e.

$\omega_{it} = \rho\omega_{it-1} + \xi_{it}$ , where  $\xi_{it}$  is an *i.i.d.* shock.

In estimating Equation 4.2.12, one may face two challenges of unobservable pro-

---

<sup>1</sup>The quantity of variable inputs is usually not available at the firm level; so the literature considers the cost of variable inputs as a proxy for the measure of variable inputs. Although the *COGS* is a of bundle of variable inputs, it does not represent all variable inputs of the production function since a portion of variable inputs is reported in the *SG&A*. To estimate the production function in Equation 4.2.12, one does not need to generate an inclusive bundle of variable inputs.

ductivity ( $\omega_{it}$ ), and a potential correlation of a realized productivity shock with both non-capital inputs ( $v_{it}$  and  $x_{it}$ ). To overcome these challenges, I rely on the works of [De Loecker and Warzynski \(2012\)](#), [De Loecker and Eeckhout \(2017\)](#) and [De Loecker, Eeckhout and Unger \(2020\)](#), which are based on the methodologies developed by [Oley and Pakes \(1992\)](#), [Levinsohn and Petrin \(2003\)](#), and [Akerberg, Caves and Frazer \(2015\)](#). Similar to [De Loecker and Eeckhout \(2017\)](#) and [De Loecker, Eeckhout and Unger \(2020\)](#), the unobservable productivity is given by a function of inputs and a control variable, i.e.  $\omega_{it} = h(v_{it}, x_{it}, k_{it})$ . To estimate the elasticities, I apply a two-stage approach. In the first stage, I calculate the predicted output and measurement errors as follows:

$$q_{it} = \phi_{it}(k_{it}, v_{it}, x_{it}) + \epsilon_{it}, \quad (4.2.13)$$

where the estimated sales ( $\hat{\phi}$ ) is given by

$$\phi_{it} = \beta_v v_{it} + \beta_x x_{it} + \beta_k k_{it} + h(v_{it}, x_{it}, k_{it}) \quad (4.2.14)$$

The functional form of productivity  $h(v_{it}, x_{it}, k_{it})$  is unknown, and therefore, I rely on a non-parametric regression to obtain  $\hat{\phi}$  in Equation 4.2.14. For the non-parametric regression, following the work of [De Loecker and Warzynski \(2012\)](#), which is suggested by [Akerberg, Caves and Frazer \(2015\)](#), I use a third-order polynomial in the three variables of  $v_{it}$ ,  $x_{it}$ , and  $k_{it}$ . In the second stage, the implied productivity can be obtained for any value of  $\boldsymbol{\beta} = (\beta_v, \beta_x, \beta_k)$  as follows:

$$\omega_{it}(\boldsymbol{\beta}) = \hat{\phi}_{it} - \beta_v v_{it} - \beta_x x_{it} - \beta_k k_{it}. \quad (4.2.15)$$

I project the function of productivity onto its lag to retrieve the innovation to productivity. Using the generated series of innovations, which is a function of the vector of the elasticities ( $\boldsymbol{\beta} = [\beta_v, \beta_x, \beta_k]$ ), I construct the following moment conditions to impose orthogonality between productivity shocks and input decisions:

$$E \left( \xi_{it}(\boldsymbol{\beta}) \begin{pmatrix} v_{it-1} \\ x_{it-1} \\ k_{it} \end{pmatrix} \right) = 0. \quad (4.2.16)$$

Equation 4.2.16 implies that any decisions on the amount of non-capital inputs in the last period ( $v_{it-1}$  and  $x_{it-1}$ ) are independent of current period productivity shock ( $\xi_{it}(\boldsymbol{\beta})$ ). Similarly, the current capital ( $k_{it}$ ), which is determined in the last period, is orthogonal to productivity shock by construction. The estimates of the elasticities,  $\boldsymbol{\beta}$ , enable me to obtain the returns to scale of the production technology as  $\hat{\gamma} = \hat{\beta}_v + \hat{\beta}_x + \hat{\beta}_k$ . The standard errors of the estimates can be calculated based on block bootstrapping.

Finally, I plug the estimated returns to scale ( $\hat{\gamma}$ ) and the generated economic profit share ( $\pi_{it}$ ) back into Equation 4.2.11 to estimate time-varying firm-level markups. In Section 4.3, I explain the details of generating firm-level economic profit from accounting data and financial statements of firms.

## 4.3 Economic profits and accounting data

### 4.3.1 Economic profits

As explained in the previous section, the economic profit of each firm is defined as

$$\Pi = TR - (P^vV + P^xX + P^kK). \quad (4.3.1)$$

The total revenue ( $TR$ ) and the cost of non-capital inputs ( $P^vV$  and  $P^xX$ ) can be easily obtained from income statements of firms. However, calculating the realized cost of capital ( $P^kK$ ) is challenging because the accumulated capital is either invested directly by firm owners or borrowed from financial institutions and households. These two sources of accumulating capital need to be separated because the rate of the cost of capital, when borrowed from financial institutions and households, is different across firms. The rate depends on some factors, such as the credit constraint of firms that cannot be evaluated at the firm level.<sup>2</sup> Therefore, it can be difficult to obtain the economic profit using Equation 4.3.1. To overcome the challenges, I use the relationship between the accounting profit and the economic profit to obtain the latter.<sup>3</sup> One advantage of this approach is that the accounting profit is reported directly in financial statements of firms, and hence, I only need to calculate the opportunity cost of owners capital, which is equivalent to the required rate of return

---

<sup>2</sup>Cao and Leung (2020) show that the heterogeneity in financial constraints may also affect the estimates of productivity at the firm level. Although I address this issue in calculating profits, I acknowledge that this essay cannot take into account the heterogeneity in financial constraints in estimating the production function, by assuming a common production across firms in each industry.”

<sup>3</sup> Practically, this approach is a shortcut to obtain the economic profit; conceptually, it is the same as equation 4.3.1.

on the capital invested by firm owners.<sup>4</sup>

In accounting terminology, the financial data reported regularly by firms are called financial statements. The financial statements consist of a cash flow statement, an income statement, and a balance sheet. The cash flow statement, which focuses on the cash or equivalent liquid assets, is an indicator of available funds for operating expenses, investments, and debt obligations of firms. Income statements provide information on revenue and expenses of firms, and hence, one can obtain the accounting profits of firms from their income statements. The balance sheet includes the accounts of assets, liabilities, stockholders' equity (assets of owners) at a point in time. To calculate the economic profit, I use data from income statements and balance sheets of firms. I define the economic profit of a firm as follows:

$$\text{Economic profit} = \text{Accounting profit} - \text{Opportunity cost.} \quad (4.3.2)$$

The accounting profit is defined as total sales net of total expenses and can be obtained directly from income statements of firms:

$$\text{Accounting profit} = \text{Sales} - \underbrace{(\text{COGS} + \text{SG\&A} + \text{Depreciation and Amortization})}_{\text{EBIT}} - \text{Interest payments,} \quad (4.3.3)$$

where earnings before interest and taxes (EBIT) is directly reported by firms. I deduct the interest expenses of firms from the reported EBIT to calculate the accounting profit before taxes.

---

<sup>4</sup>Ahmad, Fernald and Khan (2019) estimate the profit rate based on an aggregate approach using weighted average of the required equity return and the required bond-market return to calculate the cost of capital.

To generate the economic profit using Equation 4.3.2, one needs to calculate the value of the opportunity cost for firm owners' assets that are not available in financial statements. I define the opportunity costs as risk-free returns on the invested assets (or capital) of stockholders in a firm. To obtain the assets of stockholders (firm owners), I use the following accounting equation from a firm's balance sheet:

$$\textit{Total assets of a firm} = \textit{Total liability} + \textit{Total assets of owners.} \quad (4.3.4)$$

According to Equation 4.3.4, the total assets of firms are equal to the book value of stocks (that are initially issued and recorded as the assets of owners) plus the liability that firms borrowed from lenders to finance some investment projects. In future, when people buy/sell stocks in a secondary market, the price of traded stocks will not have any impact on the book value of assets. Therefore, at any point in time, I obtain the total assets of owners<sup>5</sup> as:

$$\textit{Total assets of owners} = \textit{Total assets of a firm} - \textit{Total liability.} \quad (4.3.5)$$

So, the opportunity cost of owners' capital can be defined as the following:

$$\textit{Opportunity cost} = \textit{Risk-free rate} \times \textit{Total assets of owners,} \quad (4.3.6)$$

where the risk-free rate is defined as the 10-year US government bond yield. From the

---

<sup>5</sup>Since a portion of assets of owners are invested in short term securities by firms and not used in the production process, I deducted the cash and short term investments from the total assets. Cash and short term investments represent cash and all securities readily transferable to cash as listed in the Current Asset section.

calculated opportunity cost and accounting profit, I generate economic profits using Equation 4.3.2. I divide the calculated annual economic profit over the annual sales of a firm to attain the time-varying economic profit share ( $\pi_{it}$ ).

### 4.3.2 Data

The firm-level data for both estimating production function (Equation 4.2.12) and generating economic profit (Equation 4.3.2 to 4.3.6) are obtained from the Compustat dataset. To estimate the production function, I measure the output ( $y_{it}$ ), variable inputs ( $v_{it}$ ), semi-variable inputs ( $x_{it}$ ), and fixed input ( $k_{it}$ ) by the variables of *sale*, *COGS*, *SG&A*, and *PPE*, respectively. To generate economic profits, I obtain the variables of *EBIT*, *interest payments*<sup>6</sup>, *total asset*, and *liability* from Compustat as well. Data are annual, and the cleaned data cover from 1972 to 2018. To measure the rate of return on a risk-free asset in Equation 4.3.6, I use the 10-year US government bond yield from the Organization for Economic Co-operation and Development (OECD) database. I use the annualized yield, which is the average of monthly ones.

The firms are categorized into 16 sectors (2-digit NAICS code). I exclude the sectors of Utilities, Health Care and Education because these sectors are highly affected by government policies. I also remove the finance sector because a simple Cobb-Douglas production function cannot explain well the behaviour of financial firms. For

---

<sup>6</sup>In the baseline model, I use the *interest capitalized* (variable name: *intc*) as a measure of interest expenses. This item only considers interest charges incurred to the company's fixed and current assets amortized over the life of the purchased assets. In Equation 4.3.2, I assume that firms only apply for loans to purchase assets and capital and not to pay their operating expenses. I assume that firms can pay their operating expenses using their income. However, firms may have some incentives, such as corporate tax incentives, to finance their operating expenses from debt as well.

example, a common practice in the literature is that financial firms optimally choose their portfolios to maximize their profits given the constraints from the financial markets and government regulations (see [Sealey and Lindley \(1977\)](#), [Santomero \(1984\)](#), and [Ge and Weisbach \(2019\)](#)). However, the overall result is robust when I include the sectors of Finance and Utilities. To alleviate the problem of merging and acquisition, I exclude any observations when either the growth of their capital is more than 50% or the value of an acquisition is greater than 5% of the total assets of firms. To eliminate the impacts of outliers, I winsorize the top and bottom 3% of profit rates and the variables of the production function. To convert nominal variables to real ones, I deflate the variables of *sales*, *COGS*, *SG&A*, *total asset*, and *liability* using the GDP deflator reported by National Income and Product Accounts (line 1 in Table 1.1.9, 2012=100). I also deflate the capital stock (*PPE*) using the non-residential fixed investment deflator reported in line 9 of the same table.

## 4.4 Results

### 4.4.1 The evolution of markups

Figure 4.1 illustrates the weighted average of markups using the “total-input approach” for the US publicly-traded firms from 1972 to 2018. To obtain the average markups, each firm is weighted by its sales. The weighted average markups increased by six percentage points, from around 7% in 1980 to around 13% in 2018. I find a weighted average returns to scale of  $\gamma = 1.01$ , which represents an almost constant

returns to scale technology.<sup>7</sup> The weighted average of profit share for the US firms increased by five percentage points, from around 4% in 1980 to around 9% in 2018.

My finding is in line with the results of [Ahmad, Fernald and Khan \(2019\)](#) who use aggregate data to estimate a constant returns to scale technology. Their finding corresponds to an implied markup of 12% and an implied profit share of 11% in the post-1990 period.<sup>8</sup> Qualitatively, my results confirm the statements of [Basu \(2019\)](#) and [Syverson \(2019\)](#) that a smaller markup is more consistent with the observed small profit rates of the aggregate economy. My results are also close to the findings of [Traina \(2018\)](#), [Karabarbounis and Neiman \(2018\)](#), and [Flynn, Gandhi and Traina \(2019\)](#). They use the “single-input” approach and consider the *operating expenses* ( $COGS + SG\&A$ ) as a measure of variable inputs and estimate that markups increased by around five percentage points, from around 10% in the early 1980s to around 15% in 2016. In contrast, my results are substantially smaller than the findings of [De Loecker and Eeckhout \(2018a\)](#), [Diez, Leigh and Tambunlertchai \(2018\)](#), and [De Loecker, Eeckhout and Unger \(2020\)](#) where they use the  $COGS$  as a measure of variable inputs and find that markups are increasing from around 20% in the early 1980s to more than 60% in recent years. The  $COGS$  undermeasures variable inputs because it excludes some variable inputs such as payment to labours (see [Traina \(2018\)](#) and [Basu \(2019\)](#)), and the *operating expenses* overmeasures variable inputs because it includes some semi-fixed inputs such as advertising costs.

My results are less sensitive to the choice of conventional measures of variable

---

<sup>7</sup>The returns to scale for the main sectors of Manufacturing, Wholesale Trade, and Retail Trade are 1.01, 0.99, and 1.00, respectively.

<sup>8</sup>[Barkai \(2017\)](#) uses aggregate data and demonstrates that the share of profit in value added increased by 13.5 percentage points from approximately -5.6% in 1984 to 7.9% in 2014.

inputs, i.e. the *COGS* and the *operating expenses*. In the current literature, changing the measure of variable inputs from the *COGS* to the *operating expenses* results in significant discrepancy, while in my approach, both measures demonstrate almost the same results. To study the impacts of changing the measure of variable inputs, I specify a new production function. Similar to Traina (2018), Karabarbounis and Neiman (2018), and Flynn, Gandhi and Traina (2019), I consider a production function as  $y_{it} = \beta_z z_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}$ , where  $z_{it}$  is the non-capital inputs of production which is measured by the *operating expenses*. Figure 4.2 illustrates the result of markups using this production function where the returns to scale of each industry is defined as  $\gamma = \beta_z + \beta_k$  and the profit rates of firms are the same as before. The red dashed line that represents the markups with the *operating expenses* as the measure of variable inputs follows closely the solid black line that represents the benchmark markups using the production function identified as Equation 4.2.12.

My results indicate that the “total-input” approach is not sensitive to the choice of the conventional measures of variable inputs. The findings are also close to the papers that use either aggregate data or firm-level data with the *operating expenses* as the measure of variable inputs. These findings suggest that the large markups estimated by the papers using the *COGS* are more likely to be driven by measurement issues than by firm heterogeneity. Moreover, in Section 4.5, I show that having a reliable measure of variable inputs reduces the differences between the findings of the exiting method of estimating firm-level markups and the method used in this chapter.

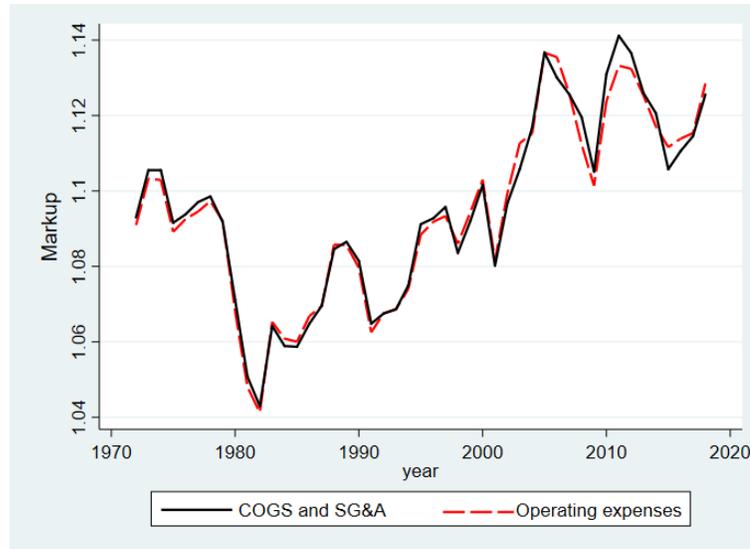


Figure 4.2: Markups using different measures of production function inputs

Notes: The figure illustrates the estimates of markups using the “total-input” approach based on two different sets of inputs from accounting data. The solid black line displays the markups using capital, COGS, and SG&A as the measures of production function inputs, while the red dashed line illustrates the markups using capital and operating expenses as the measures of production function inputs.

#### 4.4.2 The dispersion of markups

The heterogeneity across firms has increased in the past decades. [Autor, Dorn, Katz, Patterson and Van Reenen \(2020\)](#) study the rise of superstar firms and its impacts on labour share decline in the US. [Van Reenen \(2018\)](#) indicates that the increase in heterogeneity between firms is mainly due to new technologies and globalization. Having the firm-level markups allows for tracking the distribution of the markups over time. [Figure 4.3](#) illustrates the rise of superstar firms and the dispersion of markups across firms. To study the dispersion of markups across firms, I sort the firms according to their markups and find the 50, 75, and 90 percentiles of markups. To obtain the percentiles, I weigh each firm by its share of sales. The markups show a remarkably large distribution across firms, especially after the mid-1990s. The median of markups is

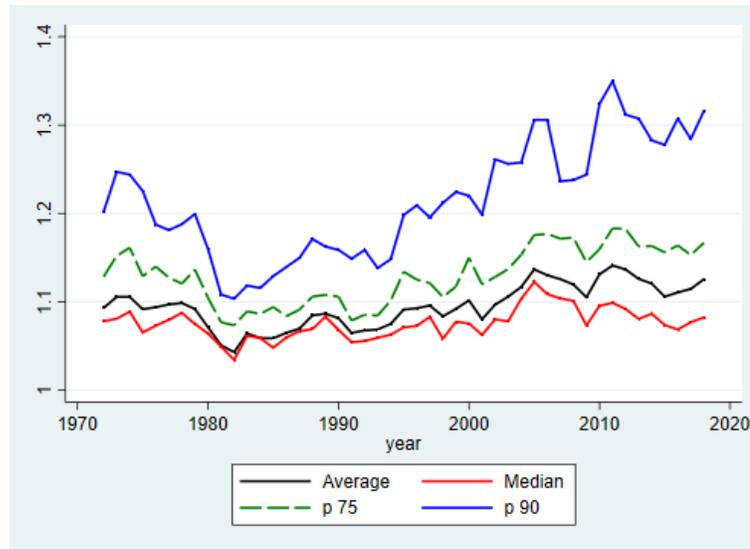


Figure 4.3: The evolution of the distribution (percentiles) of markups for the US

Notes: In this figure, Average, Median, p75, and p90 refer to the weighted average, median, 75 and 90 percentile of markups, respectively. To study the dispersion of markups across firms, I sort the firms according to their markups and find the 50, 75, and 90 percentiles of markups. To obtain the percentiles, I weigh each firm by its share of sales.

below the average and increased by two percentage points from around 2% in 1980 to 6% in 2018. In comparison, markups for the 75th percentiles increased by seven percentage points, from 10% to 17%, and those for the 90th percentiles increased by 16 percentage points from 16% to 32% over the above period. The dispersion of markups indicates that the tendency over time between 90th percentiles and the average has dramatically increased after the mid-1990s. This result is consistent with the finding of [Autor, Dorn, Katz, Patterson and Van Reenen \(2020\)](#) that the sharp fall of labour share since the early 2000s can be due to the rise of superstar firms.

#### 4.4.3 The industry-level markups

Figure 4.4 displays the average weighted markups for four major industries of publicly-traded firms. The average of markups in each industry is calculated using the sales

of firms as the weight. The industries are categorized based on the two-digit North American Industry Classification System (NAICS) codes, i.e. at the sector level. Since 1980, markups have increased in the four major sectors of Mining, Oil and Gas (NAICS 21), Manufacturing (NAICS 31, 32 and 33), Wholesale Trade (NAICS 43) and Retail Trade (NAICS 44 and 45). The largest increase in markups belongs to Manufacturing (NAICS 31) where markups increased by 11 percentage points from 7% in 1980 to 18% in 2018, and the lowest amount corresponds to Wholesale Trade with one percentage point increase, from 2% in 1980 to 3% in 2018. The overall average of markups across the study period (1972-2018) for the sector of Mining, Oil and Gas is 15%, and it is greater than the other three major sectors. Due to a 50% decline in the price of oil and gas starting in 2014 (see [Ellwanger, Sawatzky and Zmitrowicz \(2017\)](#)) the markups for Mining, Oil and Gas sharply declined by 23 percentage points, from 18% to -5%. The decline is almost double of the decrease due to the Great Recession in 2008 for that sector.

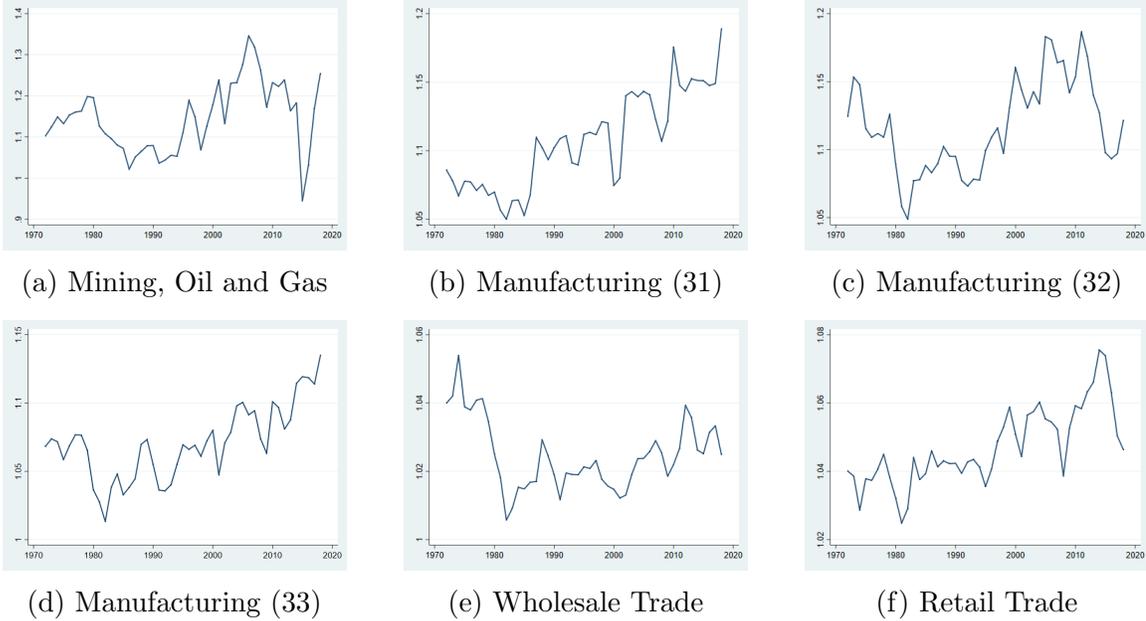


Figure 4.4: The average markups for the major sectors in the US.

## 4.5 Total-input v.s. single-input approach

According to the “single-input” approach, the firm-level markup is defined as the product of two terms, the output elasticity of variable inputs and the ratio between revenue and the cost of variable inputs ( $\mu = \beta_v \frac{PY}{PvV}$ ). An inappropriate measure of variable input not only results in a biased estimate of output elasticity of variable input but also causes an inaccurate ratio of sales (revenues) over the cost of variable inputs. The estimation of markups using this method is very sensitive to the choice of the measure of variable inputs. Since firms do not report variable inputs in their income statements, researchers usually measure variable inputs either using the *COGS* or *operating expenses*. The former undermeasures variable inputs, and the latter overmeasures them. As a result, there are significant discrepancies of up to 50 percentage points in estimating markups using the two measurements (see [Basu](#)

(2019) and Syverson (2019)).

Basu (2019) states that labour is a more reliable measure of variable inputs in the exiting “single-input” approach of estimating firm-level markups. The Compustat dataset reports the wage bill for a subset of firms that contains around 20% of the total firms. I use the wage bill as a measure of (the cost of) variable inputs and estimate value-added markups based on the “single-input” approach for the firms with available wage bill as follows:

$$\mu_{it}^{va} = \beta_l \frac{S_{it}}{W_{it}}, \quad (4.5.1)$$

where  $\beta_l$ ,  $S_{it}$ , and  $W_{it}$  are the output elasticity of variable input (*labour*), *value-added*, and the wage bill, respectively. Then, I transform the value-added markups ( $\mu_{it}^{va}$ ) to gross markups ( $\mu_{it}$ ). See Appendix C.1 for more details.

To compare the results, I also estimate markups using the “total-input” approach using accounting data (i.e. *COGS*, *SG&A*, and *PPE*) only for the firms with available wage bill,

$$\mu_{it} = \gamma \left( \frac{1}{1 - \pi_{it}} \right), \quad (4.5.2)$$

where,  $\gamma = \beta_v + \beta_x + \beta_k$  is the returns to scale of firms with available wage bill, and  $\pi_{it}$  is the economic profit of those firms.

Figure 4.5 illustrates the results of weighted average markups based on both methods for the firms with the available wage bill. The difference in the overall time-series average of markups between the two approaches for the entire study period (1972-

2018) is around three percentage points. It is 9% for the “single-input” approach and 12% for the “total-input” approach. In both cases, markups have increased since 1980. Using the “single-input” approach and considering wage bill as a measure of variable inputs, the gross markups increased by five percentage points, from around 11% in 1980 to around 16% in 2018. Using the “total-input” approach and considering the accounting data of *COGS*, *SG&A*, and *PPE*, the gross markups increased by five percentage points, from 6% in 1980 to 10% in 2018. From 2000 to 2013, the estimated markups of both methods closely follow each other, while for the rest of the years, the two series have around five percentage points gap in their markups.

Assuming the wage bill is a reliable measure of (the cost of) variable inputs, this exercise confirms that the new approach is capable of creating an acceptable estimate of markups. Then one can use this new approach to estimate markups even for those firms that do not report this most reliable measure of variable inputs.



Figure 4.5: The average markups using the two different approaches

Notes: The figure exhibits the estimates of average markups using two different approaches. The red line displays the estimates of markups using the “single-input” approach, where the wage bill is defined as the measure of variable inputs. The US Compustat dataset reports the wage bill for a subset of firms that contains around 20% of total firms. The black line shows the estimates of average markups using the “total-input” approach for the same firms which report the wage bill. I use Capital, COGS, and SG&A as the inputs of production in the “total-input” approach.

## 4.6 Conclusion

The firm-level estimates of markups in the current literature is very sensitive to the measure of variable inputs of production that are obtained from accounting data. Accounting systems report the cost of production in their income statements as direct and indirect expenses. These expenses do not necessarily correspond to the costs of variable and fixed inputs that are required for estimating markups using the “single input” approach. So, it is often difficult to measure (the cost of) variable inputs of production from the direct and indirect expenses of firms. The current literature on firm-level markups reported a discrepancy of up to 50 percentage points in estimated markups depending on the choice of measure of variable inputs from accounting data.

In this chapter, I introduce a refined “production function” approach that is less

sensitive to the choice of measure of variable inputs. This approach not only captures the heterogeneity of firms but also reports an overall average markup that is consistent with the studies on profit rates and returns to scale at the aggregate level (see [Basu \(2019\)](#) and [Ahmad, Fernald and Khan \(2019\)](#)). My results indicate that the average markups increased by six percentage points, from 7% in 1980 to around 13% in 2018. Over the same study period, the economic profit share increased by around five percentage points, from 4% to 9%, and the production function exhibits constant returns to scale.

# Chapter 5

## Conclusion

In this thesis, I focus on studying both the impacts and the estimates of rising market power. In Chapter 2, I explain the impacts of the rising market power on the secular stylized facts of the U.S. economy using a heterogeneous agent model. In Chapter 3, I estimate the evolution of firm-level markups in Canada using the novel T2-Longitudinal Employment Analysis Program (T2-LEAP) database. In Chapter 4, I introduce a new approach to estimate firm-level markups using the U.S. accounting data.

The rise of market power increases the profit share of an economy. An increase in the profit share, *ceteris paribus*, results in declining labour and capital shares of income. In Chapter 2, I demonstrate that not only the increase of profit but also the distribution of profit among heterogeneous households plays a key role in understanding the economic impacts of a rise of market power. When inequality is high, the redistribution of profit among households, who are heterogeneous in terms of income, and hence in terms of wealth and the marginal propensity to consume

(MPC), is uneven. The wealthy, who have lower MPC, gain proportionally higher profits and save more in the form of capital. The increase in aggregate capital pushes down the rate of return on capital through the marginal product of capital (MPK). Therefore, when households are heterogeneous, the rise of market power causes the key secular trends during the past several decades in the U.S. economy.

The increasing markups are associated with lower welfare and higher inefficiency in any economy. Thus, it is essential to study the dynamics of markups over time, as well as their estimates at a point in time. In Chapter 3, I estimate the firm-level markups in Canada using the existing “production function” approach which is also known as the “single-input” approach. Although estimating markups using firm-level data instead of aggregate data has some advantages, it faces some challenges as well. The sample selection bias due to using only publicly-traded firms and the errors in the measures of variable costs are the two main challenges. I address these challenges by using the novel and rich database of T2-LEAP. The database has information from the universe of Canadian firms, allowing me to include both publicly-traded and non-publicly-traded firms in my analysis. I also define the *wage bill* as the measure of variable costs in the production function. The *wage bill* is a fitter measure of variable costs than the *cost of goods sold (COGS)* and the *operating expenses* that are widely used in the literature. In general, the *COGS* undervalues the variable costs, while the *operating expenses* overvalues them. The former leads to overestimating markups while the latter results in underestimating markups. Both measurements should be net of depreciation, and therefore, both of them may also suffer from the possibility that reported accounting depreciation of firms do not correctly reflect true economic

depreciation. I find that from 2002 to 2015, the average of the value-added markups increased mildly from 24% to 29%, and its corresponding gross markups increased from 6% to 9%.

Moreover, I show that contrary to the *COGS*, the markups using the *operating expenses* are not very different from the markups using the *wage bill*. In other words, the *operating expenses* is a better measure of variable costs than the *COGS* if the *wage bill* is not available. I also argue that excluding non-publicly-traded firms may result in overestimates of markups.

The firm-level estimates of markups using the “single-input” approach is very sensitive to what researchers choose as the measures of the single (bundle of) variable costs from accounting data. On the other hand, it is often difficult to measure the variable costs that are recorded as direct and indirect expenses in accounting data. The estimated markups using the “single-input” approaches show a discrepancy of up to 50 percentage points, where most of the discrepancy can be attributed to the choice of the measure of variable costs.

In Chapter 4, I introduce the “total-input” approach as a refined “production function” approach to estimate firm-level markups. In this approach, the markup is defined as a function of economic profit rates and returns to scale of firms, and its results are less sensitive to the choice of measure of variable inputs. This approach not only captures the heterogeneity of firms but also reports an overall average markup that is consistent with the studies on profit rates and returns to scale at the aggregate level. My results indicate that the average markups increased by six percentage points, from 7% in 1980 to around 13% in 2018. Over the same study period, the

economic profit share increased by around five percentage points, from 4% to 9%, and the production function exhibits constant returns to scale.

The rise of market power not only lowers welfare and efficiency in an economy but also has negative distributional effects when inequality is considerably high. To alleviate the costly consequences of market power, the government's policies should target both market power and inequality. To promote more competition, the government can implement policies such as removing entry barriers, and regulating merging and acquisition. For example, [Grullon, Larkin and Michaely \(2019\)](#) find a positive relationship between merging, profitability and market concentration. To reduce income and wealth inequality, the government can design the income tax in favour of low-income households (see [Piketty and Saez \(2013\)](#)).

# References

- Acemoglu, D. and Restrepo, P.: 2018, The race between man and machine: Implications of technology for growth, factor shares, and employment, *American Economic Review* **108**(6), 1488–1542.
- Ackerberg, D. A., Caves, K. and Frazer, G.: 2015, Identification properties of recent production function estimators, *Econometrica* **83**(6), 2411–2451.
- Ahmad, M., Fernald, J. and Khan, H.: 2019, Returns to scale in us production, redux, *Working paper*, Carleton University.
- Aiyagari, S. R.: 1994, Uninsured idiosyncratic risk and aggregate saving, *The Quarterly Journal of Economics* **109**(3), 659–684.
- Alan, S.: 2006, Entry costs and stock market participation over the life cycle, *Review of Economic Dynamics* **9**(4), 588–611.
- Antras, P.: 2004, Is the us aggregate production function cobb-douglas? new estimates of the elasticity of substitution, *Contributions to Macroeconomics* **4**(1), 1–34.
- Auclert, A. and Rognlie, M.: 2017, Aggregate demand and the top 1 percent, *American Economic Review* **107**(5), 588–92.

- Auclert, A. and Rognlie, M.: 2018, Inequality and aggregate demand, *Working Paper 24280*, National Bureau of Economic Research.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J.: 2017, The fall of the labor share and the rise of superstar firms, *Working Paper 23396*, National Bureau of Economic Research.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C. and 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.
- Bachmann, R., Bai, J. H., Lee, M. and Zhang, F.: 2017, The welfare and distributional effects of fiscal volatility: a quantitative evaluation, *Working Paper DP12384*, Center for Economic Policy Research.
- Baldwin, J. R. and Gorecki, P. K.: 1986, *The Role of Scale in Canada/US Productivity Differences in the Manufacturing Sector, 1970-1979*, Vol. 6, Universtiy of Toronto Press.
- Barkai, S.: 2017, Declining labor and capital shares, *Working paper*, University of Chicago.
- Barkai, S.: 2020, Declining labor and capital shares, *The Journal of Finance* **75**(5), 2421–2463.
- Barro, R. J. and Furman, J.: 2018, Macroeconomic effects of the 2017 tax reform, *Brookings Papers on Economic Activity* **2018**(1), 257–345.

- Basu, S.: 1996, Procyclical productivity: increasing returns or cyclical utilization?, *The Quarterly Journal of Economics* **111**(3), 719–751.
- Basu, S.: 2019, Are price-cost markups rising in the united states? a discussion of the evidence, *Journal of Economic Perspectives* **33**(3), 3–22.
- Basu, S. and Fernald, J. G.: 1994, Constant returns and small markups in us manufacturing, *Discussion Paper 483*, Board of Governors of the Federal Reserve System.
- Basu, S. and Fernald, J. G.: 1997, Returns to scale in us production: Estimates and implications, *Journal of Political Economy* **105**(2), 249–283.
- Basu, S. and Fernald, J. G.: 2002, Aggregate productivity and aggregate technology, *European Economic Review* **46**(6), 963–991.
- Berry, S., Gaynor, M. and Scott Morton, F.: 2019, Do increasing markups matter? lessons from empirical industrial organization, *Journal of Economic Perspectives* **33**(3), 44–68.
- Blonigen, B. A. and Pierce, J. R.: 2016, Evidence for the effects of mergers on market power and efficiency, *Working Paper 22750*, National Bureau of Economic Research.
- Bond, S., Hashemi, A., Kaplan, G. and Zoch, P.: 2020, Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data, *Working Paper 27002*, National Bureau of Economic Research.
- Brun, L. and González, I.: 2017, Tobin’s q and inequality, *Available online at SSRN 3069980* .

- Caballero, R. J., Farhi, E. and Gourinchas, P.-O.: 2017, Rents, technical change, and risk premia accounting for secular trends in interest rates, returns on capital, earning yields, and factor shares, *American Economic Review* **107**(5), 614–20.
- Calligaris, S., Criscuolo, C. and Marcolin, L.: 2018, Mark-ups in the digital era, *Working Paper 2018/10*, OECD Science, Technology and Industry.
- Cao, S. and Leung, D.: 2020, Credit constraints and productivity of smes: Evidence from canada, *Economic Modelling* **88**, 163–180.
- Carroll, C. D.: 1997, Buffer-stock saving and the life cycle/permanent income hypothesis, *The Quarterly Journal of Economics* **112**(1), 1–55.
- Carroll, C., Slacalek, J., Tokuoka, K. and White, M. N.: 2017, The distribution of wealth and the marginal propensity to consume, *Quantitative Economics* **8**(3), 977–1020.
- Chamley, C.: 1986, Optimal taxation of capital income in general equilibrium with infinite lives, *Econometrica: Journal of the Econometric Society* **54**(3), 607–622.
- Conesa, J. C., Kitao, S. and Krueger, D.: 2009, Taxing capital? not a bad idea after all!, *American Economic Review* **99**(1), 25–48.
- Crouzet, N. and Eberly, J. C.: 2019, Understanding weak capital investment: the role of market concentration and intangibles, *Working Paper 25869*, National Bureau of Economic Research.
- Curry, B.: 2016, Four major changes to canada’s housing rules.

- URL:** <https://www.theglobeandmail.com/real-estate/four-major-changes-to-canadas-housing-rules/article32223470/>
- Davila, J., Hong, J. H., Krusell, P. and Ríos-Rull, J.-V.: 2012, Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks, *Econometrica* **80**(6), 2431–2467.
- De Loecker, J. and Eeckhout, J.: 2017, The rise of market power and the macroeconomic implications, *Working Paper 23687*, National Bureau of Economic Research.
- De Loecker, J. and Eeckhout, J.: 2018a, Global market power, *Working Paper 24768*, National Bureau of Economic Research.
- De Loecker, J. and Eeckhout, J.: 2018b, Some thoughts on the debate about (aggregate) markup measurement, *Working paper*.
- De Loecker, J., Eeckhout, J. and Unger, G.: 2020, The rise of market power and the macroeconomic implications, *The Quarterly Journal of Economics* **135**(2), 561–644.
- De Loecker, J. and Warzynski, F.: 2012, Markups and firm-level export status, *American Economic Review* **102**(6), 2437–71.
- De Ridder, M.: 2019, Market power and innovation in the intangible economy, *Working Paper 1931*, University of Cambridge.
- Diez, F., Leigh, D. and Tambunlertchai, S.: 2018, Global market power and its

- macroeconomic implications, *Working Paper WP/18/137*, International Monetary Fund.
- Dixit, A. K. and Stiglitz, J. E.: 1977, Monopolistic competition and optimum product diversity, *The American Economic Review* **67**(3), 297–308.
- Domowitz, I., Hubbard, R. G. and Petersen, B. C.: 1986, Market structure and cyclical fluctuations in u.s. manufacturing, *Working Paper 2115*, National Bureau of Economic Research.
- Dorn, D., Katz, L. F., Patterson, C., Van Reenen, J. et al.: 2017, Concentrating on the fall of the labor share, *American Economic Review* **107**(5), 180–85.
- Eggertsson, G. B., Robbins, J. A. and Wold, E. G.: 2018, Kaldor and piketty’s facts: The rise of monopoly power in the united states, *Working Paper 24287*, National Bureau of Economic Research.
- Ellwanger, R., Sawatzky, B. and Zmitrowicz, K.: 2017, Factors behind the 2014 oil price decline, *Bank of Canada Review* **2017**(Autumn), 1–13.
- Elsby, M. W., Hobijn, B. and Şahin, A.: 2013, The decline of the us labor share, *Brookings Papers on Economic Activity* **2013**(2), 1–63.
- Ewens, M., Peters, R. H. and Wang, S.: 2020, Measuring intangible capital with market prices, *Working Paper 25960*, National Bureau of Economic Research.
- Farhi, E. and Gourio, F.: 2018, Accounting for macro-finance trends: Market power,

- intangibles, and risk premia, *Working Paper 25282*, National Bureau of Economic Research.
- Flynn, Z., Gandhi, A. and Traina, J.: 2019, Measuring market power with production data, *Working paper*, Available at SSRN 3358472.
- Ge, S. and Weisbach, M. S.: 2019, The role of financial conditions in portfolio choices: The case of insurers, *Working Paper 25677*, National Bureau of Economic Research.
- Gomes, F. and Michaelides, A.: 2005, Optimal life-cycle asset allocation: Understanding the empirical evidence, *The Journal of Finance* **60**(2), 869–904.
- Gouveia, M. and Strauss, R. P.: 1994, Effective federal individual income tax functions: An exploratory empirical analysis, *National Tax Journal* **47**(2), 317–339.
- Grullon, G., Larkin, Y. and Michaely, R.: 2019, Are u.s. industries becoming more concentrated?, *Review of Finance* **23**(4), 697–743.
- Gu, W. and Macdonald, R.: 2020, Business sector intangible capital and sources of labour productivity growth in canada, *Working Paper 11F0019M No. 442*, Statistics Canada.
- Guo, J.-T. and Lansing, K. J.: 1999, Optimal taxation of capital income with imperfectly competitive product markets, *Journal of Economic Dynamics and Control* **23**(7), 967–995.
- Hall, R. E.: 1988, The relation between price and marginal cost in us industry, *Journal of Political Economy* **96**(5), 921–947.

- Hall, R. E.: 2018, New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy, *Working Paper 24574*, National Bureau of Economic Research.
- Hall, R. E., Blanchard, O. J. and Hubbard, R. G.: 1986, Market structure and macroeconomic fluctuations, *Brookings Papers on Economic Activity* **1986**(2), 285–338.
- Hall, R. E. and Jorgenson, D. W.: 1969, Tax policy and investment behavior: Reply and further results, *The American Economic Review* **59**(3), 388–401.
- Hong, S.: 2018, Price markups for small and large firms over the business cycle, *Economic Synopses* (17), 1–3.
- Innovation, Science and Economic Development Canada: 2019, Key small business statistics, [https://www.ic.gc.ca/eic/site/061.nsf/eng/h\\_03090.html](https://www.ic.gc.ca/eic/site/061.nsf/eng/h_03090.html).
- Jappelli, T. and Pistaferri, L.: 2014, Fiscal policy and mpc heterogeneity, *American Economic Journal: Macroeconomics* **6**(4), 107–36.
- Judd, K. L.: 1985, Redistributive taxation in a simple perfect foresight model, *Journal of Public Economics* **28**(1), 59–83.
- Kaplan, G., Moll, B. and Violante, G. L.: 2018, Monetary policy according to hank, *American Economic Review* **108**(3), 697–743.
- Karabarbounis, L. and Neiman, B.: 2014, The global decline of the labor share, *The Quarterly Journal of Economics* **129**(1), 61–103.

- Karabarbounis, L. and Neiman, B.: 2018, Accounting for factorless income, *Working Paper 24404*, National Bureau of Economic Research.
- Karabarbounis, L. and Neiman, B.: 2019, Accounting for factorless income, *NBER Macroeconomics Annual* **33**(1), 167–228.
- Kaymak, B. and Schott, I.: 2018, Corporate tax cuts and the decline of the labor share, *2018 Meeting Papers 943*, Society for Economic Dynamics.
- Keller, W. and Yeaple, S. R.: 2009, Multinational enterprises, international trade, and productivity growth: firm-level evidence from the united states, *The Review of Economics and Statistics* **91**(4), 821–831.
- Khan, H. and Kim, B.-G.: 2013, Markups and oil prices in canada, *Economic Modelling* **30**, 799–813.
- Khorunzhina, N.: 2013, Structural estimation of stock market participation costs, *Journal of Economic Dynamics and Control* **37**(12), 2928–2942.
- Knoblauch, M., Roessler, M. and Zwerschke, P.: 2020, The elasticity of substitution between capital and labour in the us economy: A meta-regression analysis, *Oxford Bulletin of Economics and Statistics* **82**(1), 62–82.
- Krusell, P. and Smith, Jr, A. A.: 1998, Income and wealth heterogeneity in the macroeconomy, *Journal of Political Economy* **106**(5), 867–896.
- Leung, D.: 2008, Markups in canada: Have they changed and why?, *Working Paper 2008-7*, Bank of Canada.

- Levinsohn, J. and Petrin, A.: 2003, Estimating production functions using inputs to control for unobservables, *The Review of Economic Studies* **70**(2), 317–341.
- Macdonald, R.: 2008, An examination of public capital’s role in production, *Economic Analysis Research Paper 2008050*, Statistics Canada.
- Martinez, J.: 2018, Automation, growth and factor shares, *2018 Meeting Papers 943*, Society for Economic Dynamics.
- Martins, J. O., Scarpetta, S. and Pilat, D.: 1996, Mark-up pricing, market structure and the business cycle, *OECD Economic Studies* (27), 71–106.
- McCarthy, J.: 1995, Imperfect insurance and differing propensities to consume across households, *Journal of Monetary Economics* **36**(2), 301–327.
- Mirrlees, J. A.: 1971, An exploration in the theory of optimum income taxation, *The Review of Economic Studies* **38**(2), 175–208.
- Mishkin, F. S. and Serletis, A.: 2018, *The Economics of Money, Banking, and Financial Markets*, 7th Canadian edn, Pearson Canada, Toronto.
- Nekarda, C. J. and Ramey, V. A.: 2013, The cyclical behavior of the price-cost markup, *Working Paper 19099*, National Bureau of Economic Research.
- Nekarda, C. J. and Ramey, V. A.: 2020, The cyclical behavior of the price-cost markup, *forthcoming Journal of Money, Credit, and Banking: 50th Anniversary Issue* .

- Norrbin, S. C.: 1993, The relation between price and marginal cost in us industry: a contradiction, *Journal of Political Economy* **101**(6), 1149–1164.
- OECD: 2019, Digitalisation and productivity: a story of complementarities, *OECD Economic Outlook* **2019**, 55–83.
- Olley, G. S. and Pakes, A.: 1992, The dynamics of productivity in the telecommunications equipment industry, *Working Paper 3977*, National Bureau of Economic Research.
- Piketty, T.: 2014, *Capital in the 21st Century*, Harvard University Press Cambridge, MA.
- Piketty, T. and Saez, E.: 2007, How progressive is the us federal tax system? a historical and international perspective, *Journal of Economic Perspectives* **21**(1), 3–24.
- Piketty, T. and Saez, E.: 2013, Optimal labor income taxation, *Handbook of Public Economics* **5**, 391–474.
- Piketty, T., Saez, E. and Zucman, G.: 2017, Distributional national accounts: methods and estimates for the united states, *The Quarterly Journal of Economics* **133**(2), 553–609.
- Piketty, T. and Zucman, G.: 2014, Capital is back: Wealth-income ratios in rich countries 1700–2010, *The Quarterly Journal of Economics* **129**(3), 1255–1310.
- Raurich, X., Sala, H. and Sorolla, V.: 2012, Factor shares, the price markup, and the

- elasticity of substitution between capital and labor, *Journal of Macroeconomics* **34**(1), 181–198.
- Rotemberg, J. J. and Woodford, M.: 1991, Markups and the business cycle, *NBER Macroeconomics Annual* **6**, 63–129.
- Rotemberg, J. J. and Woodford, M.: 1993, Dynamic general equilibrium models with imperfectly competitive product markets, *Working Paper 4502*, National Bureau of Economic Research.
- Rotemberg, J. J. and Woodford, M.: 1999, The cyclical behavior of prices and costs, *Handbook of Macroeconomics* **1**, 1051–1135.
- Saez, E. and Zucman, G.: 2016, Wealth inequality in the united states since 1913: Evidence from capitalized income tax data, *The Quarterly Journal of Economics* **131**(2), 519–578.
- Saez, E. and Zucman, G.: 2020, Trends in us income and wealth inequality: Revising after the revisionists, *Working Paper 27921*, National Bureau of Economic Research.
- Santomero, A. M.: 1984, Modeling the banking firm: A survey, *Journal of money, credit and banking* **16**(4), 576–602.
- Scheuer, F. and Slemrod, J.: 2020, Taxing our wealth, *Working Paper 28150*, National Bureau of Economic Research.
- Sealey, C. W. and Lindley, J. T.: 1977, Inputs, outputs, and a theory of production

- and cost at depository financial institutions, *The journal of finance* **32**(4), 1251–1266.
- Shapiro, M. D.: 1987, Measuring market power in u.s. industry, *Working Paper 2212*, National Bureau of Economic Research.
- Stansbury, A. and Summers, L. H.: 2020, The declining worker power hypothesis: An explanation for the recent evolution of the american economy, *Working Paper 27193*, National Bureau of Economic Research.
- Stewart, N.: 2015, Benefits benchmarking 2015, <https://www.conferenceboard.ca/e-library/abstract.aspx?did=7364>.
- Straub, L. and Werning, I.: 2020, Positive long-run capital taxation: Chamley-judd revisited, *American Economic Review* **110**(1), 86–119.
- Syverson, C.: 2019, Macroeconomics and market power: Context, implications, and open questions, *Journal of Economic Perspectives* **33**(3), 23–43.
- Traina, J.: 2018, Is aggregate market power increasing? production trends using financial statements, *Working paper*, University of Chicago.
- Van Reenen, J.: 2018, Increasing differences between firms: market power and the macro-economy, *Discussion Paper CEPDP1576*, London School of Economics and Political Science.
- Vissing-Jorgensen, A.: 2002, Towards an explanation of household portfolio choice

heterogeneity: Nonfinancial income and participation cost structures, *Working Paper 8884*, National Bureau of Economic Research.

Wolff, E. N.: 2017, Household wealth trends in the united states, 1962 to 2016: Has middle class wealth recovered?, *Working Paper 24085*, National Bureau of Economic Research.

Young, A. T. and Tackett, M. Y.: 2018, Globalization and the decline in labor shares: Exploring the relationship beyond trade and financial flows, *European Journal of Political Economy* **52**, 18–35.

# Appendix A

## Appendix for Chapter 2

### A.1 Firms

#### A.1.1 Consumption good producers

The perfectly competitive final producers in corporate sector produce  $Y_t$  from assembling a continuum of differentiated intermediate input good  $y_{i,t}$  as follows

$$Y_t = \left( \int_0^1 y_{i,t}^{\frac{\epsilon_t-1}{\epsilon_t}} di \right)^{\frac{\epsilon_t}{\epsilon_t-1}}, \quad (\text{A.1.1})$$

Where  $\epsilon_t > 1$  denotes the elasticity of substitution between input varieties,  $y_{i,t}$ .

The consumption good producers buy inputs at price  $p_{i,t}$  from the monopolistically-competitive firms.

To obtain input demand function, the final producers minimize their costs,

$$\min_{y_{i,t}} \int_0^1 p_{i,t} y_{i,t} di, \quad (\text{A.1.2})$$

subject to the technology constraint,

$$Y_t = \left( \int_0^1 y_{i,t}^{\frac{\epsilon_t-1}{\epsilon_t}} di \right)^{\frac{\epsilon_t}{\epsilon_t-1}}. \quad (\text{A.1.3})$$

As a result of this optimization, the (conditional) downward input demand functions is equal to

$$y_{i,t} = \left( \frac{p_{i,t}}{P_t^Y} \right)^{-\epsilon_t} Y_t. \quad (\text{A.1.4})$$

or

$$p_{i,t} = \left( \frac{y_{i,t}}{Y_t} \right)^{-\frac{1}{\epsilon_t}} P_t. \quad (\text{A.1.5})$$

The final good producers are price taker , i.e  $P_t^Y = MC$ . Therefore, the cost minimization implies that the exogenous price of the final consumption good is such that,

$$P_t^Y = \left( \int_0^1 p_{i,t}^{1-\epsilon_t} di \right)^{\frac{1}{1-\epsilon_t}}. \quad (\text{A.1.6})$$

### A.1.2 Intermediate good producers

The intermediate good firms are owned by the stockholders. The monopolistically-competitive firms produce each intermediate good  $i$  according to the CRS technology  $y_{i,t} = F(k_{i,t}, n_{i,t})$ .  $y_{i,t}$  is the (intermediate) output per unit of efficiency labour,  $k_{i,t}$  is

the efficiency units of capital that are rented from the bondholders at price  $R_{i,t}$  and the efficiency units of labour,  $n_{i,t}$ , are rented from households at price  $W_{i,t}$ . Stockholders own the intermediate good firms. In order to optimize the utility of stockholders, monopolistic intermediate firms maximize their profit subject to the CES technology:

$$\max_{y_{i,t}, k_{i,t}, n_{i,t}} \Pi_{i,t} = p_{i,t}(y_{i,t})y_{i,t} - R_{i,t}k_{i,t} - W_{i,t}n_{i,t} \quad (\text{A.1.7})$$

subject to their CES technology

$$y_{i,t} = \left[ \alpha(k_{i,t})^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(n_{i,t})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (\text{A.1.8})$$

Considering Equation A.1.5, I set up the Lagrangian as

$$L(\cdot) = \left(\frac{y_{i,t}}{Y_t}\right)^{-\frac{1}{\epsilon_t}} P_t y_{i,t} - R_{i,t}k_{i,t} - W_{i,t}n_{i,t} + \lambda_{it} \left[ y_{i,t} - \left[ \alpha(k_{i,t})^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(n_{i,t})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \right]. \quad (\text{A.1.9})$$

The First Order Conditions (F.O.Cs) with respect to output, capital and labour imply

$$p_{i,t} = \lambda_{it} \left( \frac{\epsilon_{it}}{\epsilon_{it} - 1} \right), \quad (\text{A.1.10})$$

$$R_{i,t} = \left( \frac{p_{i,t}}{\mu_t} \right) F_{k_{i,t}} = \left( \frac{p_{i,t}}{\mu_t} \right) \alpha \left( \frac{y_{i,t}}{k_{i,t}} \right)^{\frac{1}{\sigma}}, \quad (\text{A.1.11})$$

$$W_{i,t} = \left( \frac{p_{i,t}}{\mu_t} \right) F_{n_{i,t}} = \left( \frac{p_{i,t}}{\mu_t} \right) (1-\alpha) \left( \frac{y_{i,t}}{n_{i,t}} \right)^{\frac{1}{\sigma}}. \quad (\text{A.1.12})$$

The intermediate firms face a downward demand curve. As it is shown in Equation A.1.10, the firms set their prices about marginal cost markups, i.e.  $\mu_t = \frac{\epsilon_t}{\epsilon_t - 1} \geq 1$ .

By substituting the optimal price of labour and capital into Equation [A.1.7](#), and normalizing the output price (i.e.  $p_{i,t} = 1$ ), I can obtain the following profit function:

$$\Pi_{i,t} = \left(1 - \frac{1}{\mu_t}\right)y_{i,t}. \quad (\text{A.1.13})$$

## A.2 The rise of market power and its measurement

In recent years, several papers have studied the rise of market power in the US economy. While it is widely accepted that market power has increased since the 1980s, there are different ideas about the size of the increase. Table [A.1](#) lists papers that have studied the rise of market power in the US by estimating time-varying markups. In my model, I increase the markup from 6% in the first steady state in the 1980s to 20% in the second steady state in the last decade, and then study the effects of that on macroeconomic stylized facts. I choose this range because this is close to the median of the range from the literature. I also support this range by displaying the papers that estimate the value of non-time-varying markups around the 1980s (for calibrating the markup in the first steady state) and the 2010s (for calibrating the markup in the second steady state) in Table [A.2](#).

In addition to markup, the literature also evaluates the rise of market power by the Lerner and Herfindahl-Hirschman indices. The Lerner index, which is another measure of a firm's market power, is defined as  $\left(\frac{\text{price} - \text{marginal cost}}{\text{price}}\right)$ . The Herfindahl-Hirschman index is a measure of market concentration obtained based on the share of firms in a market. The Herfindahl-Hirschman index is high when few firms have

a large share of a market. [Grullon, Larkin and Michaely \(2019\)](#) find that the Lerner index increased from around 8% in the 1980s to around 11% in the US. They also state that 75% of the US industries have experienced an increase in the Herfindahl-Hirschman index, and consequently, the level of market concentration has increased. Similarly, [Dorn, Katz, Patterson, Van Reenen et al. \(2017\)](#) document the rise of superstar firms in the US. Therefore, regardless of the method of estimation, studies confirm that market power has increased in the US economy since the early 1980s.

In the first essay, I focus on the consequences of the rise of market power rather than its causes. In recent years a wave of studies tries to examine the drivers of market power, and some find that intangible assets play an important role in explaining the rise of market power. [Crouzet and Eberly \(2019\)](#) state that intangible capitals are not distributed equally across firms, and the legal protection of scalable intangible capitals will lead to market concentration. [De Ridder \(2019\)](#) indicates that using intangible capital as a fixed cost of production reduces firms' marginal cost, which results in a rise of market power for intangible-intensive firms. [Calligaris, Criscuolo and Marcolin \(2018\)](#) find that markup differentials between digitally-intensive and less-digitally-intensive sectors have increased significantly over time in OECD countries. See [OECD \(2019\)](#) for a summary of the studies that discuss the relationship between intangible assets and the rise of market power.

Table A.1: The estimated increase of markups for the US economy

<b>Paper</b>	<b>Estimated markup</b>	<b>Period</b>	<b>Data</b>	<b>Time varying</b>
<a href="#">Raurich, Sala and Sorolla (2012)</a>	18% (1980) to 35%(2006)	1980 - 2006	Aggregate level	Yes
<a href="#">Traina (2018)</a>	9% (1980) to 15% (2016)	1950 - 2016	Firm level	Yes
<a href="#">Hall (2018)</a>	12%(1998) to 38% (2015)	1988 - 2015	Industry level	No
<a href="#">Karabarbounis and Neiman (2019)</a>	14% (1980) to 22% (2016)	1950 - 2016	Firm level	Yes
<a href="#">Flynn, Gandhi and Traina (2019)</a>	4% (1980) to 9% (2018)	1950 - 2018	Firm level	Yes
<a href="#">De Loecker, Eeckhout and Unger (2020)</a>	21% (1980) to 61% (2016)	1950 - 2016	Firm level	Yes

Table A.2: The estimated value of markup for the US economy

<b>Paper</b>	<b>Estimated markup</b>	<b>Period</b>	<b>Data</b>	<b>First/Last decade</b>
<a href="#">Norrbin (1993)</a>	Average: 5%	1950 - 1984	Industry level	First decade
<a href="#">Basu and Fernald (1997)</a>	Implied markup: 4%	1959 - 1989	Industry level	First decade
<a href="#">Basu and Fernald (2002)</a>	Average: 5% and 13%	1961 - 1989	Industry level	First decade
<a href="#">Ahmad, Fernald and Khan (2019)</a>	Implied markup: 12%	1989 - 2014	Industry level	Last decade

# Appendix B

## Appendix for Chapter 3

### B.1 Wage bill and the US markups

In general, my finding confirms that the firm-level average markups only increased moderately using the wage bill as the measure of the variable input. However, [De Loecker, Eeckhout and Unger \(2020\)](#) do a robustness check in their published paper and claim that the average markups increased substantially by around 35% percentage points when they consider the wage bill as the measure of the variable input. To address this discrepancy, I use their Compustat dataset to estimate markups using the wage bill. The Compustat dataset reports the wage bill for a subset of firms that contains around 20% of total firms in the US, and the following analysis is done for those firms where their wage bill is available.

To this end, I follow the same methodology that is explained as the baseline model in Section 3.2. To convert nominal variables to real ones, I deflate the variables using the consumer price index (1984=100). The result indicates that the average

gross markups increased from 8% in the early 1980s to 17% in 2018 for the US publicly-traded firms (see Figure B.1). I compare my finding of the wage bill against the alternative measures of variable inputs (the *COGS* and the *operating expenses*) using the same firms which report the wage bill. Using the *operating expenses*, the average gross markups increased from 4% in the early 1980s to 10% in 2018, while using the *COGS* resulted in increasing markups from 10% to 40% over the same period of studies (see Figure B.2). Hence, the markups using the *operating expenses* closely follow the markups of the wage bill, while the *COGS* tends to overestimate the markups significantly.

Although the results of markups using the *COGS* and the *operating expenses* for the subset of firms with the wage bill information are in line with the literature of firm-level markups, I compare them with the results of markups obtained from the total firms in the US Compustat. The result is depicted in Figure B.3. It illustrates almost similar markups to the subset of firms with the wage bill report; and it is consistent with the literature of the firm-level markups. As already discussed, this finding confirms the statement of Basu (2019) that theoretically a more inclusive variable (such as *operating expenses*) is a better choice than the *COGS* when the wage bill is not available for the measure of variable inputs.

Generating intermediate inputs and value-added outputs are the main problems of the estimating markups by De Loecker, Eeckhout and Unger (2020) using the wage bill. Compustat does not provide the intermediate inputs, so I follow the work of Keller and Yeaple (2009) to generate them. The process is briefly explained in Section 3.3. In the following, I discuss further how I generate the intermediate inputs

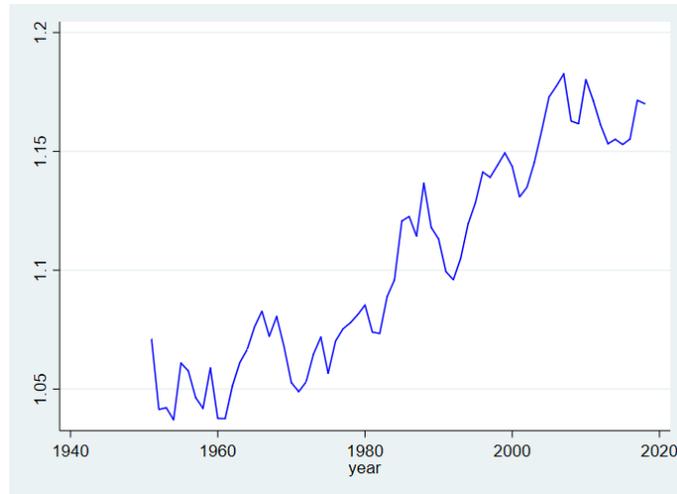


Figure B.1: The US average markups using the wage bill. The markups are weighted based on the value-added of firms.

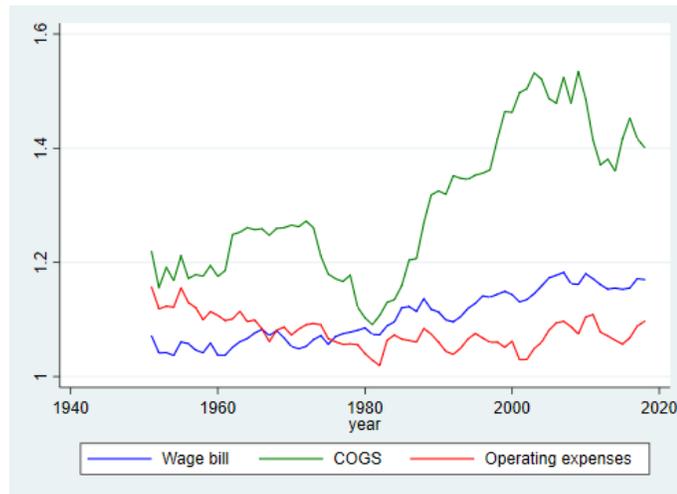


Figure B.2: The comparison of the US markups using the wage bill, the operating expenses, and the COGS.

from the income statements of firms to clarify the discrepancy between my results and the work of [De Loecker, Eeckhout and Unger \(2020\)](#).

### B.1.1 Generating intermediate inputs

As a standard practice, I classify the inputs of production into three categories of labour, capital and intermediate inputs. Any inputs that are neither labour nor

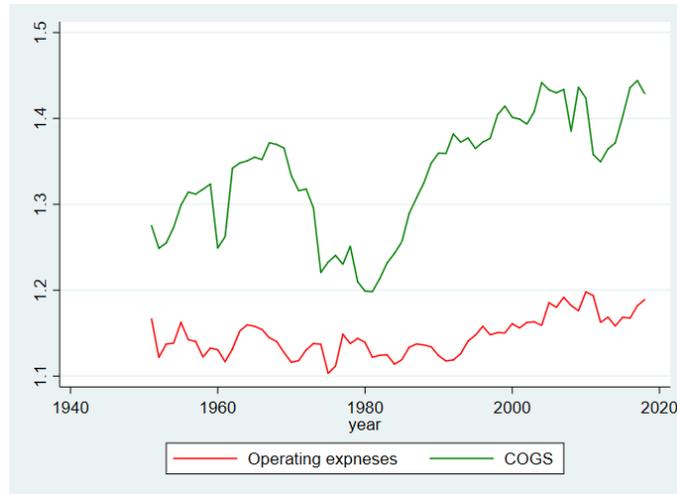


Figure B.3: The US markups using the COGS and operating expenses for the whole observation of Compustat

capital will be categorized as intermediate inputs, such as materials and energy. That being said, one can define three kinds of costs for a typical firm. The cost of labour (wage bill), the cost of capital (depreciation), and the cost of intermediate inputs. Therefore, to generate the cost of intermediate inputs, one can deduct the wage bill and the depreciation from the total cost of a firm:

$$\textit{The cost of intermediate inputs} = \textit{total cost} - \textit{the cost of labours} - \textit{the cost of capital} \quad (\text{B.1.1})$$

In accounting terminology, the total cost of a firm is called the *operating expenses*. In other words, one can obtain the cost of intermediate inputs as follows:

$$\textit{The cost of intermediate inputs} = \textit{operating expenses} - \textit{wage bills} - \textit{total depreciation} \quad (\text{B.1.2})$$

As discussed in Section 3.4, the *operating expenses* consists of the *cost of goods sold* (*COGS*) and the *selling, general and administrative* (*SG&A*) expenses. It should be emphasized that the *operating expenses* and the *COGS* in Compustat are net of depreciation.

Obtaining the total cost of a firm from its income statement is the main reason for the discrepancy between my results and the ones from De Loecker, Eeckhout and Unger (2020). I consider the *operating expenses* as the total cost of a firm, while they define *COGS* as the total cost. In particular, they generate intermediate inputs from the income statement using the following equation:

$$\textit{The cost of intermediate inputs} = \textit{COGS} - \textit{wage bills} - \textit{total depreciation} \quad (\text{B.1.3})$$

There are three main problems with the definition of the cost of intermediate inputs in Equation B.1.3. First, the *COGS* undervalues the total cost of firms because it excludes the *SG&A* expenses from the total costs of firms. Consequently, the *COGS* undermeasures the cost of intermediate inputs. The undermeasured intermediate inputs will become more severe over time because the share of *SG&A* expenses in the total costs of firms has increased since the 1980s.<sup>1</sup> Figure B.4 illustrates the size of undermeasured intermediate inputs over time. The figure demonstrates the trend of the ratio obtained using the two different measurement methods of intermediate inputs: the numerator of the ratio is the cost of intermediate inputs based on the

---

<sup>1</sup>As another problem, one can say that the undermeasured intermediate inputs affect the measure of the productivity of firms if the intermediate inputs are used as a proxy for the unobservable productivity.

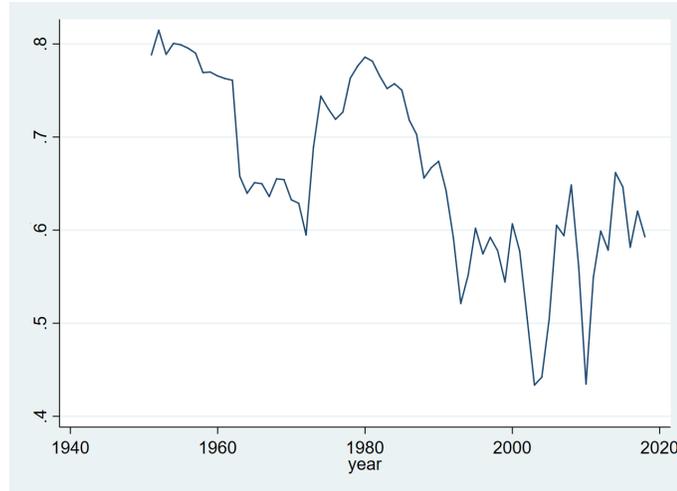


Figure B.4: The average ratio of intermediate inputs

Notes: The figure demonstrates the trend of the intermediate input ratio obtained using two different measurement methods of intermediate inputs. The numerator of the ratio is the cost of intermediate inputs based on the method used by [De Loecker, Eeckhout and Unger \(2020\)](#) and explained in Equation B.1.3 ( $intermediate\ inputs = COGS - wage\ bills - depreciation$ ) and the denominator is the cost of intermediate input that I define in Equation B.1.2 ( $intermediate\ inputs = operating\ expenses - wage\ bills - depreciation$ ).

COGS (Equation B.1.3) and the denominator is based on operating expenses (Equation B.1.2). The ratio has decreased from around 0.8 in the early 1980s to around 0.6 in 2018. Second, the undermeasured intermediate inputs will overvalue the value-added output because it is defined as gross output minus intermediate inputs. Third, the reported wage bill in Compustat represents the payments to all employees, including both direct and indirect labours.<sup>2</sup> Even if I assume that Equation B.1.3 is theoretically correct, I am not allowed to deduct the wage bill of direct labours only (for the definition of direct and indirect labour, see Section 3.4).

The undermeasured intermediate inputs (the first problem) cause to overestimate gross markups through Equation 3.5.2. In particular, the undermeasured intermediate

<sup>2</sup>Compustat denotes the wage bill as “xlr” and explains it as “This item represents direct payments to, and indirect payments on behalf of, all employees... Companies using Income Statement Model Numbers 01, 03, and 04 usually report this item in a note to the statements but the effects are included in Operating Expense.”



Figure B.5: The average markups estimated using the two different costs of intermediate inputs

Notes: The figure displays the evolution of markups using wage bill as a measure of variable inputs based on two different definitions of the intermediate inputs. The red line exhibits the markups when the intermediate input is defined as  $intermediate\ inputs = COGS - wage\ bills - depreciation$ . This definition is used by [De Loecker, Eeckhout and Unger \(2020\)](#). The black line, on the other hand, illustrates the markups when the intermediate inputs is defined as  $intermediate\ inputs = operating\ expenses - wage\ bills - depreciation$ .

inputs undervalue  $S_m$ , the share of intermediate inputs in total sales of firms, which tends to overestimate gross markups. Moreover, the overmeasured value-added output (the second problem) will result in overestimation of value-added markups by overmeasuring the ratio of value-added over the labour cost in Equation 3.2.11. It is obvious that the overestimation of the value-added markups results in overestimation of gross markups via Equation 3.5.2.

To sum up, using the *COGS* as the total cost of a firm undervalues the intermediate inputs and overvalues the value-added output. Consequently, using the *COGS* rather than *operating expenses* tends to overestimate gross markups. Figure B.5 compares the average of gross markups based on both methods of computing intermediate inputs. Using the intermediate inputs that are computed based on the *COGS*, the average markups increased from 23% in 1981 to 35% in 2018, while the markups

increased from 8% to 17% using the intermediate inputs that computed based on *operating expenses* as the total cost of firms.

## B.2 Returns to scale

Few papers have estimated the returns to scale of the Canadian economy. Among them, [Macdonald \(2008\)](#) finds that the Canadian private sector has constant returns to scale. [Baldwin and Gorecki \(1986\)](#) estimate an increasing returns to scale for the Canadian manufacturing sector. They find that manufacturing sector experiences the returns to scale of 1.155 and 1.153 in 1970 and 1979 respectively.

As a side-product of Chapter 2, I also obtain the average returns to scales (ARTS) of the Canadian economy based on firm-level data. Table B.1 summarizes the ARTS from various specifications of the model discussed in Chapter 2. The first column indicates whether the value-added or gross output production is considered. The second column indicates whether it uses the universe of firms or only the publicly-traded firms. The third column explains the measures of the variable inputs used. The fourth and fifth columns demonstrate the value of output elasticities of variable inputs ( $\beta_l$  for the benchmark;  $\beta_v$  for the others) and those of fixed inputs ( $\beta_k$ ). The sixth column illustrates the value of ARTS, and the final two columns denote the sample period and the sources of data. The value of ARTS for all models are around one which confirms the finding of [Macdonald \(2008\)](#) that Canadian industries have constant returns to scale. Surprisingly, the elasticities with respect to variable inputs demonstrate similar values regardless of the model specification used. Taking this

and Equation 3.2.11 into consideration, the value of markups is mostly determined by the ratio of sales (or value-added) over the cost of the variable input. Hence, it is very crucial to have a reliable measure of variable input of production to estimate markups precisely.

Table B.2 demonstrates the value-added returns to scale of Canadian sectors for the benchmark model. Most sectors show roughly constant returns to scale with high elasticity of the variable inputs and low elasticity of the fixed input. The measurement error in fixed investment, the declining of physical investment during the study period as well as the short period of study, i.e. 14 years, may explain the low value of fixed input elasticity. Most of these firms are established many years ago and T2-LEAP does not provide any information about the first year of the firms. Hence, researchers cannot have a good proxy for the initial investment of the firms. This problem is less severe with Compustat since the data goes back to 1950s.

Table B.1: The average returns to scale (ARTS)

Types of Production	Variety of Firms	Variable Input Measure	Elasticity ( $\beta_l$ or $\beta_v$ )	ARTS	Estimated Period	Source
Value-added	Universe	Wage Bill	0.96	0.99	2002-2015	T2-LEAP
Gross Output	Universe	OPEX	0.96	0.98	2002-2015	T2-LEAP
Gross Output	Public	OPEX	0.84	0.97	1960 - 2018	Compustat
Gross Output	Public	COGS	0.85	0.99	1960-2018	Compustat

Table B.2: The value-added returns to scale (RS) and the elasticity of variable inputs ( $\beta_i$ ) at the industry level for the benchmark model

NAICS	Sector	$\beta_i$	RS
11	Agriculture, Forestry and Fishing	0.88	0.99
21	Mining, Oil and Gas	0.93	1.01
22	Utilities	0.96	1.05
23	Construction	0.96	1.00
31	Manufacturing	1.02	1.04
32	Manufacturing	0.99	1.02
33	Manufacturing	0.99	1.01
42	Wholesale Trade	1.00	1.01
44-45	Retail Trade	0.93	0.95
48-49	Transportation and Warehousing	0.06	1.01
51	Information and Cultural	0.99	1.01
54	Professional, Scientific and Technical Services	0.96	0.97
56	Administrative and Waste Management Services	0.96	0.99
71	Art and Entertainment	0.97	1.00
72	Food Services and Accommodation	0.98	1.00
81	Other Services	1.00	1.02

## B.3 The value-added markups at the industry level

Figure B.6 illustrates the average value-added markups for 16 industries at two-digits NAICS. To estimate markups, the wage bill is considered as the measure of variable input. In order to calculate the average markups, the estimated markups are weighted over the value-added of firms at each industry.

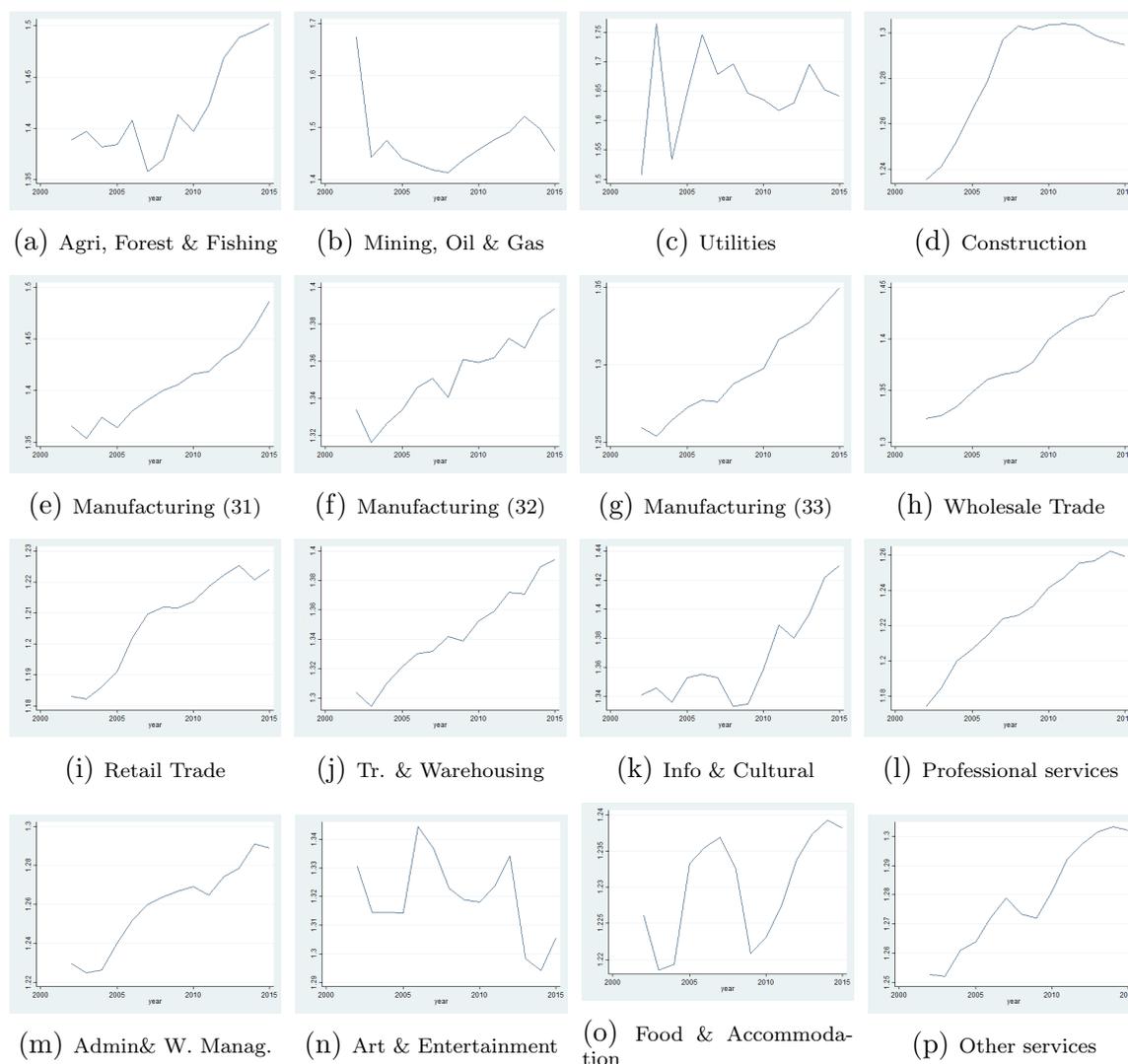


Figure B.6: The average of the value-added markups at the industry level (source: T2-LEAP, 2002-2015).

## B.4 The gross markups at the industry level

Figure B.7 displays the average of gross markups at two-digits NAICS for 16 industries. The gross markups are the transformation of the estimated value-added markups using Equation 3.5.2.

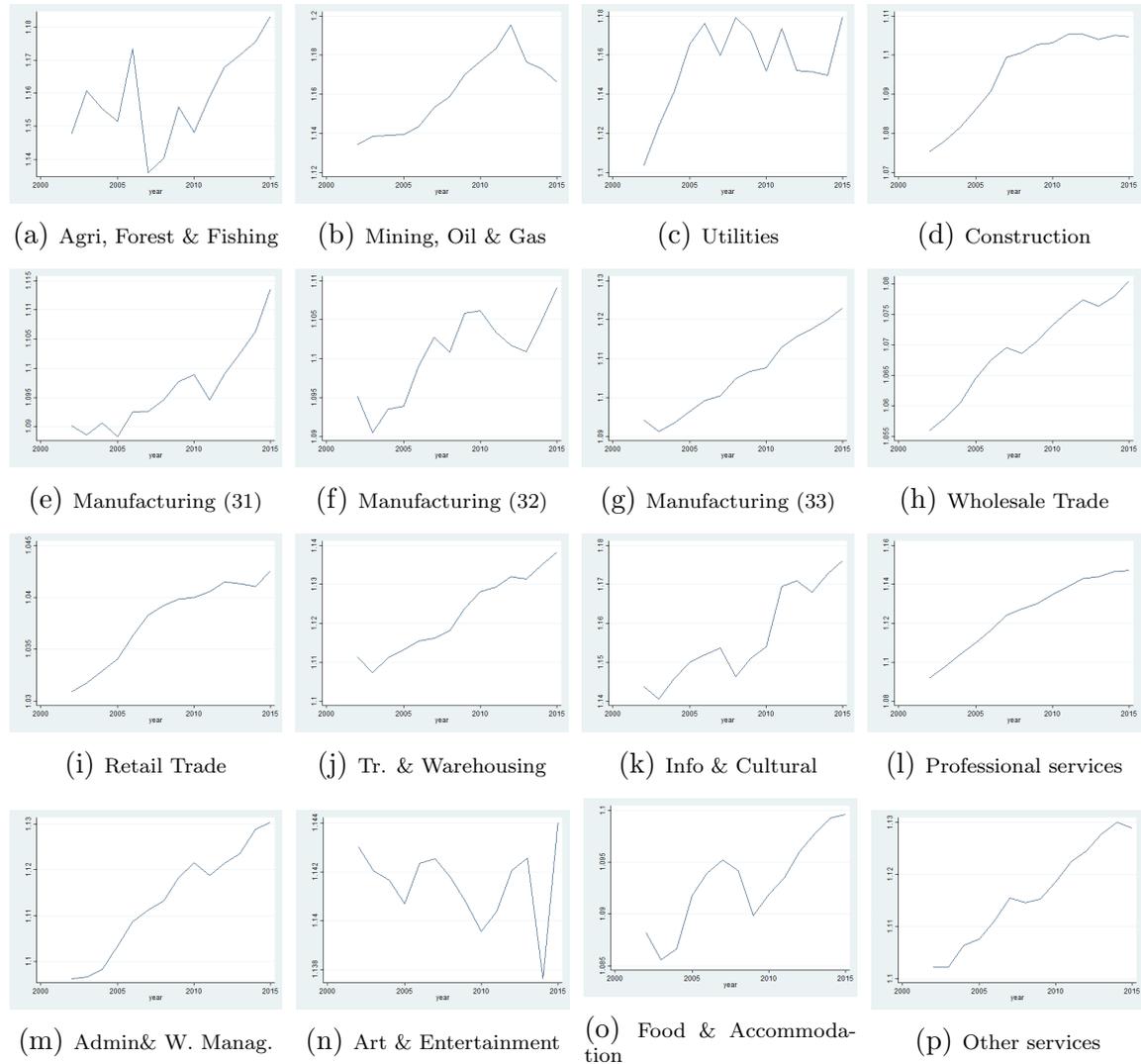


Figure B.7: The average of the gross markups at the industry level (source: T2-LEAP, 2002-2015).

# Appendix C

## Appendix for Chapter 4

### C.1 Estimating markups using wage bill

To estimate markups using the *wage bill* as a measure of variable inputs, I closely follow closely the work of [De Loecker and Warzynski \(2012\)](#), which is based on the spirit of [Hall \(1988\)](#). Particularly, in each period, firm  $i$  of each industry minimizes its costs, and the following equation for value-added markups ( $\mu_{it}^v$ ) can be obtained from the F.O.Cs:

$$\mu_{it}^v = \beta_l \frac{P_{it}Q_{it}}{P_{it}^l L_{it}}, \quad (\text{C.1.1})$$

where,  $P_{it}Q_{it}$  and  $P_{it}^l L_{it}$  are the value-added output and the cost of the variable input (labour), respectively. Both of them can be obtained from the Compustat database. However, the output elasticity of the variable input,  $\beta_l$ , needs to be estimated. In order to obtain the elasticity of variable input, following the work of [Akerberg, Caves](#)

and Frazer (2015), I estimate the following value-added production function for each industry:

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}, \quad (\text{C.1.2})$$

where  $q_{it}$ ,  $l_{it}$ ,  $k_{it}$ , and  $\omega_{it}$  are the logs of deflated firm-level value-added, variable labour input, capital, and productivity respectively; and  $\epsilon_{it}$  is an *i.i.d.* measurement error. Since the above equation is in log-log form, the estimated coefficients will be interpreted as elasticity. The estimated coefficient of the variable input,  $\hat{\beta}_l$ , is the output elasticity of variable inputs. Productivity ( $\omega_{it}$ ) captures factors such as managerial ability, expected defect rates, or soil quality, which are unobservable by econometricians but observable by firms when they take input decisions. Therefore, unobservable  $\omega_{it}$  is likely to be correlated with inputs, so omitting it will result in biased estimates of the elasticities. Similar to De Loecker and Warzynski (2012), I follow the works of Olley and Pakes (1992), Levinsohn and Petrin (2003), and Akerberg, Caves and Frazer (2015) to define productivity as a function of intermediate inputs<sup>1</sup>; and I rely on the two-stage method suggested by the authors to estimate the production function and the elasticity. In particular, I assume that productivity follows an AR(1) process,  $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$ , and impose the following moment conditions:

$$E\left(\xi_{it}(\beta) \begin{pmatrix} k_{it} \\ l_{it-1} \end{pmatrix}\right) = 0, \quad (\text{C.1.3})$$

---

<sup>1</sup>Compustat does not provide data on intermediate inputs so, I follow the work of Keller and Yeaple (2009) and define the cost of intermediate inputs as  $m_{it} = \text{operating expenses} - \text{wage bill} - \text{depreciation of capital}$ . Obviously, the *value-added* output is defined as *sales* minus the cost of *intermediate input*.

where  $\beta = (\beta_l, \beta_k)$ . The key idea behind constructing the moment conditions is that the shock to the productivity should be orthogonal to the decisions on the inputs made in the previous period ( $k_{it}$  and  $l_{i,t-1}$ ). I plug  $\hat{\beta}_l$  back into Equation C.1.1 and estimate value-added markups.

In the process of the production function estimation, intermediate inputs are used as a proxy for unobservable productivity. So, I choose the value-added output instead of the gross output to estimate the production function. However, [Hall, Blanchard and Hubbard \(1986\)](#) state that the exclusion of intermediate inputs from gross output may lead to an overestimate of the markups (see [Domowitz, Hubbard and Petersen \(1986\)](#), [Shapiro \(1987\)](#), [Rotemberg and Woodford \(1993\)](#), [Basu \(1996\)](#), and [Basu and Fernald \(2002\)](#)). To compute gross markups, I follow the above literature and assume that gross production function is Leontief in intermediate inputs, i.e. the production function is proportional to intermediate inputs as the following:

$$Q_t^g = \min \left\{ \frac{M_t}{\theta_t}, \frac{F(K_t, L_t)}{1 - \theta_t} \right\}, \quad (\text{C.1.4})$$

where  $Q_t^g$  is gross output,  $M_t$  is the amount of intermediate inputs,  $F()$  is the value-added production function with labour ( $L_t$ ) and capital ( $K_t$ ) as the inputs, and  $\theta_t$  is the weight of intermediate input in one unit of output. Given the definition of the marginal cost of production in gross and value-added output, the following relationship can be obtained between the gross markups and the estimated value-added

markups. For complete derivation steps, see [Khan and Kim \(2013\)](#).

$$\mu^g = \frac{\mu^v}{1 + (\mu - 1)S^m} \tag{C.1.5}$$

where  $S^m$  is the ratio of intermediate input costs over total revenue, and  $\mu^g$  is the gross markups using the *wage bill* displayed in [Figure 4.5](#).