Entrant Demographics, Option to Delay Entry, and Aggregate Fluctuations

Ia Vardishvili

Georgia, Tbilisi

Master of Arts in Economics, International School of Economics at Tbilisi State University, 2013

Bachelor of Business Administration, Tbilisi State University, 2011

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia

April, 2019

Abstract

Existing firm-dynamics models that quantify the role of entry in shaping aggregate fluctuations could not account for the following two salient features of the dynamics of entrants: First, entry rate is significantly procyclical and two times as volatile as aggregate employment. Second, cohort of firms that start operating during recessions employ fewer workers at entry and over time, while they are on average more productive and have higher survival rates compared to expansionary cohorts. Given the importance of start-ups for aggregate job creation and economic growth, the dissertation investigates what accounts for the observed significant and persistent effect of the initial entry conditions on the selection of entrants over the cycles and how does the observed life cycle demographics of entrants shape the aggregate fluctuations.

In the first part of the dissertation, I show that potential entrants' ability to delay entry, previously ignored by the existing firm-dynamics literature, accounts for the significant effect of the aggregate conditions on selection of entrants. Procyclical variation in the expected survival rates moderate the relationship: during recessions increased risk of post-entry failure creates positive *value of waiting* and increases relative cost of entry today on top of fixed entry cost. The mechanism generates group of potential entrants who decide to postpone entry even if entry promises positive life-time value. I show that for reasonable parameter values (i) the total cost of entry during recessions increases by 7% for medium productivity entrants, (ii) the seemingly small increase in the entry cost generates delays that could last from 1 to 8 years. The expected duration of delay is negatively correlated with entrants' productivity level. At an aggregate level the mechanism produces a countercyclical cost of entry which leads to the documented significant selection of entrants over the cycles. I show that without the ability to delay entry, existing firm dynamics models that rely on traditional entry decision rule and fixed entry cost require counterfactually large variance of the aggregate demand shock process to reconcile the observed facts.

In the second part of the dissertation, I investigate what accounts for the observed characteristics of cohorts over the cycles. I calibrate the model developed in previous section to U.S. establishment level data. The calibrated model matches average cohorts' post-entry survival, growth and employment share. The model is able to account for the observed dynamics in entry rate and the documented persistent and significant differences in cohorts' life-cycle characteristics. I find that the option value of delay channel is quantitatively and qualitatively important for generating the results. Due to the option value of delay the opportunity cost of entry almost doubles during recessions, generating cohorts of entrants with countercyclical average productivity. The medium productivity entrants who delay entry contribute to the persistent procyclical variation

in cohort-level employment, since they represent high-growth and high-survival firms. I find that the ability to delay entry allows potential entrants to choose time of entry that leads to higher profits for longer periods of time, leading to countercyclical survival rate. A model without the option value of delay implies acyclical average survival rate.

In the third part of the dissertation, I quantify the role the observed demographics of entrants play in shaping aggregate fluctuations. I find that a calibrated model that accounts for the life-cycle dynamics of U.S. establishments produces business-cycle fluctuations in aggregate employment and output that are similar to those observed in the data. Moreover, I show that the variation in the number and the composition of entrants at entry, rather than the post-entry shocks, is responsible for generating the observed persistence and variance of aggregate variables. Using the model, I re-examine the causal relationship between the persistent and significant drop in the number of entrants and the slow recovery observed after the Great Recession. A counterfactual exercise shows that if the entry rate had stayed at the pre-crisis level, the drop in aggregate employment would have been 45 percent lower, and the economy would have recovered two times faster.

In the final part of the dissertation, I show that accounting for the option to delay entry has important policy implications and could significantly alter the existing firm-dynamics models' predictions about the response of potential entrants to various policies. In a model with the option value, a final effect of a policy depends on its effect on entrants' relative cost of entry today versus tomorrow, while the traditional entry decision rule accounts only for the direct effect of the policy. Depending on the magnitude of an indirect effect, these two specification could imply quantitatively and qualitatively different responses of potential entrants depending on the timing, duration, and the magnitude of policies. For example, I show that existing models predict that a temporary or a permanent decrease in entry cost have a same effect on the number of entrants, whereas a model with the option value predicts that temporary decrease in entry cost is more effective in increasing the number of entrants during recessions since it decreases relative cost of entry more than the permanent policy. The main takeaway is that while designing policies that intend to affect potential entrants one needs to take into account very carefully indirect effects implied by potential entrants ability to delay entry.

Contents

Acknowledgments

My sincere thanks go to my advisor, Eric Young. His invaluable guidance, encouragement, and support has made the dissertation process very interesting, productive, and enjoyable. He has generously shared his time and has assisted me to develop, distill, and communicate my ideas. The dissertation would not have been possible without his insightful comments, challenging questions, and critique that helped me to broaden my perspectives and significantly improve the quality of my research. Working with him have had a formative experience for both my personal and professional growth.

I am deeply indebted to my advisor, Toshihiko Mukoyama for his invaluable assistance at every stage of the dissertation writing process. Being part of his advising group helped me keep my progress on schedule. His detailed comments, feedback, and suggestions about the paper was crucial to advance the project to the final document. I am also extremely grateful to him for his invaluable professional advice, practical suggestions, and guidance, which extended beyond the dissertation topic and will assist me throughout my academic career.

I am grateful to Sophie Osotimehin, an excellent Professor and a mentor, for her extremely helpful recommendations, feedback and advice. I want to thank Kieran Walsh for taking time to serve as my outside committee member and for his insightful comments.

I am extremely grateful to Leland Farmer, Zach Bethune, and Jae Won Lee for their challenging questions, useful comments, and suggestions throughout the duration of the project. I want to thank my fellow graduate students: Jorge Miranda Pinto, Diego Legal Canisa, Miguel Mascarua, Jeffrey Schafer, Guillermo Hausmann, Daniel Savelle, Allison Oldham, Nick Embrey, and Ilhan Guner for insightful conversations and feedback on my research. I want to additionally thank Ilhan for being a very caring and supportive friend throughout my PhD.

I would also like to recognize the invaluable feedback that I received during my visit to the Federal Reserve Bank of St. Louis. Special thanks go to Alexander Monge-Naranjo, B. Ravikumar, Juan Sanchez, Sungki Hong, and Yongseok Shin for their useful suggestions, and conversations. I am very grateful to the participants of the workshops at the Federal Reserve Bank of St. Louis, UVA-Richmond Fed Jamboree, Midwest Macroeconomics Meetings Fall 2018, 13th Annual Economics Graduate Students Conference at Washington University in St. Louis, the 6th Lindau Nobel Prize Laureate Meeting on Economic Sciences, and the Macro Research Group at University of Virginia for their insightful comments and constructive critique. I also gratefully acknowledge the financial support that I received from the Bankard Fund for Political Economy.

Next, I would like to extend my deepest gratitude to all faculty members at the University of Virginia who helped me to challenge my intellectual capacity, to grow personally, and excel professionally. Special thanks are due to Federico Ciliberto, and Simon Anderson for their enthusiastic encouragement, support and help during my PhD studies. I believe that their positive attitude and care had played an important role in successful completion of my PhD degree. Additionally, I would also like to thank Sage Bradburn, and Patty Futrell for their constant, and patient assistance throughout my stay here.

I am very fortunate to have a great group of friends and peers in my graduate program with whom I share many unforgettable and fun memories. Special thanks go to 'Four Wind' group: Ekaterina Khmelnitskaya, Devaki Ghose, Taheya Tarannum, for their wonderful friendship and constant support. I want to additionally thank Taheya for being always very understanding and considerate while reflecting on my irrelevant or relevant concerns. I want to acknowledge Sean Sullivan, Divya Pandey, and Miguel Mascarua for their friendship, and encouragement. I also want to thank Emily Cook, Diego Legal, Alex Gross, Hundanol Kebede, Salim Ergene, Fran Murphy, Shreejaya Shrestha, Jigme Norbu, Tre Wells, Sven Hanold, Abby Teshome, Moogdho Mahzab, Ramiro Burga, Geoggrey Carr, Hanna Charankevich, Jaeki Jang, Pooja Khosla, Seunghyeon Lee, Melissa Moore, Grace Toufeili, Xintian Wang, Gizem Kutlu, Snigdha Das, Fiorella Pizzolon, and Joaquin Saldain for many unforgettable moments that will stay in my (phone) memory forever. I would also like to acknowledge each member of the Cohort 2017 for their warm friendship and encouragement. Being their Teaching Assistant taught me (in a hard way) that jokes, unless they are funny, are not funny!

I would like to extend my sincere thanks to my friends, Nikoloz Pkhakadze, Robizon Khubulashvili, Irma Lapiashvili, David Tsirekidze, John Samkharadze, and David Nikolaishvili for their constant support and encouragement. I feel extremely lucky to have a wonderful friend like Irina Timchenko during my stay in Charlottesville. I would like to thank her for her invaluable guidance and help, positive and fun character, and thoughtful comments.

Finally, I would also like to say a heartfelt thank you to my parents Iza Chikvaidze, and Soso Vardishvili, and to my siblings Nonika Vardishvili, Oliko Vardishvili, Mariam Vardishvili, and Luka Vardishvili for their unconditional love, and support.

All errors are my own.

... I dedicate this dissertation to my grandma, Lili.

1 Introduction

Existing firm-dynamics models that quantify the role of entry in shaping aggregate fluctuations do not account for the following two salient features of the dynamics of entrants: First, entry rate is significantly procyclical and two times as volatile as aggregate employment.[1] Second, cohort of firms that start operating during recessions employ fewer workers at entry and over time, while they are on average more productive and have higher survival rates compared to expansionary cohorts.[2] Despite the importance of entrants for the aggregate job creation and economic growth, we lack the theoretical framework that rationalizes and quantifies the observed significant and persistent effect of the initial entry conditions on selection of entrants over the cycles.[3]

This paper shows that the existing firm dynamics models can be reconciled with the data by allowing potential entrants to delay entry after observing the aggregate state. The feature, previously ignored by the firm dynamics literature, generates a non-negative option value of delay and significantly alters potential entrants entry decisions over the cycles.[4] The mechanism leads to a countercyclical opportunity cost of entry in equilibrium, which significantly amplifies the response of potential entrants to aggregate shocks and generates the observed different composition of entrants over the cycles. Using a calibrated model that matches the life-cycle dynamics of U.S. establishments I am able to re-examine and quantify the role of entry in shaping the business-cycle fluctuations in aggregate employment and output. Finally, I show that accounting for the option to delay entry has important policy implications and could significantly alter the existing firm-dynamics models predictions about the response of

[1] Source the Business Dynamic Statistics (BDS) dataset. Establishment level data. Period 1977-2015.

[2] Moreira (2015) and Sedlacek and Sterk (2017) document that cohort-level employment is significantly and persistently procyclical. Lee and Mukoyama (2015) and Moreira (2015) find that entrants that are born during recessionary periods are on average more productive at entry and over time. Using the BDS database over 1977-2015 period, I show that entrants that are born during recessionary periods have persistently higher survival rate, compared to cohorts that are born during expansionary periods. These facts are robust to different measures of the business cycles. These facts are also true for the firm-level data.

[3] Standard firm-dynamics models inability to account for the microeconomics life-cycle characteristics of entrants is even considered as a puzzle in Lee and Mukoyama (2018).

[4] The mechanism corresponds to the considerable theoretical microeconomics literature that dates back to Arrow (1968) and shows that the ability to delay entry significantly alters individual firms' investment decisions under aggregate volatility. Examples, Bernanke (1993), McDonald and Siegel (1986), Pindyck (1991). See Dixit and Pindyck (1994) for detailed review.

potential entrants to various policies.

Existing business cycle firm-dynamics models, based on Hopenhayn (1992) framework, could not account for the documented variation in the entry rate and the persistent differences in cohorts' life-cycle characteristics without generating excessive volatility in aggregate variables. In these models the entry decision follows a traditional neoclassical investment decision rule. Potential entrants enter the market if lifetime benefits from entry are greater than the fixed entry cost. When entry cost does not vary over the cycles, aggregate conditions affect the selection of entrants only through affecting the lifetime benefits of entry. To produce the documented significant selection of entrants, these models require large variance of the aggregate shocks, which leads to excessive variation in aggregate variables and counterfactual predictions of the role of entry. With the ability to delay entry, an aggregate shock indirectly affects the entry decision through the option value of delay from its additional affect on the future benefits of entry. Due to the latter channel, even a small shock that generates a difference between the benefits of entry today and tomorrow can have a substantial effect on pre-entry selection of entrants.

I build model on Moreira (2015), a Hopenhayn (1992)-style firm-dynamics model with endogenous firm entry and exit. In the framework, exogenous aggregate demand shocks drive business cycles. Heterogeneous firms compete in monopolistically competitive markets and make decisions about production and exit. Potential entrants hold heterogeneous signals about their post-entry initial productivity. Signals are persistent, which allows potential entrants to postpone entry without losing the signal. The feature of the model generates a non-zero option value of delay in entry decision that varies with the signal and with the aggregate demand level.[5] I use the framework to analyze and quantify the implications of the ability to delay entry.

I find that the option value of delay significantly modifies individual potential entrants entry decisions. The ability to delay entry allows potential entrants to endogenously choose

[5]If the signals have zero persistence the option value of delay equals to zero and the entry decision reduces to the traditional neoclassical decision rule (e.g. Moreira (2015), Clementi and Palazzo (2016)). A nondegenerate entry rate requires the option value of delay to equal to zero in firm-dynamics models with homogeneous potential entrants too (e.g. Bilbiie, Ghironi and Melitz (2012)).

expected post-entry survival rates. In particular, by waiting during recessions and entering during expansions potential entrants are able to get higher period profits for longer periods of time. In other words, the procyclical variation in the expected survival rates, acts as a procyclical discount factor that increases value of waiting during recessions by increasing the net benefits of entry during expansions.[6] I find that potential entrants with a medium range of the signals, whose expected survival rate varies significantly over the cycles, are the ones that benefit from the ability to delay entry. Potential entrants that hold low and high productivity signals do not find waiting beneficial since their survival rates are respectively close to zero and one.

The option value of delay channel endogenously generates a countercyclical opportunity cost of entry in equilibrium. Recessions decrease survival rates of potential entrants which increase downside risk of entry. As a result, potential entrants find it optimal to wait and enter the market when the benefits from entry is significantly higher than the fixed entry cost.[7] This feature of the model increases procyclical variation in the number of entrants compared to firm-dynamics models with fixed entry cost. Moreover, since the option value of delay mainly affects potential entrants with medium-range of the signals, the selection mechanism additionally increases the variation in the productivity composition of entrants over the cycles. I show that the latter effect significantly increases the model's internal propagation mechanism.[8]

To evaluate the quantitative importance of the mechanism, I calibrate the model using establishment level data over the period 1977-2015 from the Business Dynamic Survey (BDS) dataset. The calibrated model matches average cohorts' post-entry survival, growth, and

[6]The procyclical variation in the expected survival rates is due to endogenous exit. During the recessions, aggregate demand and as a result incumbent firms continuation value is low. As a result, the probability of observing a fixed cost that is higher than the continuation value is high. Thus, during recessions probability of exit (survival) increases (decreases).

[7]Using a calibrated model, I find that during sever recessions the opportunity cost of entry could well reach double the fixed entry cost.

[8]The mechanism is consistent with Decker et al.'s (2014) empirical findings, who show that entrant cohorts contribution to the aggregate employment comes from the small share of the high-growth firms. The mechanism corresponds to the 'missing generation' effect described in Clementi, Khan, Palazzo and Thomas (2014), Clementi and Palazzo (2016), and Siemer (2016). In the paper, I also provide an exercise that decomposes the missing generations into the variation in the low and high productivity entrants. The exercise shows that specifically the procyclical variation in the latter group of entrants are responsible for propagation of the aggregate shocks.

employment share. I calibrate the exogenous aggregate demand shock process to match the dynamics of the entry rate in the model and in the data. I find that the baseline model requires an aggregate demand shock process with a seven times lower variance to generate data-conforming variation in the entry rate compared to a model without option value of delay.

The calibrated model accounts for the documented persistent and significant differences in cohorts' life-cycle characteristics. I find that the option value of delay channel is quantitatively and qualitatively important for generating the results. Indeed, due to the option value of delay the opportunity cost of entry almost doubles during recessions relative to expansions. As a result, entrants during recessions hold significantly higher signals compared to those that enter the market during expansions. Consequently, the model produces significantly countercyclical productivity composition of entrants. The model generates the documented persistent procyclical variation in cohort-level employment due to procyclical variation in the entry rate of high productivity entrants. Moreover, consistent to the empirical findings by Sedlacek and Sterk (2017), and Pugsley, Sedlacek and Sterk (2016), I find that the majority of the procyclical variation in cohort-level employment is due to the selection at entry rather than the post-entry shocks.

The model produces a countercyclical average survival rate, which is a direct implication of the option to delay entry. Increase in average survival rate due to selection of high productive entrants dominates the negative affect of recessions on potential entrants post-entry survival rates and as a result the model generates countercyclical average survival rate. In a model without the option to delay entry, the two effects cancel each other out, leading to an acyclical average survival rate. The difference can be explained by the additional selection generated through entrants' ability to delay entry. This result provides a clear testable implication for the option value of delay mechanism. Using the Business Dynamic Survey database over the period 1997-2015, I show that cohorts that start operating during recessions have on average persistently higher survival rates compared to cohorts that enter the market during expansions. I interpret the result as providing support to the option value of delay mechanism.[9]

Using the model I investigate how the observed macro and micro life-cycle characteristics of entrants shape the dynamics of aggregate variables. The main finding is that a model that accounts for the cohorts' life-cycle demographics (selection at entry, growth, survival) produces business-cycle fluctuations in aggregate employment and output which are similar to those observed in the data. In addition, I find that the variation in the number and the composition of entrants at entry, rather than the post-entry shocks, is responsible for generating the observed persistence and variation in aggregate variables.

Next, I re-examine the amplification and propagation mechanism of the change in the number and the composition of entrants.[10] In the baseline model, variation in the composition of entrants has a long-lasting effect on the recovery of aggregate employment: After a one-time negative aggregate demand shock, the baseline economy takes 3 years to recover half-life, and an extra 12 years to recover additional 25 percent of the decline. By contrast, an economy in which the shock does not affect the entry margin, takes only 2 years to recover two-thirds of the decline.[11] Further examination shows that the significantly protracted recovery for the additional 25 percent of the decline is due to a small number of high-productive entrants that choose to delay entry as a response to the shock.

I find that if the shock is persistent, the composition effect has a substantial effect on the depth and the long-run recovery of the economic aggregates. Motivated by these findings, I study how the persistent and significant drop in number of entrants in the U.S., over the period 2007-2015, accounts for the slow recovery in the aggregate employment observed after the Great Recession. I construct an aggregate demand shock series that matches the dynamics of entrants in the data over the period 1977-2015.[12] A counterfactual exercise

[9] Bayard et al. (2018) using the newly developed Business Formation Statistics dataset that collects information about business applications and formation, show that the firms that entered into the market in year 2013 consists of firms that applied for the Employer Identification Number within the period 2004-2013. Their findings provide additional evidence for the option value of delay mechanism.

[10] Examples of the existing firm dynamics models that study the same question are Clementi, Khan, Palazzo and Thomas (2014), Clementi and Palazzo (2016), Siemer (2016).

[11] The mechanism is consistent to empirical findings by Gourio, Messer and Siemer (2016). Using an annual panel of US states over the period 1982-2014, they show that change in number of entrant firms have a persistent effect on the dynamics of the aggregate variables.

[12] Clementi and Palazzo (2016), Gourio, Messer and Siemer(2016) and Sedlacek and Sterk (2017) show

shows that, if the number and the composition of entrants had stayed at the pre-crisis level, the drop in aggregate employment would have been 45 percent smaller, and the economy would have recovered two times faster.[13]

In the last section, I show that accounting for the option to delay entry has important policy implications and could significantly alter the existing firm-dynamics models predictions about the response of potential entrants to various policies. The reason is that the ability to delay entry implies qualitatively and quantitatively different responses of potential entrants to a change in the expected value of entry depending on the timing, duration, and the magnitude of the change. For example, permanent decrease in entry barriers does not affect potential entrants that choose to delay entry during recessionary periods, since they will benefit from the lower entry cost in the future. However, if the decline in entry cost is temporary then potential entrants choose to enter right away. The existing firm dynamics models predict the same response of potential entrants to the temporary and permanent decrease in entry barriers. Besides, due to the significantly countercyclical opportunity cost of entry, I find that the marginal entrants who respond to the reduction of fixed entry cost are mostly high productive entrants during recessions and low productive firms during expansions. Furthermore, I show that the additional, option value of delay channel significantly shapes the dynamics of the economy as a response of the news shocks.[14] The main takeaway of the exercises is that the option value of delay is very important for understanding the response of potential entrants to various policies. And, any policy that intend to affect potential entrants should very carefully take into account the channel.

---

that during the Great Recession entry was affected more at extensive margin (number of entrants) rather than the intensive margin (average size of entrants relative to incumbent firms). Secondly, using cross-state (cross-MSA) data Gourio, Messer and Siemer (2015) show that the states (MSAs) that experienced higher drop in entry rate over the period 2008-2010 experienced slower recovery over the period 2010-2013, even after controlling for house prices and leverage.

[13]The result also indicates that the combination of the persistently low aggregate demand and the observed entrant demographics are not sufficient for explaining the depth of the Great Recession.

[14]Constantini and Melitz (2007) also show that potential entrants respond differently to the news about trade liberalization depending on the timing and the implementation of the policy. However, with the option value of delay the news shock also affects relative profits that potential entrants can get by entering the market in the future. Thus, analyzing the relative variation in the benefits today versus tomorrow is crucial to quantify the news affect on firm-dynamics.

1.1 Contributions

The paper relates and contributes to the several areas of the research. *First,* the paper contributes to the theoretical literature that studies macro and micro life cycle characteristics of entrants over the business cycles. Lee and Mukoyama (2015) show that average size and average productivity of U.S. manufacturing plants are significantly countercyclical. Lee and Mukoyama (2018) show that generating the documented significant selection of entrants in Hopenhayn and Rogerson (1993) framework is a puzzle. They show that introducing entry cost that varies over the cycles in a particular way can solve the puzzle. I contribute to the literature by proposing a mechanism that microfounds the variaiton in entry cost. In particular, the option value of delay endogenously generates countercyclical opportunity cost of entry in equilibrium, which significantly amplifies the effect of the aggregate shocks and generates documented selection of entrants over the cycles.

Sedlacek and Sterk (2017) and Moreira (2015) use demand side factors to explain the persistent procyclical variation in cohorts' employment. Moreira (2015) accounts for the fact through persistent demand dynamics. The mechanism hinders entrants that start operating during recessions to catch up to cohorts that start operating during expansions. I build model on Moreira's (2015) framework and find that accounting for the documented significant selection of entrants decreases the role persistent demand dynamics play on variation in cohort-level employment over the cycles. In turn, the ability to delay entry increases procyclical variation in high productivity firms and contributes to persistent procyclical variation in cohort-level employment. The mechanism is related to Sedlacek and Sterk's (2017) theoretical explanation. They allow potential entrants to choose between high-growth and low-growth business profiles at entry. The pro-cyclical variation in the number of entrants with high-growth business profiles generates persistent differences in cohort-level employment. Additionally, I find that the selection of potential entrants only through productivity is not enough to account for the procyclical variation in average size that is documented in Sedlacek and Sterk's (2017).[15]

[15] Counteryclical average size generated in the baseline model is in line to Lee and Mukoyama (2018). I expect that further extending the model to account for the procyclical average size at entry will increase the difference in cohort-level employment over the business cycles. Indicating that the quantitative results of the project represents the lower bound of the contribution of entry to aggregate dynamics.

Second, the paper is closely related to a large body of theoretical literature that studies the role of entry and exit in the amplification and propagation of aggregate shocks.[16] I contribute to the literature by building a model that accounts for cohorts' average life-cycle characteristics, matches to the observed dynamics of the entry rate over the business cycles, and also generates the documented persistent differences in cohorts' characteristics that are born at different cycle states of the economy. Using the model, I re-examine and quantify the role of the entry and exit on the dynamics of the aggregate employment and output.[17] I find that the variation in the number and the composition of entrants explains very closely the business cycle dynamics in aggregate employment and output. Additionally, I show that not accounting for these facts significantly weakens the effect of entry on the propagation of the aggregate shocks.[18]

Third, The paper also contributes to the theoretical literature that studies the causal relationship between the significant and the persistent drop in entry rate and the slow recovery in aggregate employment observed after the Great Recession. The paper is most closely related to Clementi and Palazzo (2016) and Siemer (2016). They develop a firm-dynamics model with endogenous entry and exit and show that a one-time decrease in the number of entrants generates a long-lasting recession due to a missing generation of entrants.[19] Using an empirically sensible aggregate demand shock series and a model that generates data-conforming life cycle dynamics of entrants I am able to fully quantify the role persistent drop in entry rate over the period 2009-2015 played in aggregate dynamics after the Great Recession.

Fourth, the paper relates to a considerable amount of theoretical work on irreversible invest-

[16]For example, Samaniego (2008), Lee and Mukoyama (2008), Bilbiie, Ghironi and Melitz (2012), Clementi and Palazzo (2016) and others.

[17]Haltiwanger et al.(2013) find that young firms exhibit distinct life-cycle dynamics compared to their mature counterparts and emphasize the importance of accounting for not only the entry process but the subsequent post-entry dynamics (growth, survival, job creation) to quantify the contribution of entry to aggregate dynamics.

[18]For example, Samaniego (2008) predicts that the dynamics of aggregate output are insensitive to variations in entry and exit, in a model where the entry and exit are not sensitive to an aggregate demand shock process.

[19]In particular, Siemer (2016) develops a firm-dynamics model where entrants are required to finance a fraction of the fixed entry cost through debt. To generate data-conforming drop in the number of entrants he uses a counterfactually large financial shock. Clementi and Palazzo (2016) considers impulse response to the drop in the number of entrants as a response to a negative aggregate productivity shock and exogenous entry specific shock. Moreira (2015) considers a one time aggregate demand shock to quantify the role of entry dynamics.

ment decision under aggregate uncertainty that dates back to Arrow (1968).[20] The study shows that the ability to choose timing of investment has a substantial effect on investment decision. Veracierto (2002) studies aggregate consequences of investment irreversibility using a real-business-cycle model with exogenously determined entry and exit and finds that investment irreversibility has no effect on aggregate fluctuations. I contribute to the literature by extending the analyzes on an entry margin.[21] I show that entrants ability to delay entry under aggregate state volatility substantially modifies individual entry decision and significantly amplifies and propagates business cycle fluctuations. In particular, to study aggregate consequences of the option value of delay I use a model that is calibrated to U.S. data and closely accounts for the business cycle dynamics in aggregate employment. I find that shutting down the option value of delay channel decreases variance of aggregate employment by three times. At the same time, persistence of aggregate employment decreases from 0.58 to 0.43.

The paper also relates to recent work by Fajgelbaum, Schaal and Taschereau-Dumouche (2017). They show that the ability to delay entry can generate endogenous uncertainty in a framework where entrants strategically choose to wait so that they can learn from actions of others. In my paper, entrants find beneficial to delay entry even if there is no strategic interaction between firms. In particular, the volatility in aggregate state that affects procyclical variation in entrants' survival rates acts as a procyclical discount factor. The feature of the model increases net benefits of entry during expansions and consequently increase value of waiting during recessions.

Fifth, the paper also relates to the literature that points out the weak internal propagation mechanism of standard business cycle models (Cogley and Nason (1995), King and Rebelo (1999)). I find that augmenting entry decision with the option value of delay significantly decreases reliance of the model to external dynamics. For example, I find that a model without option value of delay requires first order autocorrelation in aggregate demand shocks

[20] Bernanke (1993), McDonald and Siegel (1986), Pindyck (1991). See Dixit and Pindyck (1994) for detailed review.

[21] In the model the decision to enter into the market is an irreversible in the following sense: once the potential entrant enters into the market, the potential entrant loses the initial signal and do not get another chance to make entry decision again.

as high as 0.80 to generate the data-conforming persistence in aggregate employment. While a model with the option value of delay it only needs to be 0.57.

The rest of the paper is organized as follows. Section 5.2.3 documents differences in recessionary and expansionary cohorts' survival rates. Section 2 and Section 3 describe the model and explain entry decision with persistent signals. Section 5 outlines the calibration procedure and shows model fit. Section 5 studies performance of the standard firm-dynamics models. Section 7 studies propagation of the aggregate shocks through entry rate. Section 8 studies implications of option value of delay on entrant dynamics. Section 9 concludes.

2 The model

The model follows closely to a partial equilibrium, firm-dynamics model developed in Moreira (2015). The model features endogenous firm entry and exit in the style of Hopenhayn (1992). Exogenous aggregate demand shocks that affect firms' profitability and selection of entrants are the only source that drive business cycles in the model. The economy consists of incumbent firms and potential entrants. Incumbent firms produce differentiated products and are heterogeneous over idiosyncratic productivity and customer capital. They make decisions about production and exit. Potential entrants hold heterogeneous signals about their initial post-entry productivity. I deviate from the original framework by allowing potential entrants to keep the signals over time until they enter the market. The modification gives potential entrants the option to delay entry in the future after observing the aggregate state. Latter I show that due to the mechanism the aggregate conditions at entry have a significant effect on the selection of potential entrants at entry. Description of the model is given below.

2.1 Setup

Time is discrete and is indexed by $t = 1, 2, ...$ Agents face an infinite horizon. There are two types of agents: incumbent firms and potential entrants.

2.1.1 Technology

At time t, a positive measure of heterogeneous firms produce differentiated products on a monopolistically competitive market using the following production function:

$$y_{it} = s_{it} n_{it}$$

The production function is linear in labor n_{it}. Labor supply is infinitely elastic. Wage is exogenous and constant. s_{it} is a time-varying idiosyncratic productivity specific to a firm i and evolves according to a persistent $AR(1)$ process:

$$log(s_{i,t}) = \rho_s log(s_{it-1}) + \sigma_s \varepsilon_{it}$$

where $\varepsilon_{i,t} \sim i.i.d.\ N(0,1)$ for all $t \geq 0$. Idiosyncratic productivity is distributed independently across firms.

Every period, firms that are operating in the market incur fixed cost $c_f > 0$, drawn from a time-invariant distribution $c_f \sim G(c_f)$.

2.1.2 Demand

In each period, demand for firm i's differentiated good is determined according to the following isoelastic demand function:

$$y_{it} = p_{it}^{-\rho} b_{it}^{\eta} \alpha z_t$$

where p_{it} is price set by firm i and $\rho > 1$ is the price elasticity of demand. $\eta \in (0,1)$ measures the elasticity of demand with respect to customer capital b_{it}, which evolves according to:

$$b_{it} = \begin{cases} (1-\delta) b_{it-1} + (1-\delta) p_{it-1} y_{it-1} & \text{if firm } i \text{ produced in } t\text{-1} \\ b_0 & \text{entrant firm} \end{cases}$$

where b_0 is the initial level of customer capital, common across all entrants. $\delta \in (0,1)$ is the depreciation rate of customer capital. Process of customer capital which is tied to past sales hinders firms' ability to freely adjust their demand over time, which creates persistence in the dynamics of production and employment. z_t represents common aggregate demand shock that evolves as a persistent $AR(1)$ process,

$$log(z_t) = \rho_z log(z_{t-1}) + \sigma_z \epsilon_t$$

where $\epsilon_t \sim i.i.d.N(0,1)$ for all $t \geq 0$. $\alpha > 0$ is a scale factor.

2.1.3 Dynamics

At the beginning of every period, the market consists of a heterogeneous mass of incumbent firms who make production and exit decisions. The mass of incumbent firms are distributed over the different level of customer capital and idiosyncratic productivity. The distribution function is given by $\Omega_t(s,b)$.

At the beginning of every period, there is a constant mass of potential entrants M. Potential entrants are endowed with heterogeneous signals q about their first period idiosyncratic productivity. For a given signal q the distribution of the initial period productivity is given by $H_e(s|q)$ and the distribution decrease with the signal q. The signal is persistent over time, which gives potential entrants the ability to keep the signals and enter the market in the future instead of today. The aggregate distribution of potential entrants over signals is time-invariant and is given by $W(q)$.[22]

At the beginning of every period $t > 0$ the vector of the aggregate state variables is given by $\Gamma_t = \{\ z_t, \Omega_t(b,s), W(q)\}$.

Given the setup, the incumbent firm's and the potential entrant's problem at time t are described below. The summary of the incumbent firm's timing is illustrated in Figure 5. Figure 6 describes the potential entrant firm's timing.

[22]Underling the restriction is an assumption that there is a restricted number of business ideas that can be implemented in the market in each period. Appendix A.1.1 provides more detailed description of the entry phase that justifies the constant mass of potential entrants in a framework where entrants are able to delay entry. In Appendix A.1.2 I show that main results of the paper is robust if I extend the model and allow accumulation of potential entrants over time. In fact, I find that allowing accumulation of potential entrants amplifies the effect of the aggregate state on the selection of potential entrants.

2.2 Incumbent firms

At the beginning of each period, an incumbent firm i, with predetermined customer capital b_{it}, observes aggregate demand shock z_t, and idiosyncratic productivity s_{it}. Using the information, the incumbent firm makes decisions about the optimal production level, price and next period's customer capital. At the end of the period the incumbent firm draws fixed cost c_f and makes the continuation decision. Even if the firm decides to stay in the market, it may be hit by a random exit shock with probability $\gamma \in (0,1)$. The outside value is normalized to zero.[23] Firms discount future profits at the time-invariant factor β.

The incumbent firm solves the following functional equation:

$$V^I(b,s,z) = \max_{y,p,b'} \quad \left(p - \frac{w}{s}\right) y + \int max \left\{0, -c_f + \beta(1-\gamma)E[V^I(b',s',z')|s,z]\right\} dG(c_f)$$

$$\text{s.t.} \quad b' = (1-\delta)(b+y)$$

$$y = \alpha p^{-\rho} b^{\eta} z$$

2.3 Potential entrants

Each potential entrant with a signal q, observes an aggregate state of the economy (z_t) and makes an entry decision. Entry into the market is subject to a fixed entry cost c_e.

The value that the potential entrant expects to receive from entering the market equals to the expected value of being an incumbent at time t, $\int\limits_s V^I(b_0, s, z_t) dH_e(s|q)$ minus the fixed cost of entry, c_e. Upon entry the potential entrant observes actual idiosyncratic productivity (s), receives initial customer capital stock (b_0), and behaves like an incumbent with (b_0, s) state variables.

The potential entrant has the option to postpone entry until next period. In that case, with probability $\tau \in [0,1]$, the potential entrant gets the same signal, observes new aggregate state, and makes an entry decision in $t+1$. With probability $1-\tau$ the potential entrant loses the signal, and obtains outside option value.

[23] Assume that, if incumbent firm decides to exit from the market the probability that the firm gets initial productivity signal and becomes a potential entrant again is zero.

An entry decision is summarized with the following functional equation:

$$V^e(b_0, q, z) = max\left\{\tau\ \beta E[V^e(b_0, q, z')|z],\ \ -c_e + \int V^I(b_0, s, z)dH_e(s|q)\right\}$$

Where

$$\beta E[V^e(b_0, q, z')|z] = \int_{z'} V^e(b_0, q, z')\pi(z'|z)dz'$$

Note, that when the signal has zero persistence ($\tau = 0$), the option value of delay equals 0. In that case, all potential entrants with positive net expected value of being an incumbent choose to enter the market. This case corresponds to standard firm-dynamics models with a fixed entry cost.

2.4 Recursive competitive equilibrium

For a given Γ_0, a recursive equilibrium consists of: (i) Value functions $V^I(b, s, z)$, $V^e(q, z)$, (ii) Policy functions $y(b, s, z)$, $p(b, s, z)$, $n(b, s, z)$ and $b'(b, s, z)$, (iii) Distribution of operating firms $\{\Omega_t\}_{t=1}^{\infty}$, such that:

1. $V^I(b, s, z)$, $y(b, s, z)$, $p(b, s, z)$, $n(b, s, z)$ and $b'(b, s, z)$ solves incumbent's problem.
2. $V^e(q, z)$ solves the entrant problem.

3 Entry decision, the option to delay entry and aggregate risk

3.1 The traditional entry decision rule

I start analyzing entry decision of a potential entrant that is endowed with a signal q. The signal provides the potential entrant a noisy information about the post-entry initial productivity. Once the potential entrant decides to enter the market, the potential entrant receives the first period productivity and loses the signal, as well as the option to use the signal for re-entry in the future. Therefore, the potential entrant's decision to exercise the signal by entering the market is irreversible. The expected net benefits(NPV) from exercising the signal equal to the gross value of entry minus fixed entry cost. The aggregate demand shocks affect the net expected benefits of entry only through effecting gross value of entry, since entry cost is fixed over the cycles.

Proposition 3.1.1. *(The properties of the net benefits of entry)*

(a) For given aggregate demand level z, NPV(q,z) strictly increases with the signal q.

(b) For given signal q, NPV(q,z) strictly increases with the aggregate demand level z.

Proof. See Appendix B.1.1. □

Initially, I study an optimal entry decision of a potential entrant when the signal have zero persistence $\tau = 0$. In this scenario, if a potential entrant does not enter the market, the potential entrant loses the signal and drops out from the pool of potential entrants.[24] Therefore, the potential entrant always chooses to enter the market when the gross value of entry is greater than or equal to fixed entry cost. In that regard, when the signal has zero persistence the entry decision coincides to the traditional, neo-classical investment decision rule, according to which investment should be made when the net present value of investment is

[24]Note that if $\tau = 0$ potential entrants have the ability to delay entry however with the following two limitations: they receive 0 from that period on and they do not get to make entry decision in the future. The delay that I am after in the dissertation provides a potential entrants the ability to make a similar entry decision tomorrow. Thus, modeling entry decision as two options: enter or drop out when $\tau = 0$ enables me to refer the ability to delay entry only to the cases when $\tau > 0$.

nonnegative.

$$\textit{Enter (exercise the signal q) if} \quad : \quad NPV(z,q) \geq 0 \tag{1}$$

Using the individual entry decision described in Equation 1 and properties of the net benefits of entry, I characterize the group of potential entrants that enter the market when the aggregate demand is z. Since all entrants receive the same level of customer capital b_0 and observe the same aggregate demand level z, the selection of entrants depends only on the level of signal q: potential entrants with signal q which satisfies Equation 1 are going to enter the market, while the rest are going to stay outside the market.

To characterize selection of potential entrants across aggregate states define a signal $q^*_{\tau=0}(z)$ such that $NPV(z,\ q^*_{\tau=0}(z)) = 0$. The subscript $\tau = 0$ refers to a scenario when the signal has zero persistence. The threshold signal does not exist in the following two cases: a) if the aggregate demand level z is such that $NPV(z,q) < 0$ for all q. This is a case when nobody enters the market. b) if the aggregate demand level z is such that $NPV(z,q) > 0$ for all q, in which case everybody finds it profitable to enter the market. In all other cases, $q^*_{\tau=0}(z)$ exist and is unique, which is due to the fact that the net benefits of entry are strictly increasing function of the signal q.

Proposition 3.1.2. *Suppose for an aggregate demand level z there exist a signal $q^*_{\tau=0}(z)$ such that $NPV(z,\ q^*_{\tau=0}(z)) = 0$, then all potential entrants with $q \geq q^*_{\tau=0}(z)$ decide to enter the market, while the rest decide to stay outside.*

Proof. Gross value of entry and hence the net benefits from entry is a strictly increasing function of a signal q. As a result, for any $q \geq q^*_{\tau=0}(z),\ \ NPV(z,q) > NPV(z,\ q^*_{\tau=0}(z)) = 0$. Thus, potential entrants who hold signals from the range $q \geq q^*_{\tau=0}(z)$ decide to enter the market. For any $q < q^*_{\tau=0}(z),\ \ NPV(z,q) < NPV(z,\ q^*_{\tau=0}(z)) = 0$. Potential entrants that hold signals from the range $q < q^*_{\tau=0}(z)$ do not enter the market. □

In the rest of the paper, I refer $q^*_{\tau=0}(z)$ as a *threshold signal* when the aggregate demand level is z. One of the main goals of the paper is to understand what drives the change in the composition of entrants over the cycles, thus I concentrate on describing cases where entry has an interior solution (the threshold signal exist). Figure 9(a) and Figure 9(b) display gross value of entry (solid red line) and fixed entry cost (solid horizontal line) for different aggregate demand levels. Potential entrants with the signal for which gross value of entry is higher than the fixed entry cost enter the market. In each figure the intersection of the two lines indicates the threshold signal level.

Proposition 3.1.3. *The threshold signal $q^*_{\tau=0}(z)$ is countercyclical.*

Proof. According to Proposition 3.1.1(a) $NPV(q, z)$ strictly increases with the aggregate demand level. The higher the aggregate demand level z, the lower the required signal level that ensures nonnegative net benefits from entry. As a result, the minimum signal $q^*_{\tau=0}(z)$ required for a potential entrant to enter the the market decreases with the aggregate demand level. Blue dashed line on Figure 9 displays the threshold signal $q^*_{\tau=0}(z)$ for each aggregate demand level. The figure illustrates that the threshold signal is countercyclical. □

Countercyclical threshold signal leads to endogenous variation in the productivity distribution of entrants over the cycles: during the low aggregate demand periods potential entrants that enter the market hold relatively higher level signals compared to the group of potential entrants that start operating during high aggregate demand periods. Besides, since the distribution of potential entrants across the signal is fixed, countercyclical threshold signal also implies procyclical variation in the number of entrants. Hence, accounting for the empirically documented significant effect of the initial aggregate conditions on the composition and the number of entrants requires high elasticity of the threshold signal with respect to aggregate demand level.

When entry decision follows Equation 1 and the sunk entry cost is fixed over the cycles, the aggregate demand affects threshold signal and hence, selection of entrants only through its direct effect on potential entrants life-time stream of profits (expected value of being an incumbent). Meaning that increasing the elasticity of the threshold signal with respect to the aggregate demand level requires significant variation in the net benefits of entry with the aggregate demand level. However, the effect of the mean-reverting aggregate demand shocks on the life-time profits is limited unless the magnitude of the shocks are large. Increasing the variance of the aggregate demand shocks increases variance of incumbent firms' employment and output and leads to excessive variance in aggregate variables. Consistent to the claim, in Section 6.2, I am going to show that existing business cycle firm-dynamics models that employ traditional entry decision rule find it challenging to account for the documented variation in the entry rate and the persistent differences in cohorts' life-cycle characteristics without generating excessive volatility in aggregate variables.

Besides, the correlation between the variation in the expected life-time profits and probability of entry has been found to be weak in the data. For example, Geroski (1995) summarizes stylized facts about entry and points out that the existing conventional measure of entry decision, where entry depends on expected post entry profits and entry barriers does not explain much of a variation in entry rate over time.[25] Moreover, the variation in expected stream of profits over time are found to be minor.

With the above described shortcomings in mind I am going to study entry decision of potential entrants that allows potential entrants to delay entry. With the ability to delay entry an aggregate shock affects the entry decision in two ways. First, the aggregate shock directly affects benefits from entry today. Second, the shock affects the future benefits of entry, which in turn indirectly affects the entry decision through the option value of delay. Due to the latter channel, even a small shock that generates a difference between the benefits of entry today and entry tomorrow could have a substantial effect on pre-entry selection of entrants. The latter feature significantly increases the elasticity of threshold signal with respect to aggregate state and enables existing firm-dynamics models to account for the observed cyclical variation in the number and the composition of entrants without generating excessive volatility in aggregate variables.

[25] Even more, for example O'Brien, Folta and Johnson (2003) use 1993 National Survey of Small Business Finance(NSSBF) and the 1998 Survey of Small Business Finance(SSBF) and find that the probability of entrepreneur entry in the market decreases with the increase in the current rate of industry profitability. The latter value proxies the present value of future cash flows and is measured by the total operating profit divided by total sales for the median business segment in each industry.

3.2 Entry decision with the ability to delay entry

To study entry decision with the ability to delay entry I endow persistent signals to potential entrants. Persistent signal provides potential entrants an option to keep the signal and enter the market in the future instead of today.[26] Decision about exercising the signal today and exercising the signal in the future are mutually exclusive options. By entering the market today potential entrants incur sunk entry cost and also give up the net present value from exercising the signal in the future, later referred as *Option Value of Delay*. Hence, the ability to delay entry alters the traditional decision rule of entry described in Equation 1. Now, the potential entrant requires not only the NPV to be greater than zero but also greater than the option value of delay.[27]

$$\textit{Enter (exercise the signal q) if} \quad : \quad NPV(z,q) \geq \textit{Option Value of Delay}(z,q) \tag{2}$$

Figure 7 displays *Option Value of Delay* (z,q) across the signal q and for different aggregate state z. The figure illustrates properties of the option value of delay that are formally described in Proposition 3.2.1 below.

Proposition 3.2.1. *(The properties of the option value of delay)*

(a) Option Value of Delay(q,z) *is non-negative for all* q *and* z.

(b) For a given aggregate demand level z, *Option Value of Delay*(q,z) *is a weakly increasing function of the signal* q.

(c) For a given signal q, *Option Value of Delay*(q,z) *weakly increases with the aggregate demand level* z.

Proof. See Appendix B.1.2. □

Intuition behind the properties of the option value of delay are as follows. In the model,

[26] Meaning that if a potential entrant with signal q decides not to enter the market today, the potential entrant gets the same signal q tomorrow with probability $\tau = 1$. In comparative statics I provide discussion how $\tau \in (0,1)$ effects the entry decision and modifies the discussion below.

[27] The modified entry decision corresponds to the investment decision rule under uncertainty, when investors have the ability to choose time of investment. The analysis of the decision rule dates back to Arrow (1968). Bernanke (1993), McDonald and Siegel (1986), Pindyck (1991).

every time potential entrants make entry decisions they have the option to stay outside the market and receive outside option value, which equals to zero. Thus, the expected value of exercising the signal tomorrow can not be negative. Option value of delay is a net present value of life-time profits if the signal is exercised in the future instead of today. As a result, the higher the signal, the higher the value that each potential entrant expects to get in the future. The value of the option to exercise the signal in the future depends on the expected aggregate state tomorrow, which increases with the today's aggregate demand level. Thus, the value of the option increases with the aggregate demand level.

Using the properties of the option value of delay and numerical solution, I study how the increased opportunity cost of entry modifies the selection of potential entrants across aggregate demand levels.

Result 3.1. *Suppose for an aggregate demand level z, there exist a signal $q^*_{\tau=1}(z)$ such that*

$$NPV(z,\ q^*_{\tau=1}(z)) = \textit{Option Value of Delay}(z, q^*_{\tau=1}(z))$$

then all potential entrants with $q \geq q^*_{\tau=1}(z)$ decide to enter the market, while the rest decide to delay entry the market.

Examine Figure 8 that displays the gross value of entry and the option value of delay for each signal level for the lowest ($z_{low} < 1$) and for the highest ($z_{high} > 1$) aggregate demand levels.[28] First, study the selection of potential entrants during the lowest aggregate demand level, illustrated on Figure 9(a). With the ability to delay entry, potential entrants that enter the market are those for whom the gross value of entry is greater than the total opportunity cost of entry. With the ability to delay entry the latter value equals to the option value of delay plus fixed entry cost. Note that potential entrants that decide to enter the market are the ones that hold signals higher than the signal $q^*_{\tau=1}(z_{low})$ for which the gross value of entry equals to the total opportunity cost of entry. All potential entrants that hold signal less than the signal $q^*_{\tau=1}(z_{low})$ does not enter the market. In other words, in the model where the persistence of the signal τ equals to 1, there exist a threshold signal $q^*_{\tau=1}(z)$ equivalent

[28]The latter values are determined within the numerical solution. For given variance and persistence of the aggregate demand shock process I find gird points and transition matrix for the aggregate demand shock process using the Rouwenhorst method. I choose the aggregate demand shock process to match the variance and the persistence of entry rate in the model to the data counterpart.

to the case of $\tau = 0$ which we can use to describe selection of entrants at each aggregate demand level.

Potential entrants who hold signals that are lower than the threshold signal do not enter the market. I decompose the latter group of potential entrants into two groups: The first group of potential entrants are ones that hold signals from the following range: $q \in [\underline{q},\ q^*_{\tau=0}(z_{low})]$. For the group of potential entrants expected value of being an incumbent is less than the fixed entry cost, they earn negative net present value if they enter the market. Note, that this are the group who stays outside the market when they do not have the ability to keep their signals over time.[29] Aggregate demand affects the group of entrants through directly affecting profits from entry today. Later I refer the selection as a *direct effect* of the aggregate demand. Second group of potential entrants hold signals from the following range: $q \in [q^*_{\tau=0}(z_{low}),\ q^*_{\tau=1}(z_{low})]$. For the potential entrants net present value of entry today is greater than zero, thus without the ability to delay entry they decide to enter the market right away. However, with the ability to delay entry aggregate demand today also affects indirectly expected net present value of benefits from exercising the signal tomorrow. Due to the *indirect effect* this group of entrants decide to wait for the future. Thus, the ability to delay entry generates additional group of potential entrants with the medium range of signals that decide not to enter the market.

Next consider the other extreme, selection of potential entrants during the highest aggregate demand period, illustrated on Figure 9(b). During the peak waiting is not optimal, since observing higher aggregate demand level tomorrow is highly unlikely and everybody who expects non-negative profits is going to enter the market immediately. Thus, the threshold signal during the peak is the same with or without persistent signal and the *indirect effect* of the aggregate state on selection of potential entrants equal to zero.

Using the numerical solution I examine selection of entrants for the rest of the aggregate demand levels and I find that the threshold signal $q^*_{\tau=1}(z)$ exist and with the increase in the aggregate demand level the threshold signal decreases. Also, I find that with the increase in

[29] With the ability to delay entry waiting is a dominant strategy for the potential entrants that incur negative value from entry. while they also do not enter market today, they might still use the option to keep signal and wait for better aggregate demand conditions.

the aggregate demand level selection of potential entrants due to *indirect effect* $(q^*_{\tau=1}(z) - q^*_{\tau=0}(z))$ becomes less and less significant. Figure 9 displays threshold signals for $\tau = 1$ and $\tau = 0$.[30]

Result 3.2. *The signal level $q^*_{\tau=1}(z)$ is countercyclical and the difference $(q^*_{\tau=1}(z) - q^*_{\tau=0}(z))$ weakly increases with the decrease in the aggregate demand level.*

The ability to delay entry ($\tau = 1$) significantly increases elasticity of entrants with respect to the aggregate demand level compared to the case without persistent signal due to the following argument: Signal that generates expected profits that covers just fixed entry cost is countercyclical (see Proposition 3.1.2). Moreover, potential entrants with higher signals have higher option value of delay and thus higher opportunity cost of entry. As a result, when the aggregate demand level decreases, required signal to cover the total opportunity cost of entry increases.

Finally, I study the minimum level of benefits that potential entrants require to enter the market for each aggregate demand level z, when they have the ability to delay entry. Toward the goal, for each threshold signal $q^*_{\tau=1}(z)$ described in Figure 9, I compute the total opportunity cost of entry. The latter value equals to the fixed entry cost plus the option value of delay of a potential entrant with signal $q^*_{\tau=1}(z)$. The total opportunity cost of entry turns out to be also threshold opportunity cost of entry. All potential entrants with expected entry value higher than the threshold opportunity cost enters the market. While all potential entrants with entry value less than the threshold opportunity cost decide to delay entry.

Figure 10 shows that if the signal is persistent, the threshold opportunity cost is countercyclical. This figure implies that the lower the aggregate demand level, the higher the total opportunity cost of entry. In contrast, if the signal is not persistent, the value equals to fixed entry cost in all aggregate states. Thus, comparing the two models elucidates the mechanism how the ability to delay entry increases elasticity of entrants with respect to aggregate demand level and generates significant selection of entrants at entry, relative to the models without the ability to delay entry. This result corresponds to Lee and Mukoyama (2018)

[30]Note, that for $\tau \in (0, 1)$ the threshold signal is in between these two scenarios, since the option value of delay weakly increases with τ.

who show that variation in the entry cost in particular manner is important to generate the documented significant selection at entry in the existing models.

For reasonable parameter values, I find that the ability to delay entry implies that during the recessionary periods potential entrants decide to postpone entry until the benefits from exercising the signal is double the fixed entry cost. Meanwhile during the higher aggregate demand periods, the cost reduces to the fixed cost of entry.

3.3 Magnification of the entry cost: the value of waiting

The *Option Value of Delay* represents the net present value of benefits associated with entering the market in the future, while NPV represents net present value of benefits a potential entrant receives by entering today. Since exercising the signal today and in the future is mutually exclusive options, by entering today the potential entrant receives NPV but gives up the value associated with the option to exercise the signal in the future. Thus, the ability to delay entry weakly increases opportunity cost of entering for all potential entrants by the amount of the *Option Value of Delay*. However, note that the increased opportunity cost of entry does not necessarily translates into cost of entry since some of the profits that the potential entrant expects to get from entry in the future can still be recovered by entering the market today. The actual cost/benefit that the potential entrants incur by entering the market today instead of waiting the future is the change in the net present value of benefits that potential entrants earn from these two options.[31] The value, later referred as *Value of Waiting*, is given in the following equation.

$$\textit{Value of Waiting(q,z)} = \textit{Option Value of Delay(q,z)} \text{ - } \textit{NPV(q,z)} \tag{3}$$

Figure 11 illustrates value of waiting for potential entrants with six different signal levels ($q_1 < q_2 < ... < q_6$) across all aggregate demand. The figure shows that for given signal level q, the *Value of Waiting(q,z)* decreases with the aggregate demand level: the higher the aggregate demand level, the lower the gain from waiting. Note also that for given aggregate

[31] The definition is in line with Pindyck (2009) defines value of waiting as the reduction in NPV by investing today instead of waiting.

demand level z, the *Value of Waiting(q,z)* decreases with the signal level: the lower the signal the lower the net present value from entry today and the higher the rewards from entering during the higher aggregate demand levels.[32]

Positive *Value of Waiting(q,z)* represents the amount of net present value of benefits that potential entrants with signal q gives up by entering the market today instead of exercising the signal in the future. Hence, positive *Value of Waiting(q,z)* increases cost of entry on top of fixed entry cost. Negative *Value of Waiting(q,z)* represents the loss (in terms of the net present value of benefits) due to waiting and not entering today. However, even if the value is negative it does not decrease cost of entry since its not something that potential entrants actually receive by entering today. Thus, in the latter scenario cost of entry coincides the fixed entry cost.

$$\textit{Total Cost of Entry} = 1\left\{\textit{Value of Waiting(q,z)} > 0\right\} + c_e \tag{4}$$

Figure 12 illustrates the *Total Cost of Entry* for the same potential entrants as considered on Figure 11. We see that for some of the signal q total cost of entry increases above fixed cost of entry and the *Total Cost of Entry* increases with the decrease in the aggregate demand level. Meaning, that during recessions, increased risk associated with post-entry profits increase cost of investing today.[33] Following Pindyck (2009) one can think about the *Total Cost of Entry* as 'full', 'risk adjusted' cost of entry.[34]

[32] Note that if for a potential entrant with signal q entering at aggregate demand level z results negative net benefits from entry $NPV < 0$, then it increases above 0 *Value of Waiting* (even if the option value of delay equals to 0). Interpretation of positive *Value of Waiting* still stays the same: entering at an aggregate demand level z not only results reduction in net present value due to investment at sub optimal time but also incurs actual cost, since fixed cost of entry is higher than net benefits from entry. For example, consider signal q_1 on Figure 11. Value of waiting is always positive for the considered aggregate demand levels, thus the potential entrant does not enter for the given range of aggregate demand levels.

[33] One can also think about the effect of the aggregate demand on the relative cost of investing today versus tomorrow in the following manner. What is the cost of getting same NPV today versus tomorrow? The cost of exercising the signal in the future equals to ($c_e -$ *Value of Waiting*), while the cost of investing today equals to c_e. If the *Value of Waiting* is positive, the relative cost of investing today versus tomorrow equals to $\frac{c_e}{c_e - \textit{Value of Waiting}} > 1$. Meaning, that increased risk associated with post-entry returns during recessions increases relative cost of investing today versus future.

[34] Pindyck (2009) analyzes how different type of risk could magnify the cost of entry. I find the similar mechanism in my paper with two crucial distinctions: First, my mechanism depends on the post-entry failure risk that varies with the aggregate demand conditions, in his paper post-entry failure risk is independent from the aggregate demand conditions. Second, I use my model to study firm-dynamics over the business cycles and their contribution to dynamics of aggregate variables. Pindyck (2009) concentrates on explaining

Why the aggregate demand conditions matter at entry? What drives the value of waiting? and why the value differs for potential entrants across different signals? To answer the questions, I study how the initial aggregate demand level affects potential entrants post-entry returns. Consider an Equation 5 that decomposes gross value of entry into expected first period profit and expected continuation value conditional on survival, later referred as expected long run value.

The aggregate demand level at entry has a persistent effect on entrant's lifetime value for two reasons. First, the persistent aggregate demand process implies that entrants entering during high aggregate demand periods experience high aggregate demand levels for several periods after entry. Second, due to the persistent demand accumulation through customer capital, high aggregate demand level at entry leads to higher firm-specific demands in the following periods. Consequently, high aggregate demand level at entry increases not only entrants' first period profit but also entrants' long run value. In contrary, lower initial aggregate demand, decreases not only the first period profit but also long run value. Expected long-run value in turn determine expected post-entry survival rate.[35]

$$
\begin{aligned}
V^{\text{gross}}(b_o, q, z) &= \int\limits_s \left(\Pi(b_o, s, z) + \int\limits_{c_f} max \left\{ 0, -c_f + \beta(1-\gamma) E[V^I(b', s', z')|s, z] \right\} dG(c_f) \right) dH_e(s|q) \\
&= \underbrace{\int\limits_s \Pi(b_o, s, z) \quad dH_e(s|q)}_{\textit{Expected short run profit}} \\
&+ \underbrace{\int\limits_s \beta(1-\gamma) G\left(c_f^*\right) \left[E(V^I(b', s', z')|s, z) - \frac{1}{(1-\gamma)\beta} E(c_f \mid c_f \leq c_f^*) \right] \; dH_e(s|q)}_{\textit{Expected long run profits conditional on survival}}
\end{aligned} \tag{5}
$$

Where $c_f^* = \beta(1-\gamma)E[V^I(b', s', z')|s, z]$ is an expected value of fixed cost which makes incumbent firm indifferent between exiting or staying in the market.

industry dynamics and concentration to quantify the full cost .

[35]To remind reader, each period incumbent firms face random fixed cost c_f. If the expected continuation value is higher than observed fixed cost incumbent firms decide to stay in the market and continue operation. However, if the value is lower than the observed fixed cost than incumbent firm decides to exit from the market. For a potential entrant long-run value equals to expected continuation value.

In Equation 5, $G(c_f^*)$ represents expected survival rate of a potential entrant with signal q across different initial productivity levels s. The threshold fixed cost level c_f^* equals to incumbent firm's expected continuation value. Since the latter value increases with the aggregate demand level, expected survival rate $G(c_f^*)$ also increases with the aggregate demand level. Now, note that the procyclical expected survival rate acts as a procyclical discount factor, which one can easily see in the above equation. Due to the feature, potential entrants post-entry returns during recessions decrease relative to expected returns during expansionary periods. The result in turn explains why for given signal q, the *Value of Waiting(q,z)* decreases with the aggregate demand level z: by waiting potential entrants expect to receive higher level profits for longer periods of time.

The variation of the *Value of Waiting(q,z)* across signal levels depends on the relative variation in the entrants' expected survival rates. For any given aggregate demand, potential entrants expected survival rate and hence *Value of Waiting(q,z)* decreases with the signal q.

Finally, I study the consequences of the option to delay entry for an individual firm level. Toward that goal, I am going to compare optimal entry decision of a potential entrant with signal q with and without the option to delay entry.

Definition 3.1. Define a *threshold aggregate demand* level $\underline{z}_\tau(q)$ as the minimum aggregate demand for which the following inequality is satisfied:

$$NPV(q, \underline{z}^\tau(q)) \geq \tau \, Option\ \ Value\ of\ Delay(q, \underline{z}^\tau(q)) \tag{6}$$

Consider first a case when $\tau = 0$. The $NPV(q, z)$ is a strictly increasing function of an aggregate demand level z. If for a given q , there exist z for which $NPV(q, z) > 0$ then the threshold aggregate demnad level $\underline{z}_{\tau=0}(q)$ exist. Otherwise, the potential entrant with signal q never enters the market. Similar logic applies for a case when $\tau = 1$. Potential entrant with signal q decides to enter the market if *Value of Waiting(q,z)* < 0. The latter value is a decreasing function of an aggregate demand level z. If for a potential entrant with signal q there exist z for which *Value of Waiting(q,z)* < 0 then the numerical solution identifies a threshold aggregate demand level $\underline{z}_{\tau=1}(q)$. From the discussion follows the following result:

Result 3.3. *For any $\tau \in (0, 1)$, potential entrant with signal q decides to enter the market when the aggregate demand level is z if $z \geq \underline{z}_{\tau}(q)$ and not enter the market otherwise.*

Figure 14(a) shows that when $\tau = 1$ the threshold aggregate state for each potential entrant with signal q is weakly greater that the threshold aggregate state when the signal has zero persistence $\tau = 0$. The figure shows that the ability to postpone entry in the future has no effect on potential entrants that hold low-range and high-range signals.[36] While potential entrants with a medium-range of signals find it profitable to wait for a higher aggregate demand periods if they have the option to postpone entry.

Figure 14(e) shows the ratio of the first period profit to long run profit for each signal level. Potential entrants that hold low level signals put higher weight on the first period profit, since their survival rates are quite low. Due to this reason they decide to enter the market only during high aggregate demand periods irrespective whether they have the ability to delay entry or not.[37] The long run profit have the highest weight for potential entrants with high-range of signals. Initial aggregate state has a minor effect on their entry benefits since they survive with probability $1(1-\gamma)$. Thus, they decide to enter the market immediately, irrespective of the initial aggregate state. The option to wait allows potential entrants with medium-range of signals to choose an aggregate state that maximizes the value of the signal. Figure 14(c) and Figure 14(d) show that by entering during higher aggregate demand periods this group of entrants increase the first period profit and the long-run profit. Moreover, Figure 14(e) shows that the share of the long-run profit increases in the total value of entry, which implies higher expected survival rate (see Figure 14(f)).

For reasonable parameter values I find that during recessions the *Value of Waiting* increases the cost of entry from 0% to 8% of the fixed entry cost. I investigate the implications of the

[36] For potential entrants with a low-range of signals the option value of delay is very close to zero. Required increase in the NPV to cover the option value of delay is minor. As a result, the numerical solution reports the same threshold aggregate state for the case $\tau = 0$ and $\tau = 1$.

[37] More precisely, if the potential entrants do not have the ability to delay entry then during low aggregate demand periods they decide to stay outside the market and they drop out from the pool of potential entrants. With the ability to delay entry potential entrants with low level signals that are born during low aggregate demand periods still choose to stay outside the market but due to the option they could choose to wait for a high aggregate demand periods.

increased cost of entry on the expected number of periods that potential entrants are ready to stay outside the market and wait for higher aggregate demand conditions. I calculate the expected number of periods a potential entrant with signal q expects to get from a minimum aggregate state where he expects positive net benefits from entry ($\underline{z}^{\tau=0}(q)$) to an aggregate state where he optimally chooses to enter the market: $z \geq \underline{z}^{\tau=1}(q)$.[38] Figure 14(b) shows that the expected duration of waiting for potential entrants with medium-range of signals varies on average from one to six periods.[39] Among the potential entrants who choose to delay entry the expected number of periods is negatively correlated with the signal level. The smaller the signal the smaller the range of the aggregate state that potential entrants choose to enter the market. Thus, after delaying entry the number of periods that it takes to get to preferred aggregate demand conditions increases. Period in the calibration is defined as a year, indicating that the seemingly minor increase in the cost of entry due to aggregate risk generates significant number of inaction periods for certain group of potential entrants.

3.3.1 Comparative Statics

In this section I investigate how different parameters affect additional cost generated due to entrants ability to delay entry. The parameters of interest are: (i) The fixed entry cost (c_e), (ii) The variance of the aggregate demand shock process σ_z, (iii) The persistence of the aggregate demand shock process ρ_z, (iv) The persistence of the signal τ.

Change in the fixed entry cost (c_e) I consider three levels of fixed entry cost: $c_{e,\ low} < c_{e,\ baseline} < c_{e,\ high}$. I find that the higher the actual sunk cost of entry the higher the increase in the cost of entry today due to aggregate risk. Figure 15(a) shows that increase in c_e increases value of waiting for all potential entrants for all aggregate demand levels compared to the case where fixed entry cost is lower.

Figure 15(b) illustrates change in the threshold signal with and without the ability to delay

[38] More specifically, in the exercise I find how many number of periods medium-range of potential entrants are ready to stay outside of the market after they are given the option to wait. The figure depicts the expected number of periods for the aggregate demand process with $\rho_z = 0.57$ and $\sigma_z = 0.0022$. The process is calibrated to match the business-cycle dynamics of the entry rate in the model and in the data

[39] Note, that very small group of entrants the time of waiting spikes to 20 periods. These are potential entrants with very small productivity signals and wait for the highest aggregate demand periods.

entry across different levels of fixed cost. Increase in the fixed cost of entry directly decreases life-time benefits from entry for each potential entrant in any state and increases threshold signal for each aggregate demand level. One can see that for $\tau = 0$ the threshold signal lines shift parallelly upward with the increase in the fixed cost of entry, indicating that change in the fixed cost of entry does not affect elasticity of threshold signal with respect to the aggregate demand level. Next consider threshold signal when $\tau = 1$. The direct effect of increased fixed cost of entry is still dominant force affecting the threshold signal. I find that the effect of increased cost of entry on the magnitude of the indirect effect (difference between threshold signal in a case with and without delay entry) increases but the increase is moderate.[40] Thus, increase in the fixed cost of entry has a minor effect on increasing the elasticity of threshold signal with respect to aggregate demand level even if the entrants have the ability to delay entry.

Additionally, I find that the mechanism still works when the fixed entry cost equals to zero. Variation in the relative life-time profits across aggregate states creates the room for potential entrants to use profitable the option to delay entry.

The variance of the aggregate demand shock process σ_z I study how the variance of the aggregate demand shock process affects the selection of entrants. I consider two different levels of the variance: $\sigma_{baseline} < \sigma_{high}$. I find that increase in the variance have minor effect on the net benefits of entry today, which one can see on Figure 16(b) where the threshold signal lines in the $\tau = 0$ case does not vary with the level of the variance. However, increase in the variance of the aggregate demand shock process have a significant effect on the threshold signal when $\tau = 1$. The reason is twofold: First, increased variance increases the probability of low aggregate demand levels. I discussed above, the indirect effect significantly increases with the decrease in the aggregate demand level. Second, by effecting the aggregate demand shock process, increase in the variance also affects relative returns today versus tomorrow. The latter effect increases cost of entry during recessions.

To conclude, change in the variance has a significant effect on business-cycle variation in

[40] To check the magnitude of the indirect effect I rescaled threshold signal so that the threshold signal equals each other during the highest aggregate demand periods, and I find that during recessionary periods the magnitude of the indirect effect is highest in the case with the highest fixed cost of entry.

the number and the composition of entrants, while change in the sunk entry cost does not have a significant effect on the business-cycle variation in the number and the composition of entrants. The result has important policy implications. If for any reason one wants to stabilize the entry rate the most effective way to do so is stabilizing post-entry profits rather than changing the fixed cost of entry. The policy implications of accounting potential entrants ability to delay entry is discussed in the last section of this paper.

The persistence of the aggregate demand shock process ρ_z Increase in the persistence of the aggregate demand shock process has two effects: it increases persistence and at the same time it decrease the variance of the aggregate demand shocks. The latter effect is due to the fact that $\sigma_z = \dfrac{\sigma_{\varepsilon_z}}{\sqrt{(1-\rho^2)}}$.

Figure 17(a) illustrates that the level of the value of waiting decreases and the elasticity of the value of waiting increases with the persistence level. Potential entrants expect that the current state is going to be very persistent, which decreases the return from postponing entry. Figure 17(b) displays threshold signal for different levels of the persistence of the aggregate demand shock process: $\rho_{baseline} < \rho_{high}$. As expected threshold signal weakly decreases with the persistence of the signal for each aggregate demand level at entry.

Change in the persistence of the signal τ Persistence of the signal τ describes probability that potential entrant with signal q expects to get the same signal tomorrow. Figure 18(a) illustrates the option value of delay across signal levels for different τ. The maximum discounted net present value of profits that potential entrants can get by delaying entry today is described by the level of the option value when $\tau = 1$. While the minimum level of benefits from delaying entry is described by the option value when $\tau = 0$ and it equals to 0. Returns from delaying entry when $\tau \in (0, 1)$ is in between these two scenarios. The higher the persistence level of the signal the higher the return that each potential entrant expects to get tomorrow. However, note that when τ decreases the option value of delay starts to decrease faster.

Figure 18(b) displays threshold value of signal for different τ across different aggregate states. Elasticity of threshold signal with respect to aggregate demand level increases with τ. In the rest of the paper I am going to concentrate on a case where $\tau = 1$.

3.3.2 General environment

So far I have discussed the main finding of the paper: the ability to delay entry accompanied with the aggregate risk that affects potential entrants post-entry returns endogenously generates countercyclical cost of entry and significantly affects selection of entrants over the cycles. In this section, I am going to argue that the result is not specific to the model environment and the mechanism will work successfully in any business-cycle firm dynamics models where post-entry risk of failure procyclically varies with the aggregate conditions. I concentrate on the following features of the model: (i) Customer capital accumulation; (ii) Noisy versus perfect information about the first period productivity; (iii) Partial equilibrium; (iv) Aggregate demand shock; (iv) The option to delay for incumbent firms.

Customer capital accumulation In the model the persistent customer capital accumulation contributes to the persistent effect of the initial aggregate conditions on potential entrants post-entry dynamics. I investigate how crucial is the feature of the model for generating the significant effect of the aggregate demand conditions on selection of entrants. After setting the elasticity of the customer capital η to zero I find that the feature has a minor effect on the selection of potential entrants through indirect effect. The result is partly because the variance of the calibrated demand shock process is quite low, thus the difference between the optimal customer capital over the business cycles are minor. However, the higher the difference between the aggregate demand conditions the higher the contribution of the persistent customer capital dynamics to the option value of delay mechanism.[41]

Signal informativeness In the model, potential entrants are endowed heterogeneous signals that give them noisy information about their first period productivity. The feature of the model generates additional risk associated to post-entry profits. However, since the risk associated with the first period productivity does not vary with the aggregate demand

[41] Incorporating the customer capital in the model is important for generating salient features of the micro-level data. More specifically, Foster et al. (2016) finds that the size difference between new and incumbent firms that close very slowly is due to the differences in their individual demand dynamics rather than difference in their productivities.

conditions the feature of the model does not contribute to the relative variation of the cost over the cycles. In fact, a model where heterogeneous potential entrants have a perfect information about their post-entry productivity increases the importance of the option to delay entry since each potential entrants are able to evaluate more precisely post-entry profits and failure risks and they choose optimal aggregate demand conditions accordingly.

On the other extreme, the mechanism does not work in a model where the signal carries no information about the post-entry productivity, since in that case potential entrants end up to be homogeneous. In these models if one potential entrant decides to delay entry then all potential entrants want to delay entry too. Thus, the models to have an interior solution for entry rate require the option value of delay to equal to zero.[42] Thus, for the mechanism to describe selection of entrants at an aggregate conditions its important that signal carries some information about the post-entry productivity.

Finally, the feature of the model is important to account for the composition variation of entrants over the cycles. With perfect information about the post-entry productivity selection would have happened from lower range of productivity entrants, pure cleansing effect. I show that the share of potential entrants that would have ended up receiving high productivity signals are crucial to explain the persistent effect of entry to aggregate dynamics.

General equilibrium effects The model is developed in partial equilibrium where firm-dynamics does not affect aggregate prices and hence relative prices. Appendix A.2 describes general equilibrium version of the model.[43] In general equilibrium, both wages and the stochastic discount factor become procyclical.[44] The procyclical discount factor makes delay favorable, since potential entrants give more weight to high conditions of aggregate demand. The procyclical variation in wages makes delay less favorable during recessionary periods. While the former component contributes to the selection of potential entrants through the option value of delay, the latter component decreases the importance of the option value of

[42]For example, see paper Bilbiie, Ghironi and Melitz (2012).

[43]Note that the model presented in the main body of the paper is a reduced form of a general equilibrium model with infinitely elastic labor supply $\chi(L_t) = \psi L_t$ and where the demand of aggregate consumption basket is given by $P_t = C_t^{\rho}$.

[44]Hong (2018) expands Moreira (2015) to general equilibrium framework and shows that the stochastic discount factor is procyclical.

delay mechanism. However, since the option value of delay is nonnegative (due to entrants' ability to stay outside) the channel weakly increases the threshold cost of entry in any aggregate state compared to a model with fixed entry cost. As a result, a model with option value, whether in partial or general equilibrium, amplifies the effect of aggregate shocks relative to a model with a neoclassical investment decision rule.[45]

More specifically, potential entrants are not going to affect aggregate prices significantly since the competition for entry will happen at the first stage of entry phas while allowing free entry for individuals that want to decide to enter the market each of the person competes for their own signal category. As a result, business-cycle variation in the market attractivness for potential entrants is absorbed at free entry stage whith the change of probability to end up become a individuals

Aggregate demand shock Finally, the mechanism holds for any aggregate shock process that generates procyclical variation in entrants' expected survival rates. That feature of the model acts as a procyclical stochastic discount factor, which increases the value of waiting during recessions by increasing the net benefits of entry during expansions.

The option to delay for incumbent firms In the model incumbent firms are able to delay production in the same respect as potential entrants, since incumbent firms productivity process is persistent and they stay market every period. However, incumbent firms optimally choose not to use the option to delay production. The crucial difference between potential entrants' and incumbent firms' decisions is the fact that for potential entrants delaying is costless, while its costly for incumbent firms. In particular, incumbent firms need to pay fixed cost to stay in the market irrespective whether they undertake production or not. Thus, they optimally decide to stay if the expected value from staying is higher than observed idiosyncratic fixed cost. For potential entrants delaying entry and waiting better aggregate

[45]The firm-dynamics models developed in general equilibrium (in addition to low elasticity of life-time profits with respect to aggregate demand level) face additional challenge: pro-cyclical variation in the wages and response of potential entrants due to free entry condition mitigate the effect of the aggregate conditions on the selection of entrants. In Appendix A.1.1 I propose a simple modification of a two-stage entry phase developed initially in Lee and Mukoyama (2008) that also addresses the problem by endogenously restricting potential entrants effect on the aggregate prices.

conditions costs nothing. Thus, the real options are important when going inaction is not costly, and turns out that the mechanism is very important when one considers entry or investment decision at extensive margin.

3.4 Suggestive evidence

I use Newly developed Business Formation Statistics dataset to provide suggestive evidence about the delays in potential entrants decisions to enter the market and become employer businesses. The is a unique dataset that enables to track entrepreneurs from the beginning of the time when they first initiate to start business activities.[46] In particular, the dataset collects information about the applications for Employer Identification Numbers (EINs) submitted in the United States, known as IRS Form SS-4 filings.[47] In most of the cases applicants provide detailed information about the reasons why they apply for EIN. The information is used to identify applications that have high likelihood to become employer businesses. For example, applications that are for corporate entities, indicate the number of employs, provide first wage pay date and planned wages, and that belong to some specific industries are refereed as *High Propensity Business Applications (HBA)*.[48]

I use quarterly time series of the number of HBA applications over the period $2004Q3 - 2014Q4$ to evaluate the delays in aspiring start-ups market entry decisions. Figure 18 illustrates the share of the applications that translate into employer businesses within 4 or 8 quarter period from the application date. Main takeaways from the figure are as follows: (i) on average, only 26% of the applicants start market operations within a year, (ii) on average, 4% of the applications start operation with one year delay, and (iii) major share (70%) of the applicants delay more than two years to enter the market.[49] Bayard et al. (2018) use private owned part of the BFS dataset and additionally show that the delay behavior of as-

[46]At this point, the information provided in the publicly available part of the dataset does not allow me to conduct more detailed analyzes about the mechanism.

[47]EIN is a unique number assigned to most of the business entities. EIN is required when the business is providing tax information to Internal Revenue Service (IRS). Note that EIN applications describe start-up and not establishment level activities, since opening a new establishment does not require new EIN.

[48]For detailed description of the dataset and subcategories of applications please visit the Business Formation Statistics dataset on the U.S. Census Beauro *website*.

[49]Figure 19 illustrates average duration (in quarters) from business application to formation within 4 and 8 quarters increases over time.

piring startups prevail across industries of application and types of entity. While keeping in mind that EIN (even HBA) applications provide noisy information about the actual business formation, I use the findings as an evidence that aspiring start-ups have and use the ability to delay entry.

Bayard et al. (2018) also shows that the firms that become employer business in year 2013 consists of firms that applied for the EIN starting from the year 2004 and that the application activity is positively correlated to aggregate economic conditions. Suggesting that the BFS can be used in the future to study in more details factors that affect entrants market entry decisions, especially timing of entry over the business cycles.

At this point there is no direct link between the model describe in the main part of the paper and the information provided in BFS dataset. However, in Appendix A.1.1 I propose an extension of entry phase that provides a starting point to draw connections between the model and the data. In fact, the entry phase even proposes a possible way to start thinking about the large share of the applications that never transitions into employer business.

Other evidence Considerable amount of theoretical and empirical microeconomics literature that emphasize the importance of the ability to postpone an irreversible investment/entry decision under aggregate state volatility provides additional support of the mechanism proposed in the paper. The analysis dates back to Arrow (1968)[50]. As a one example of the empirical literature that studies how the uncertainty affects entrepreneur entry decision is O'Brien, Folta and Johnson (2003). They use 1993 National Survey of Small Business Finance(NSSBF) and the 1998 Survey of Small Business Finance(SSBF) and show that high uncertainty in the target industry decreases the probability of entrepreneur entry. They explain the link by the irreversibility of the entry decision.

Next, I move on and quantify the role the option value of delay plays in selection of entrants over the business cycles and aggregate dynamics.

[50]See Dixit and Pindyck (1994) for detailed review.

4 Functional forms and calibration

4.1 Functional forms

The fixed operating cost is distributed log normally with parameters μ_f and σ_f. Aggregate distribution of the signal $W(q)$ is set to be Pareto with location parameter $\underline{q}$ and Pareto exponent $\xi > 0$. For given signal, idiosyncratic shock in the first period of operation is Normally distributed and follows the process $log(s) = \rho_s \ log(q) + \sigma_s^e \ \epsilon, \quad \text{where } \epsilon \sim N(0,1)$.

4.2 Calibration

Estimating the model requires calibrating the following 17 parameters:

$\left\{\beta, \rho_s, \rho, \eta, \delta, b_o, \sigma_s, \sigma_s^e, \underline{q}, \xi, \mu_f, \sigma_f, \gamma, c_e, \alpha, \rho_z, \sigma_z\right\}$

This section describes the calibration procedure of the parameters. The summary of the identification strategy and the final values of the parameters is given in Table 2.

To be consistent with the BDS timing a period is assumed to be one year. The unit of analysis is the establishment. I set the time preference parameter $\beta = 0.96$ to match a *4* percent annualized average riskless interest rate. I choose estimates of the parameters driving the persistence of the idiosyncratic productivity process ρ_s from Foster et al. (2008) which uses the same specification of the production function. Technology in Foster et al. (2008) is linear in inputs and productivity: $q_i = s_i x_i$ where x_i is the input and s_i is producer-specific productivity. Using establishment-level data from the Census of Manufactures they directly measure Total Physical Factor Productivity defined as $TFPQ_i = \frac{s_i x_i}{x_i} = s_i$.[51] Autoregressive properties of the measured TFPQ implies persistence rate $\rho_s = 0.814$.[52]

To calibrate parameters for the demand function I relied on Foster et al. (2016) which uses the same specification of the demand function and production technology. To identify parameters, they jointly estimate demand and Euler equation, using the dataset from Foster

[51] Foster et al. (2008) uses establishment-level data for producers of eleven manufacturing products for the following census years: 1977, 1982, 1987, 1992 and 1997. The data gives information about producer-level quantities and prices separately.

[52] Foster et al. (2008) finds that persistence of TFPQ is very close to the persistence parameters generated from other measures of TFP (e.g. Traditional measure of TFP and revenue TFP)

et al. (2008).[53] Based on their estimation results, I set price elasticity of demand (ρ) equal to *1.622*, elasticity of demand to customer capital (η) equal to *0.919*, and depreciation rate of reputation (δ) equal to *0.188*.

The rest of the parameters that drive potential entrants distribution ($\underline{q}$, ξ), selection at entry (c_e), survival function (μ_f, σ_f, γ), average size of entrants (b_0, σ_s^e), growth of entrants (σ_s), and average size of all active establishments (α) are jointly determined to make sure that the simulated moments that characterize cohorts' post-entry characteristics in the stochastic steady state of the model match the data counterparts calculated from the U.S. establishment data.

In particular, cohort characteristics at age 0 is captured by the following moments: average entry rate[54], share of entrant employment in total employment, average size of active establishments[55], and average size of entrant establishments (cohort at age 0). To capture cohorts' post-entry characteristics, I target the following moments: average size of cohort at age 5 and between 21 and 25 years old, average survival rate until 5 years old, survival between 21 and 25 years old, establishment exit rate at age five and cohort employment share in total employment at age five. I calculated the moments using the economy-wide establishment level data from the BDS dataset over the 1977-2015 period. Detailed description of the construction of the empirical moments is described in Appendix D.1.1. The first and the second column of Table 3 reports the values of the calibration targets and the model simulated counterparts.

The parameters that drives the aggregate demand shock process (ρ_z, σ_z) are calibrated to match the autoregressive properties of the cycle component of entry rate in the model and in the data. The entry rate data comes from the BDS dataset and covers the period 1977-2015. To calculate the cycle component of the log entry rate I apply the HP filter with smoothing parameter 100. The autocovariance and standard deviation of the later series are reported in the second column of Table 4.

[53] In Foster et al. (2016) firms need to pay constant fixed cost of operation, while in my model operational fixed cost is drawn randomly each period. However, since they estimate Euler equation conditional on survival, the final estimated parameters represent good fit to the model parameters.

[54] I calculate average entry rate over the $1991-2007$ period. For the evolution of the entry rate over the full 1977-2015 period see the Figure 20.

[55] Size is defined as the total number of employment by entrants/incumbents/all establishments over the total number of entrants/incumbents/all establishments.

To generate the model counterparts of the data moments, I simulate the economy over 10000 periods and apply the same detrending method to the model simulated entry rate. The values of the calibration targets derived from the simulation of the model is reported in the third column of Table 4. Finally, I set $\rho_z = 0.57$ and $\sigma_z = 0.0022$.

Finally, I set the persistence of the signal τ equal to one. In Appendix D.2, I propose a strategy to identify τ using the time-series of the aggregate employment. I find that $\tau = 0.965$. I find that the dynamics of the economy with the value of $\tau = 0.965$ is very close to the one when $\tau = 1$.

5 Entrants' life-cycle dynamics

Recent firm-dynamics literature emphasize the importance of accounting entrants post-entry demographics in measuring and understanding their contribution to the aggregate dynamics. Haltiwanger et al. (2013) findings emphasize the critical role start-ups play in U.S. employment growth dynamics. Their finding that young firms exhibit distinct life-cycle dynamics compared to their mature counterparts emphasize the importance for accounting not only the entry process but the subsequent post-entry dynamics (growth, survival, job creation). Empirical findings by Sedlacek and Sterk (2016) and Moreira (2015) show that the persistent differences in cohorts' life-cycle characteristics over the business cycles could account for the significant amplification and propagation of the aggregate shocks.

In the rest of the section, I show that the model simulated cohorts' post-entry characteristics in the stochastic steady state and over the business cycles match to the data counterparts. Later, I use the good fit of the model to evaluate the role of entrants demographics in shaping the dynamics of the aggregate employment, output and total number of firms over the business cycles.

5.1 Cohorts' post-entry characteristics in the stochastic steady state

I begin by evaluating the performance of the model in the stochastic steady state. Table 3 shows that the model simulated entry rate, average cohort size at entry, average size of all establishments (hence, the relative entrant size), and entrants employment share in total employment is close to the data counterpart.

Figure 22(a) and Figure 22(b) compare the evolution of the steady state cohort survival rate and exit rate by age in the model and in the data over the thirty years of operation.[56] The figures show that the simulated average survival rate and average exit rate by age closely mimics to the data counterpart. Figure 22(c) shows that the model simulated cohorts' average size over the thirty years of operation exhibits same dynamics as in the data.

The model also has strong predictions about relative contribution of cohort employment to the aggregate employment. Figure 22(d) shows that the share of the simulated cohort's employment in aggregate employment up to five years of operation generates close match to the data counterpart.[57]

5.2 Cohorts' post-entry characteristics over the business cycles

This section shows that the model simulated cohorts that are born at different stages of business-cycles exhibit documented distinct post-entry characteristics. To describe the business cycle conditions at entry, I use the aggregate demand shock process. The aggregate demand is in the stochastic steady state when $z \;= 1$. Unless otherwise defined, I refer periods recessionary (expansionary) when an aggregate demand is below (above) the stochastic steady state level $z < 1$ ($z > 1$). I define cohorts as recessionary (expansionary) if they start operation during the recessionary (expansionary) periods.[58]

[56]The data moments are described respectively in Appendix D.1.3 and in Appendix D.1.2.

[57]The BDS dataset does not give possibility to identify individual cohort employment after five years of operation.

[58]The model simulated results are robust to the definition of the business cycles. In particular, results are similar if I define recessionary and expansionary periods using deviations from the average log employment (output) or if I use cycle component of the HP filtered log employment (output). The robustness of the results are due to the fact that the model generates more all less symmetric business cycles.

To quantify the role of the persistent signal, I set τ equals to zero in the baseline scenario. When $\tau = 0$, option value of delay reduces to zero and all potential entrants choose to enter into the market if expected value of being an incumbent is more than the fixed entry cost. Additionally, I re-calibrate the case with $\tau = 0$ so it generates a match to the same set of moments in the stochastic steady state as the baseline model. Column (b) of Table 5 summarizes the parameter values used in the case with $\tau = 0$. Note that the only parameter that is different between the baseline scenario and the case with $\tau = 0$ is fixed entry cost, which determines threshold signal in the stochastic steady state.[59] In $\tau = 0$ the opportunity cost of entry is constant over the cycles and equals to the fixed entry cost. While in the baseline scenario due to the option value of delay, the total opportunity cost of entry is countercyclycal. The latter implies different threshold signal, hence the different number and composition of entrants in the baseline model compared to the case with $\tau = 0$. As a result, the differences between the performance of the two scenarios come when we go beyond the stochastic steady state. In the following sections, I interpret the differences between the performance of the baseline model and the case with $\tau = 0$ as a consequence of the additional selection generated due to the entrants ability to delay entry.

5.2.1 Average productivity

Recent empirical findings show that recessionary cohorts are on average more productive compared to expansionary cohorts. Lee and Mukoyama (2015) document that entrants Total Factor Productivity (TFP) relative to incumbent firms are significantly higher during the recessionary periods compared to the expansionary periods. Moreira (2015) finds that the counter-cyclical difference in entrants average labor productivity persists for several years after entry.

Figure 23(a) depicts entrants distribution over initial productivity for different aggregate demand level. Consistent to the empirical findings, the aggregate economic conditions at

[59]It turns out that setting the fixed entry cost in the $\tau = 0$ case to the steady state total opportunity cost level(fixed cost plus option value of delay for threshold signal when $z = 1$) from the baseline model, without changing any other variables is enough to produce exactly same results in the stochastic steady state. The choice of the fixed entry cost makes sure that the threshold quality of signal, hence the number and composition of entrant firms coincides in the steady state across cases. Column (b) of table 5 summarize parameter values used in the case with $tau = 0$.

entry have a significant and persistent effect on the productivity composition of entrants in the benchmark model. Figure 23(a) depicts entrants distribution over initial productivity for different aggregate demand level. The productivity distribution of entrants is positively skewed. The productivity distribution of entrants is positively skewed. The skewness significantly decreases with the aggregate demand level, producing countercyclical average productivity. Average productivity, measured as average Total Physical Factor Productivity (s) of entrants that enter during recessionary periods is around 3% higher compared to the entrants that enter during the expansionary periods. The difference in cohorts' productivity persists in later years due to the persistent idiosyncratic productivity process.

Figure 23(b) shows the productivity composition of entrants across aggregate states for the case when the signal has zero persistence. In the case with $\tau = 0$, the difference between the average productivity is at most 0.4%. Comparing the benchmark model to the case shows that the significant effect of the aggregate state on the productivity composition of entrants is due to the significantly countercyclical threshold signal generated through the option value of delay component in entry decision (as discussed in Section 3 and given in Figure 9). Which agrees to the point made by Lee and Mukoyama (2016) that selection only through constant entry cost is not sufficient for generating documented counte-cyclical productivity.[60]

5.2.2 Average survival rate

Figure 24(a) plots the average survival rates for the expansionary and the recessionary cohorts for the baseline model.[61] The figure shows that the model generates countercyclical average survival rate.

The aggregate demand has two counteracting effects on cohorts' survival rates. On the one hand, lower aggregate demand directly decreases each entrant firms' survival rate by decreas-

[60] In later sections I show that to generate significant selection in the model without option value of delay requires 7 time increase in the variance of the aggregate demand shock process, which in turn produces unreasonable implications for the volatility of the aggregate variables.

[61] To compare average survival rate generated by the model to the data counterpart I define recession (expansion) as the period when the aggregate demand is 1% below (above) the steady state level. I define cohorts as recessionary (expansionary) if they started operation during the recessionary (expansionary) periods. The definition correspondes to the second recession indicator defined in Section 5.2.3. After simulating the economy over 400 periods I calculate average survival rate for the recessionary and expansionary cohorts up to 15 years of operation.

ing individual demand level and continuation value. On the other hand, the lower aggregate demand level implies increase in average productivity of the entrants, which increases average survival rate of the cohort. The benchmark framework implies that the change in productivity composition dominates the direct effect of the aggregate demand level and as a result the model simulated cohorts that are born during the recessionary periods have higher average survival rates compared to the cohorts that are born during the expansionary periods.

Figure 24(b) plots the same statistics when the signal has zero persistence. The minor change in the productivity composition of entrants implies that the two counteracting effects described above cancel each other out and generate same average survival rate for the expansionary and recessionary cohorts.

The option value of delay component in entry decision is essential to generate the countercyclical survival rate. The reason is as follows. Persistent signal gives potential entrants the ability to enter into the market when the aggregate state maximize the exercise value of the signal. As I described in Section 3, potential entrants prefer to wait and enter during the higher aggregate states. Higher aggregate state implies relatively higher first period profit and long run value. If I calculate the expected survival rate at threshold aggregate state I see that by delaying entry potential entrants with medium size signals increase the expected survival rate, (relative to the case $\tau = 0$), refer to Figure 14(f). As a result, the additional selection through the option value of delay generates cohorts with higher survival rates.

This result provides a clear testable implication for the option value of delay mechanism. Using the Business Dynamic Survey database over the period 1997-2015, I show that cohorts that start operating during recessions have on average persistently higher survival rates compared to cohorts that enter the market during expansions. I interpret the result as providing support to the option value of delay mechanism.

5.2.3 Average survival rate: empirical evidence

This section provides evidence that cohorts of U.S. establishments born during the recessionary periods have persistently higher survival rate compared to the cohorts that start operation during the expansionary periods.[62]

To describe business cycle conditions at entry, I use the cyclical component of the annual log real GDP over the period 1977-2015. To find the cyclical component of the annual log real GDP, I use the Hodrick and Prescott (1997) (HP) filter with smoothing parameter 100.[63] Using the series, I define the following two business cycle indicators: [1] A period is referred to a recession (expansion) if the cyclical component of the log real GDP is negative (positive),[64] [2] A period is referred as recession (expansion) if the cyclical component of the log real GDP is below (above) trend by 1 percent. Cohorts born in years when the absolute deviation of the log real GDP from the trend is less than 1 percent are referred as mean cohorts.[65]

To construct survival rates, I use economy-wide establishment level data from the BDS database over the 1977-2015 period.[66] For each cohort born at year t survival rate up to g years of operation is defined as $S_{t,g} = \dfrac{N_{t,g}}{N_{t,0}}$ where $g = 0, 1, 2, 3, 4, 5$.[67] $N_{t,g}$ is number of establishments from cohort t in age g and $N_{t,0}$ is the number of entrant establishments in cohort t at birth (age 0).[68]

Cohort born at year t is referred as recessionary (expansionary) cohort if t is indicated as a recessionary (expansionary) year based on the above described indicators. The average survival rates of the recessionary (expansionary) cohorts is defined as the average survival rates of the cohorts born during the recessionary (expansionary) years.

Figure 2 and Figure 3 plot the average survival rates of the recessionary and the expansionary cohorts defined using the first and the second indicators respectively. Both of the figures show that the average survival rates of the recessionary cohorts are persistently higher over the first five years of operation compared to the survival rates of the expansionary cohorts. Moreover, Figure 3 shows that the difference between cohorts survival rates is higher when

[62] The result also holds for the firm level data.

[63] The source and the construction of the annual real GDP data is described in Appendix C.1.

[64] The indicator takes value 1 if a year is defined as recession and 0 if a year is defined as expansion.

[65] The indicator takes value 1 if a year is defined as recession, 0 if year is defined as mean year and -1 if a year is defined as expansion. The choice of the magnitude of the deviation equally divides 39 observation into three groups.

[66] The data is described in details in Appendix C.2.

[67] Assuming no censored observations the estimation coincides to Kaplan-Meier estimator.

[68] After age 5 data groups cohorts into 5 year bins, which makes decomposing cohorts into recessionary and expansionary groups unfeasible.

one compares cohorts born in more distinct aggregate conditions. To further check the hypothesis, I calculate the correlations between the cyclical component of the log real GDP and the average survival rate at different ages of operation. The result is given in the third column of Table 1. Negative values of the numbers indicate that increase in the cycle component is negatively correlated to the survival rate at any age up to 5 years of operation.

In addition to the two indicators described above, I consider the following definition of the business cycle indicator: [3] NBER based Recession Indicators for the United States from the Period following the Peak through the Trough.[69] The indicator is based on the Peak and the Trough dates defined by the NBER. I take the indicator from the Federal Reserve Economic Dataset (FRED) in quarterly frequency that covers the period 1977-2015. Using the quarterly data I define a year t as recession if at least two quarters from the second quarter of year $t-1$ to the first quarter of year t is indicated as recessionary period. This definition of recessionary years make sure that the indicator matches to the timing of the BDS dataset. In particular, in the BDS dataset establishment level activity at year t covers the second quarter of year $t-1$ to the first quarter of year t. Based on the definition the recessionary years turns out to be $1981, 1982, 1983, 1991, 2002, 2009$.[70] All other years are defined as expansionary. Figure 4 displays the expansionary and the recessionary cohorts' average survival rates based on this indicator. The results also indicate the persistent difference in cohort survival rates over the cycle.

Taking into account that recessionary cohorts survival rate is negatively affected by the persistent post-entry aggregate conditions, the documented counter-cyclical survival rate indicates the substantial effect of the initial aggregate state on the composition of the cohorts at entry.

[69]The other NBER based recession indicators do not allow me to decompose establishments that start operation during the Trough or Peak periods since the indicators either include Peak with the Trough periods or exclude the Trough periods. Examples of the indicators are (a) NBER based Recession Indicators for the United States from the Peak through the Trough and (b) NBER based Recession Indicators for the United States from the Peak through the period preceding the Trough either include Peak or exclude Trough.

[70]For example, year 1981 covers the period from the second quarter of year 1980 to the first quarter of year 1981.

5.2.4 Cohort-level employment

Variation of the productivity composition of entrants over the business cycle have a long-lasting effect on the post-entry cohort employment. Figure 25(a) plots expansionary and recessionary cohorts total employment at entry and over time for the baseline model. The model predicts that the recessionary (expansionary) cohorts employ 5.7% less (5.0% more) after entry compared to the average cohort and the difference does not disappear even after fifteen years of operation. The result is consistent to the empirical findings by Sedlacek and Sterk (2016): The employment created by entrant firms is volatile and pro-cyclical and that the variation persists as the cohorts age.

The aggregate conditions at entry have a negative effect on individual entrants' employment level (size) at entry and over time due to the persistent demand dynamics. Nevertheless, due to the significantly countercyclical average productivity of entrants it turns out that cohorts' average size is countercyclical. The result is in line with Lee and Mukoyama (2015), who show that average size of U.S. manufacturing plants is significantly countercyclical. However, the result is at odds to Sedlacek and Sterk's (2016) finding. Using the BDS dataset they show entrants' average size is procyclical. I expect that extending the model to account for the procyclical average size at entry will increase the difference in cohort-level employment over the cycles.[71] Consequently, one can interpret the results as the lower bound of the contribution of the entry to the aggregate dynamics. The rest of the section shows that the variation in the types of entrants accounts major part of the persistent pro-cyclical variation in cohort-level employment.

Propagation through variation in entry margin In the model the persistent effect of the initial economic conditions on cohort level employment can be attributed to two factors. First, firms that enter during recessionary periods employ fewer workers due to lower aggregate demand level. The persistent demand dynamics creates persistence in employment dynamics and the cohort that starts during recessionary periods find it difficult to catch up to a cohort which enters into the market during better economic conditions. Second,

[71] One potential extension to generate procyclical average size of entrants is to assume that the first period level of the customer capital is procyclicial.

the procyclical variation in cohort level employment comes from the selection of potential entrants at entry stage.

To identify how much the persistent demand dynamics explains the persistent variation in cohort level employment, I consider only the effect of the aggregate demand on cohorts' post-entry employment by shutting down the selection of entrants over the business cycles. In particular, I simulate the baseline economy while fixing the number/composition of entrants at stochastic steady state level. Row (a) and row (b) of Table 7 summarize the variation in cohort-level employment over the cycles respectively for the baseline and the described scenario. Row (b) shows that the persistent demand dynamics have a persistent effect on cohort level employment, however the magnitude of the effect is almost 15 times less than the baseline scenario.[72] The difference shows that the variation in cohort level employment is determined at entry stage, through variation in composition/number of entrant firms over the aggregate states.

Propagation through indirect effect Aggregate demand at inception affects selection of entrants through direct and indirect effect. Figure 25(b) displays cohorts employment for the case with $\tau = 0$ and shows that shutting down the option value of delay in entry decision reduces the difference between cohort level employment over the business cycles to 1%. This fact implies that the major part (80%) of the deviation in cohort employment is explained by the entrants that delay entry.

Propagation through indirect effect: number vs composition Next, augmenting entry decision with indirect effect has two effects. First, it increases volatility of entrants (by seven times) due to increase in elasticity of threshold signal with respect to aggregate demand. Second, the composition of entrants changes since the potential entrants that delay entry have higher range of productivity signals. To find out the importance of the composition of entrants I consider two additional counterfactual scenarios where the variation in the number of entrants over the cycles coincides to the baseline model. However, I let the composition of entrants to vary systematically across the scenarios. In particular, the first

[72]In Moreira (2015) persistent difference in cohort level employment is mostly due to the persistent demand dynamics.

scenario systematically adjusts lowest productive entrants from the distribution generated by the case with $\tau = 0$.[73] The second scenario systematically adjusts the highest productive firms from the distribution generated by the case with $\tau = 0$.[74] Rows (d) and (e) of table 7 summarizes the dynamics of the cohort level employment in these counterfactual scenarios. Since variation in the number of entrants are equalized, contrasting the baseline model to counterfactual scenarios helps us evaluate the role of the composition of entrants in the post-entry cohort dynamics. One can see, that if the aggregate demand at entry affects only low-productive entrants, variation in cohort level employment at entry increases by 3 times relative to the case with $\tau = 0$. However, over time the difference dies out. By the end of the year fifteen, variation in cohort level employment in the first scenario is similar to the case with $\tau = 0$, case with 7 times less volatile number of entrants. Indicating that variation in the number of entrants can explain magnitude of the difference at entry (age 0) but does not explain the persistent difference in cohort level employment over time. On the other hand, additional selection through high-productive entrants significantly increases variation and persistence in the dynamics of the cohort level employment as one can see in row (d). High-productive entrants survive with higher probability and contribute to the cohort employment for longer periods of time.[75] Indicating that the variation in these type of entrants is crucial in explaining persistent differences in cohort level employment.

Potential entrants selection at entry in the model with the option value of delay is in between these two scenarios. Potential entrants that choose to delay entry have relatively higher level of signals, which increases the probability that those potential entrants receive higher initial productivity levels after entry. Comparing columns (d) and (e) to the baseline scenario shows that the high share of those potential entrants still end up to be low productive ones. However, exactly the small share of the high-productive entrants that postpone entry produces persistent differences in cohort level employment. The mechanism is consistent with

[73]The scenario implies less low-productive entrants during recessionary periods, more low-productive entrants during expansionary periods compared to the case with $\tau = 0$. For more details refer to the Appendix E.2.

[74]The scenario implies less high-productive entrants during recessionary periods, more high-productive entrants during expansionary periods compared to the case with $\tau = 0$. For more details refer to the Appendix E.2.

[75]The mechanism corresponds to the 'missing generation' effect initially discussed in Gourio, Messer and Siemer (2015).

Decker et al.'s (2014) empirical findings, who show that the entrant cohorts contribution to the aggregate employment comes from the small share of the high-growth firms. Pugsley, Sedláćek and Sterk (2016) shows that major share of the entrant cohorts post-entry performance is due to ex-ante differences in the types of entrants. Moreira (2015) and Sedlacek and Sterk (2017) also document that the aggregate state at intercept drives the variation in cohort level employment.[76]

6 Entrant demographics and aggregate fluctuations

6.1 A model with the option to delay entry

Next, I investigate how the model that accounts for the life-cycle dynamics of the cohorts of the U.S. establishments explains the business cycle dynamics (autocorrelation and variance) of the economic aggregates, such as total number of firms, aggregate employment and output.

To compute business cycle moments from the data, I use time series of the natural logarithm of the aggregate employment, real GDP and total number of establishments that covers 1977-2015 period. The time series of the aggregate employment and the real GDP is constructed to be consistent with the timing of the BDS dataset.[77] I apply the HP filter with smoothing parameter 100 to find cycle component of the variables. The autocorrelation and the variance of the detrended time series of the aggregate employment, real GDP and total number of establishments are given in column (a) of Table 7.

I use the same methodology to compute the moments from the model simulated time-series. In particular, I run the baseline economy over large number of periods. I find the cyclical component of the natural logarithm of the simulated aggregate employment, output and total

[76] Sedlacek and Sterk (2017) and Moreira (2015) use demand side factors to explain the persistent differences in cohorts' employment. Sedlacek and Sterk (2017) allow potential entrants to choose between high-growth and low-growth business profiles after observing the aggregate demand conditions. The pro-cyclical variation in the number of entrants with high-growth business profiles generates persistent differences in cohorts employment. Moreira (2015) explains the fact through persistent demand dynamics that hinders entrants that start operating during low aggregate demand periods to catch up cohorts that start operating during high aggregate demand periods. While, in my framework, types of entrants varies across productivity which determines their life expectancy and contribution to the aggregate employment.

[77] Detailed information about the source and the construction of the aggregate variables is described in Appendix C.1.

number of firms using the HP filter with smoothing parameter 100. I use the latter time series to compute the standard deviation and autocorrelation of the variables. The statistics are described in column (b) of Table 7.

Table 7 shows that the variance and the autocovariance of the simulated total number of firms is very close to the data counterpart. The variation in exogenous aggregate demand shock has two effects on cohorts post-entry growth and survival. First, aggregate demand condition affects composition/number of entrants at entry. Second, aggregate demand affects incumbent firms decisions about production and continuation. Aggregation of these two effects by adding up cohorts at different stages of their life-cycle creates dynamics of the total number of firms that is similar to the data. The result can be interpreted as an external validation of the exogenous aggregate demand shock process.

Interestingly, a model that accounts for the life-cycle demographics (selection at entry, growth, survival) of firms generates dynamics of the aggregate employment and output that is very close to the one observed in the data. In particular, Table 7 shows that the autocorrelation of the aggregate employment in the model is 0.57 while in the data it equals to 0.61. The standard deviation in the model and in the data is 0.012 and 0.015, respectively. The result shows that firms' life-cycle demographics (selection at entry, growth and survival) over the business cycles explain major share of the business cycle fluctuations in aggregate variables. Motivated by the result, in the rest of the section I investigate the mechanism how does the firm-dynamics shapes the dynamics of the aggregate variables over the business cycles.

Propagation through variation in entry margin First, I study how much the selection of entrants at entry (rather than the post-entry choices of firms) accounts for shaping the dynamics of aggregate variables. To answer the question, I construct a counterfactual economy, where the aggregate demand shock has the same affect on the selection (composition/number) of entrants as in the baseline model. However, I set aggregate demand shocks equal to zero for all the firms that operate in the market. This assumption implies that in the counterfactual scenario, the variation of the aggregate demand affects selection but has no affect on firms' post-entry decisions. The dynamics of the counterfactual economy is summarized in column (c) of Table 7. One can see that the counterfactual scenario exhibits similar dynamics as the baseline scenario. In the baseline scenario, the aggregate demand shocks additionally affect firms' post-entry decisions about production and exit which additionally increase the variance in aggregate employment, and output. However, the minor difference between the baseline and the counterfactual scenario indicates that the post-entry shocks have minor explaining power in cohort's post-entry performance. Indicating that the composition/number selection at entry accounts for most of the business-cycle dynamics of the total number of firms, aggregate employment, and output. The mechanism is consistent with the empirical findings by Sedlacek and Sterk (2017). They find that cohorts contribution to aggregate employment is determined by selection of firms at entry stage, rather than the post-entry choice made by the firms.

Interestingly, in the scenario where the aggregate demand affects selection of entrants only through direct effect, post-entry aggregate demand shocks explain major share of the autocorrelation and variance of the economic aggregates. For example, unlike the baseline scenario, shutting down the effect of the aggregate demand shocks on operating firms in the case with $\tau = 0$, have a significant affect on the dynamics of the economy. In particular, the volatility of the aggregate employment decreases to one fourth. At the same time the autocorrelation increases from 0.40 to 0.58.[78] The result it at odds to the recent empirical findings mentioned above.

Selection through indirect effect Next, I study the role the persistent signal plays in propagating business cycle fluctuations. Toward that goal, I investigate the dynamics of the economy when the signal has zero persistence.[79] The case with $\tau = 0$ is summarized in column (d) of Table 7. One can see that shutting down the option value of delay decreases the volatility of entry almost seven times. The different variation of entrants generates significantly different dynamics of the economic aggregates. In particular, the volatility

[78] In the next section, I show that the claim is true for any level of the variation in aggregate demand. In particular, a model without persistent signal that is calibrated to generate the same set of facts as the baseline model explains the dynamics of the aggregate variables by the variation in the post-entry decision of firms over the business cycles.

[79] The difference between the case with $\tau = 0$ and the baseline economy comes solely due to the difference in composition/number of entrants, generated through option value of delay. Since shutting down option value of delay only affects selection of entrants and does not have an effect on the firms post-entry choices.

of the aggregate employment decreases by three times, while the autocorrelation decreases from 0.61 to 0.40, indicating that option value of delay significantly amplifies the effect of the aggregate demand on selection of entrants. And in turn, selection through option value of delay increases persistence in the dynamics of the aggregate variables.

Selection through indirect effect: number vs composition This paragraph shows that not just variation in the number of entrants but variation in the composition of entrants drives the contribution of entry to the aggregate dynamics.

I analyze the dynamics of the aggregate variables in the counterfactual scenarios (considered in Section 5.2) that features same variation in number of entrants as the baseline model. The counterfactual scenarios start from the selection generated in the case with $\tau = 0$ and adjust systematically low(high) productive entrants to increase the variation of the number of entrants over the business cycles.[80] Columns (e) and (f) of Table 7 summaries dynamics of the aggregate variables in the counterfactual scenarios. One can see that if the indirect effect only affects low-productive entrants variation in the aggregate employment (output) increases twice over the case with $\tau = 0$ but is still below the baseline scenario. Autocorrelation of the aggregate employment experience minor increase over the case with $\tau = 0$ and is far below the baseline case. However, if the indirect effect only affects the highest productive firms, the variation in aggregate variables increases more than 10 times over the baseline scenario (around 8 times over the data counterpart). The autocorrelation increases significantly from 0.40 to 0.65, indicating that accounting for the pro-cyclical variation in high productive firms, that creates long-lasting difference between recessionary and expansionary cohorts employment is important to explain autocorrelation in aggregate employment and output.

Combining the results from the previous paragraphs show that the indirect effect of the aggregate demand that has a disproportional effect on the selection of the low and high productive firms, significantly shapes the dynamics of the aggregate variables. Therefore, accounting not only for the variation in the number of entrants but also variation in the

[80] The scenario implies less low-productive (high-productive) entrants during recessionary periods, more low-productive (high-productive) entrants during expansionary periods compared to the case with $\tau = 0$. For more details refer to the Appendix E.2.

composition of entrants that creates the persistent differences between cohorts' characteristics over the business cycles is crucial to fully quantify the contribution of the entry margin in the business cycle fluctuations of aggregate variables.

6.2 A model without the option to delay entry

In this section I compare the baseline model performance to the existing business cycle firm-dynamics models that employ a traditional neoclassical entry decision.[81] Toward this goal, I evaluate the implications of the model without the option value of delay that is calibrated to match the same set of facts (described in Section 4.2) as the baseline model. Column (c) of Table 5 summarizes the parameter values that accomplish the goal. Note that the model without persistent signal differs from the baseline scenario with three parameters. In particular, I modify fixed entry cost to equate the threshold signal across these two models in the stochastic steady state, which ensures that the dynamics in the stochastic steady state is same across these two models. I calibrate the parameters that govern the aggregate demand shock process, ρ_z and σ_z to match the process of the simulated entry rate in the model without persistent signal to the data counterpart.

Comparing the final values of the aggregate demand shock parameters across these models helps us to quantify the role the option value of delay plays in the propagating of the aggregate demand shocks. In particular, generating the volatility of the entry rate in the model without persistent signal that is consistent with the data requires the variance of the aggregate demand shock process to be almost seven times higher compared to the baseline scenario. In particular, $\sigma_z^{\tau=0} = 0.015$ while in the baseline scenario $\sigma_z^{\tau=1} = 0.0022$.[82] The persistence of the aggregate demand shock is highly correlated to the persistence of the entry rate and ends up to be the same in both of the models $\rho_z^{\tau=0} = \rho_z^{\tau=1} = 0.57$.[83]

[81] Examples of the models are Clementi, Khan, Palazzo and Thomas (2014), Clementi and Palazzo (2016), Siemer (2016), Lee and Mukoyama (2018) Lee and Mukoyama (2008).

[82] Volatility of the entry can be also achieved by creating steeper distribution of entrant around the steady state threshold signal, for lower variation of aggregate demand shocks. Though, the procedure increases variation in entry rate but it does not generate significantly different composition of entrants over the business cycles. As we show, variation in composition of entrants is important to explain the propagation of the variation in entry margin.

[83] To be more precise, for $\rho_z^{\tau=1} = 0.57$ persistence of the entry rate is higher in the model without persistent signal compared to the data counterpart. However, for the sake of comparison I let the persistence of the

In the model without persistent signal (hence in the standard firm-dynamics models) potential entrant decides to enter into the market if the expected lifetime value from entry net of the fixed entry cost (NPV) is more than zero.[84]

The choice of the exogenous aggregate demand process, in turn, implies counterfactual (excessive) variance of the aggregate output and employment which one can see in column (c) of Table 8. Increased variation in economic aggregates also significantly decreases autocorrelation in aggregate employment and output. Interestingly, the autocorrelation of the aggregate employment in the model without persistent signal is very close to the case with $\tau = 0$ that features almost 7 times less volatile entry rate.

To further analyze what determines the dynamics of the aggregate variables over the business cycle consider column (d) of Table 8. The column describes the dynamics of the aggregate variables in the model without persistent signal, when aggregate demand affects selection but not post-entry dynamics of firms. I find that the variation in the aggregate demand accounts for the major share of the variation in cohort level employment. Increased volatility of the aggregate demand shock increases variation in entry, but at the same time it increases variation in the employment of the firms in the market. As a result, post-entry decisions made by firms explains major share of the variation in cohort level employment. In turn, post-entry decisions made by firms explains major share of the dynamics in the aggregate employment and output. While in the baseline model the dynamics is explained by the composition/number selection of entrants at entry.

To conclude, the standard business-cycle firm-dynamics models rely heavily on the high variance of the aggregate shocks. Additionally, in the model with persistent signal option value of delay provides significant internal propagation mechanism of the aggregate demand shocks. As a result, augmenting entry decision with option value of delay enables standard firm-dynamics models to produce observed volatility of the entry rate without generating

aggregate demand shock to be the same across models for two reasons: Firstly, the required decrease is minor. Secondly, decreasing the persistence of the aggregate demand shock not only does not impair analyzes of the section, but also strengthens the claim that the model without persistent signal generates counter-factual persistence in economic aggregates.

[84]For detailed discussion see Section 3.

counterfactual volatility in aggregate variables.[85]

7 Implications of entry dynamics

During the Great Recession, the number of entrant establishments dropped by 23.2 percent compared to the 2004-2006 period average. After six years (March 2014), the number remained 15.6 percent below the same pre-crisis average. In 2010, aggregate employment was 4.8 percent lower compared to March 2008 and took more than 6 years to achieve the same level of employment.[86] Figure 1 plots the evolution of the number of entrant establishments and the total employment in the U.S. over the 1977-2015 period. The figure highlights the historical drop in entry and unprecedented slow recovery of aggregate employment during the Great Recession.

Motivated by this comovements, I use the model to study causal relationship between the change in entry margin and slow recovery of the aggregate variables. Consistent with the data, entrant cohort employ only 6% of the total employment and the share decreases as the cohort ages (see Figure 22(d)). Nevertheless, I find that change in the composition of entrants at entry have a long-lasting effect on the dynamics of total employment and output. If the drop in entry rate is persistent, the cohort effect accumulates and has a substantial effect on the depth and the long-run recovery of economic aggregates. After investigating the mechanism, I investigate how the model explains the depth and the slow recovery in the aggregate variables observed during and after the Great Recession. Lastly, I quantify how much the persistent drop in number of entrant establishments over 2007-2015 period contribute to the depth and slow recovery observed after the Great Recession.

[85] Even in general equilibrium settings the model with persistent signal performs at least as good as standard firm-dynamics models. The reason is as follows. The option value of delay is always non-negative due to entrants ability to get outside option by not entering into the market. As a result, for any initial aggregate states the threshold value of entry is weakly higher in the model with persistent signal compared to the models without persistent signals.

[86] Statistics are computed using U.S. economy-wide establishment level data from Business Dynamic Statistics (BDS) dataset, covering 1977-2014 period.

7.1 Amplification and propagation of aggregate shocks

The following section investigates the mechanism how the change in the composition of entrants as a response of a negative aggregate demand shock, propagate the effect of the shock.

First, I investigate the response of the economy to a one-time negative demand shock described in Figure 25. The magnitude of the shock is chosen to yield 25% decline in number of entrant establishments observed during the Great Recession. The path for the aggregate demand process and number of entrants are given in Figure 26(a) and Figure 26(b), respectively. At impact entry rate drops by 2.7*pp* as shown in Figure 26(c), which is also close to the decline observed in the data.

Figure 26(d) shows that total number of firms decreases by 2.9% at the time of the shock. Total number of firms takes three years to recovery 50% of the drop and 8 years to recovery 75% of the drop. The persistent decline can be explained by the initial reduction in entrants and increased exit rate due to the lower aggregate demand. The former accounts more than 90% of the persistence.

Figure 26(e) and Figure 26(f) display the path for the total employment and aggregate output. Combination of the lower aggregate demand level and lower number of entrants translates into 1.8% and 2.0% immediate drop in aggregate employment and output, respectively. To evaluate the importance of the entry margin on the propagation of the aggregate shocks columns (a)-(b) of Table 9 summarizes the key features of the response of the economy to a one-time aggregate demand shock for two scenarios.[87] Column (a) describes the full response of the baseline economy. Column (b) describes the response of the baseline economy while the number/composition of entrants are fixed at the stochastic steady state level. The differences between these two scenarios are fully driven by the change in composition/number of entrants. Comparing column (a) to column (b) shows that change in composition/number of entrants accounts 60% of the initial drop in total employment, while the rest comes from the decline in incumbent firms employment.[88] After a one-time negative aggregate demand shock, the

[87]The discussion concentrates on the dynamics of the employment since the path of the aggregate output closely mimics to the path of the aggregate employment.

[88]In the table 'depth' describes the highest deviation of the interested variables from trend

baseline economy takes three years to recover half-life, and an extra 12 years to recover additional 25% of the decline, whereas an economy in which the shock does not affect the entry margin takes only two years to recover two-thirds of the decline. Further examination shows that the significantly protracted recovery for the additional 25% of the decline is due to a small number of high productivity entrants that choose to delay entry as a response to the shock. Due to the latter effect the recessionary cohorts' employment decreases persistently and that in turn has a long-lasting effect on the decline in aggregate employment.

To make sure that quick recovery of the economy without selection described in scenario (b) is not a repercussion of the initial lower drop in employment, column (c) of table 9 describes response of an economy to a one-time shock that generates the similar magnitude drop in total employment as the baseline model without affecting entry margin.[89] Comparing column (c) to column (a) and column (b) shows that when entry margin is not affected, economy recovers quickly and recovery takes almost same number of years irrespective of the magnitude of the shock.

The importance of the change in entry margin on aggregate dynamics significantly increases when the change is persistent. Second half of Table 9 reports the response of the economy to a persistent shock of the original magnitude. Comparing column (d) to column (a) shows that change in composition/number of entrants accumulates over time and the depth is reached after four years. Even though the depth is just 8% higher, the recovery up to 50% takes 5 times more (16 years) compared to the case with one-time change in entry margin. With the persistent drop in entrants, recovery up to 75% of the depth doubles and happens in 28 years. To conclude, entrants contribute only 6% to aggregate employment. However, if the shock persistently effects number/composition of entrants, the contribution of entry margin to aggregate employment accumulates over time and has a substantial effect on the long-run recovery of employment.

After analyzing the importance of the entry margin on the propagation of the aggregate shocks, I investigate what accounts for the selection of entrants at entry and whether the

[89] The magnitude of the negative shock necessary to generate same magnitude drop in employment without changing entry is 7 standard deviation of the respective exogenous aggregate demand shock process.

effect of the entry margin on employment is driven by change in number or the composition of entrants.

Aggregate demand shock affects entry decision through directly affecting the market profitability today and affecting relative profitability of the market today relative to tomorrow through option value of delay. Comparison of the baseline model dynamics to the case with $\tau = 0$ shows that option value of delay significantly amplifies the effect of the aggregate demand on the selection of the cohort at entry. Figure 26(b) shows that for the described one-time aggregate demand shock, 2.5% drop in entrants is due to the direct effect of the shock, while the rest 22.5% of the drop comes from the entrants that delayed entry into the market. Interestingly, even though the drop in number of entrants is 10 times higher in the baseline scenario compared to the case with $\tau = 0$, the persistent effect of the change in entry margin on aggregate dynamics comes through the small number of high-productive entrants that decided to postpone entry into the market.

To analyze the importance of the composition of entrants versus number of entrants, I consider two counterfactual selection of entrants that generate same magnitude drop in number of entrant establishments as the baseline model as a response of the one-time aggregate demand shock. The key difference between these two scenarios come from the variation in types of entrants. In particular, I generate 25% drop in number of entrants, by additionally cutting 22.5% of the lowest (highest) productive entrants from the entrants distribution in the case with $\tau = 0$.[90] Columns (c) and (d) of table 10 summarizes the dynamics of the economies in the described scenarios. Contrasting column (c) and (d) to column (a) shows that the major effect of the option value of delay comes through changing the composition and not the number of entrants. In particular, if aggregate demand conditions at entry only affects and reduces low-productive firms then the counterfactual evolution of the aggregate employment mimics the case with $\tau = 0$ closely, despite the fact that in the counterfactual scenario 10 times less number of firms enters into the market (see Figure 26(e)). Due to lower number of entrants (extensive margin) initial drop (depth) in employment increases by 1.6 times relative to the case with $\tau = 0$, however, the recovery period takes same number

[90] For more details refer to the appendix E.2

of years as in the case with $\tau = 0$. Aggregate employment recovers up to 50% and 75% of the depth respectively in 2 and 5 years. Figure 26(e) shows that if the aggregate demand at entry reduces only high productive entrants, the effect significantly amplifies and propagates one time aggregate demand shock, which shows that high-productive entrants are the ones that are the important to explain the contribution of entry margin on the long-run dynamics of the aggregate variables. The selection generated by option value of delay is in between of the above described scenarios. From the additional 22.5% of the entrants that do not go into the market, the propagation happens due to the small share of the high productive firms that decided to postpone entry into the market due to low aggregate demand conditions.[91]

To sum up, the aggregate shock through option value of delay affects the composition of entrants which creates long lasting effect of the entry margin on the aggregate dynamics.

7.2 The Great Recession

Next, I evaluate the model performance in explaining the depth and the slow recovery in aggregate variables observed during and after the Great Recession.

Using the linear trend over 1991-2007 period, I predict the evolution of the aggregate variables starting from 2008 as if the Great Recession had not happened.[92] Figure 28(a) and Figure 28(b) show the evolution, pre-crises trend and the prediction respectively for the number of entrant establishments and total number of establishments over the period 1991-2015. Figure 28(c) and Figure 28(d) show the same information for the total employment and the real GDP over the period 1991-2018. I explain the deviation of the actual time series of the variables from the predicted trend as the response to the unobserved shock process that hit

[91]The claim of the exercise is robust to the case when the entry margin changes persistently. Table 11 summarizes dynamics of the aggregate variables as a response of the persistent shock process. As one can see, persistent change in low-productive entrants still do not have a persistent effect on aggregate employment. In fact economy recovers faster than in the case with $\tau = 0$. However, in the scenario when economy looses persistently high-productive firms recovery is highly correlated with the number of entrants. Figure 26 shows that selection due to the option value of delay is in between these two scenarios.

[92]The National Bureau of Economic Research (NBER) dates the beginning of the Great Recession as December 2007. In the BDS the year 2007 characterizes establishment level activity from March 2006 to March 2007. To be consistent with the NBER, I choose year 2008 as the beginning of the Great Recession. Year 2008 in our sample covers economic activity from March 2007 to March 2008. To remind the reader the Total Employment and the real GDP data used in the paper is constructed to be consistent with the BDS timing.

the U.S. economy over the 2008-2015 period.

To uncover the unobserved shock process, I construct the aggregate demand process that matches change in the model simulated number of entrant establishments to the data counterpart over 2008-2015 period.[93] Figure 30(a) illustrates the evolution of the number of entrant establishments in the model and in the data and Figure 30(b) displays the latent process for the aggregate demand that generates the match.[94] Note that I set shocks equal to zero starting from year 2016, which means that in the following exercise, number/composition of entrants are fixed at the stochastic steady state level starting from year 2016. That enables me to exclusively identify the role of the observed decline in number of entrants over 2008-2015 period in the slow recover of the aggregate variables after 2015.

Figure 29 describes response of the aggregate variables in the baseline economy to the constructed aggregate demand process.

Figure 30(e) shows that the change in the simulated number of firms over 2008-2015 quite closely follows to the data counterpart. Which again emphasize the fact that the baseline parametrization of the model and constructed aggregate demand shock process generates data consistent demographics (entry, survival, exit) of the establishments. Figure 30(f) shows, that even after 15 years, total number of establishments will stay 2% (20% of the depth) below trend due to the combination of the low aggregate demand level and persistent decline in number of entrant establishments.

[93]Matching the dynamics of the number of entrants to quantify the role of the entry margin in the dynamics of the aggregate variables can be justified due to the following reasons. Firstly, empirical findings by Clementi and Palazzo (2016), Gourio, Messer and Siemer(2016) and Sedlacek and Sterk (2017) show that during the Great Recession entry was affected more at extensive margin (number of entrants) rather than the intensive margin (average size of entrants relative to incumbent firms). Secondly, using cross-state (cross-MSA) data Gourio, Messer and Siemer (2015) show that the states (MSAs) that experienced higher drop in entry rate over the period 2008-2010 experienced slower recovery over the period 2010-2013, even after controlling for house prices and leverage. Besides, the existing literature also points out that the change in the composition and the structure of entrants during the Great Recession might have played important role in the protracted recovery in aggregate variables. In the exercise matching the extensive margin of the entrant dynamics also implies change in the composition of entrants, and qualitatively the change is consistent with the average recessionary cohorts behavior in the data. To argue against structural change, Figure 28 shows that all sectors experienced significant and persistent drop in the number of entrants compared to pre-crisis level (similar argument is used in Gourio, Messer and Siemer(2016).

[94]Simulating the economy for the constructed shock process shows that the simulated entry rate matches quite closely to the data counterpart too, see Figure 30(c).

Contrasting the evolution of the aggregate employment from the data to the simulated series over 2008-2018 period shows that the model explains around 60% of the depth reached in 2012. Figure 30(g) shows that the negative effects of the persistently low aggregate demand level and change in composition/number of entrants accumulates over time and by 2015 the model explains 95% of the deviation in aggregate employment. Starting from 2015 up to 2018 (the latest available data), the model prediction about the magnitude and persistence of the recovery is close to the data counterpart. Further examining Figure 30(h) shows that the model predicts half-life recovery of the aggregate employment will happen in year 2024, in 9 periods from the depth. For additional 25% recovery, the economy needs additional 11 years. More than 80% of the deviation is driven by the entrants that delayed entry.

Repeating the same exercise for the aggregate output shows that again the model explains quite well persistence in the evolution of the aggregate output see figure 30(i). Long-run evolution of the post-Great Recession output is depicted on figure 30(j). The model predicts that 50% recovery in output happens in 2025.

The exercise shows that the model that accounts for the firm demographics observed in the data is capable to fully explain the slow recovery observed after the Great Recession. However, the results also indicate that without modeling any additional frictions, the combination of the persistently low aggregate demand and the observed entrant demographics are not enough to explain the depth of the Great Recession.

7.2.1 Lost generations of entrants and the slow recovery

In the baseline scenario, the slow recovery of the aggregate variables as a response to the constructed shock process combines two effects: change in the number/composition of entrants and change in incumbent firms decisions (about employment, production and exit) as a response to the persistently low aggregate demand level. To uniquely identify the contribution of the dynamics of entrant establishments over $2008-2015$ period, I consider the following counterfactual. Using the constructed shock process from the previous section, I simulate the economy over $2007-2030$ period, while keeping the number/composition of entrant firms at the stochastic steady state level. Figure 32(b) contrasts the evolution of the number of entrant establishments in the data, the baseline model and the counterfactual scenario. I use the counterfactual to answer the following question: what would have been the dynamics of the aggregate variables after the Great Recession if entry rate stayed at the pre-crsis level?

Figure 31 describes evolution of the aggregate variables in the data, baseline model and counterfactual scenario.

Comparing counterfactual dynamics of the aggregate employment to the baseline scenario given in Figure 32(e) shows that dynamics of entrants explains 76% of the total deviation of employment from the stochastic steady state in year 2012. Over time, the contribution increases and by 2018 the persistent change in entry margin explains 84% of the deviation. Further examining the figure shows that the persistent dynamics of the entrant establishments over 2008-2015 period implies that aggregate employment stays 2% below trend by 2030. The number corresponds to 30% of the depth in employment reached in 2015 in the baseline scenario.

Comparing the dynamics of employment in the counterfactual scenario to the data counterpart shows that the persistent decline in the number/composition of entrant establishments explains only 45% of the depth of the aggregate employment reached in 2012. By 2018, the contribution of the persistent drop of entrants accumulates and it explains 75% of the deviation from trend of the aggregate employment.

To conclude, change in number/composition of entrants observed during the Great Recession contributed significantly to the slow recovery observed after the Great Recession.

8 Policy implications

In the section, I show that the option to choose time of entry has an important implications on firm demographics outside business-cycle dynamics too. In particular, the ability to delay entry implies qualitatively and quantitatively different responses of potential entrants to a change in the expected value of entry, depending on the timing, duration, and the magnitude of the change. This result indicates that accounting for the option to delay entry has potentially important policy implications.

To demonstrate the idea, I consider response of potential entrants to the contemporaneous/future, temporary/permanent, gradual/abrupt decrease in fixed entry cost. Note that decrease in the fixed cost of entry affects only the opportunity cost of entry and does no have an effect on the expected stream of profits after entry. Still, as I show below, the variation in the timeline of the decline translates into rich dynamics of potential entrants.

8.1 Permanent decline in the fixed cost of entry

I begin by analyzing the response of potential entrants to a permanent reduction of entry barriers when the policy is initiated across different aggregate states.

The permanent decrease in fixed entry cost affects opportunity cost of entry in two counteracting ways. First, a decrease in fixed cost of entry today directly increases net expected value of being an incumbent (gross value of entry minus fixed entry cost) and encourages entry into the market.[95] Secondly, the policy indirectly increases option value of delay due to the permanent decrease in entry cost which discourages entry. Figure 33(a) shows that in equilibrium, the former (direct) effect dominates and the policy decreases threshold quality of the initial productivity signal; hence, increases number of entrants in all aggregate states.[96] One can see the effect by comparing black solid line that depicts threshold signal for the baseline economy to the blue dash-dot line, that illustrates threshold signal for the economy with lower fixed entry cost.

Potential entrants that respond to the permanent policy in different states of the business cycles are qualitatively and quantitatively different.

During expansionary periods, the minimum level of the signal that a potential entrant requires to enter into the market is very low. Potential entrants with signals around the threshold signal put higher weight on the first period profit, enter only during high aggregate demand periods and does not respond to the opportunity to delay entry into the market. As a

[95] Note that since the option to exercise the signal by entering the market is irreversible, reduction of entry cost does not affect gross expected value of being an incumbent and as a result decrease in entry cost translates into one to one increase in net expected value of being an incumbent.

[96] The figure describes a response of the potential entrants after permanently decreasing the fixed entry cost. The policy modifies net expected value of being an incumbent, thus, I recalculate entrants value function and find implied threshold signals for the new fixed entry cost.

result, during expansionary periods, the effect of the permanent decrease in fixed entry cost decreases threshold signal only through directly increasing net present value of entry and has no indirect negative effect through option value of delay. A major share of those additional firms that decide to enter into the market due to permanent reduction of the fixed entry cost end up being low productive firms. These firms have relatively low survival rate and maximize lifetime value through charging full monopolist prices, instead of investing in future customer capital.

Next, consider the permanent decrease in entry cost during recessionary periods. During low demand periods potential entrants that decide to enter into the market are endowed with relatively higher productivity signal levels with higher option value of delay. The permanent decrease in entry barriers encourages entry through direct increase in net entry value but at the same time discourages entry through indirect increase in option value of delay. Quantitatively, due to the latter effect, permanent policy has less significant effect on the decrease in threshold signal during recessionary periods compared to the expansionary periods (or in the case when signal has zero persistence). Intuition behind the result is the following. Permanent reduction of fixed entry cost does not affect potential entrants that choose to postpone entry into the market since they can still wait for the better aggregate demand conditions and at the same time, take advantage of the lower level of the fixed entry cost in the future periods. Qualitatively, in contrast to the expansionary periods, a major share of the additional potential entrants that decide to enter after the policy end up becoming high productivity firms, which invest in future customer capital at the expense of setting lower than monopolist prices today.

Figure 33(b) translates the threshold signal into the number of entrants using the assumed distribution $W(q)$ of potential entrants. The black solid line reports initial number of entrants for each aggregate state. The blue dash-dot line displays number of entrants after the permanent decrease in entry barriers across aggregate states. Comparing change in the number of entrants during expansionary and recessionary periods illustrates the quantitatively different response of entrants to the policy.

8.2 Temporary decline in the fixed cost of entry

Next, I analyze how a temporary decrease in fixed entry cost affects potential entrants' entry decisions. Contrary to the permanent policy, temporary decrease in fixed entry cost affects the opportunity cost of entry only through directly increasing net expected value of entry today. The temporary policy has no effect on the option value of delay, since potential entrants face same level of fixed entry cost tomorrow. The red dash line on figure 33(a) shows that the temporary decrease in fixed entry cost decreases threshold productivity signal and increases number of entrants for all aggregate demand levels.[97]

Further, examining figure 33(b) shows that the temporary policy has relatively larger effect on the number of entrants during recessionary periods compared to the permanent policy. The intuition behind the result is as follows. During recessionary periods, the entrants know that if they delay entry the market, they do not receive an advantage from the lower entry cost tomorrow. Thus, significant share of the potential entrants who previously decided to postpone entry the market, does not find waiting optimal any more. The potential entrants immediately enter into the market if the decline in entry cost offsets decline in the expected value of entry due to low initial aggregate demand level. During expansionary periods, a temporary reduction has similar effect as a permanent reduction of fixed entry cost. The reason is as follows. The marginal potential entrant during expansionary periods holds low productivity signals and their entry decision are not affected by option value of delay. As a result, both of the policy effects potential entrants decision only through directly affecting net expected value of entry.

To sum up, if a goal is to increase the number of entrants, temporary decline in fixed entry cost does a better job during recessionary periods and has same affect during expansionary periods compared to a permanent decline in fixed entry cost. Also, marginal entrants who respond to the reduction of fixed entry cost are mostly high productive firms during recessionary periods and low productive firms during expansionary periods.

[97] The exercise analyzes how the decline in entry cost effects entry decision, however unlike the previous (permanent) case the decline happens one-time, only in the current period. In that case, I just modify net expected value of being an incumbent and reevaluate entry decisions using the option value of delay from the baseline case and net expected incumbent value after accounting for the decline in entry cost.

8.2.1 Comparing results across models

Before moving forward, I also check how the permanent and temporary reduction of entry barriers affects entry decision in a standard firm-dynamics models by setting persistence of the signal to zero. Figure 33(c) and figure 33(d) illustrate the effect of the permanent and temporary reduction of entry barriers on the threshold signal and number of entrants for the case when signal has zero persistence. When $\tau = 0$ option value of delay equals to zero. Since the option to exercise the signal by entering into the market is irreversible, reduction of entry cost does not affect gross expected value of being an incumbent and as a result decrease in entry cost translates into one to one increase in net expected value of being an incumbent. As a result, permanent and temporary reduction of entry barriers has the same effect on the threshold quality of signal and number of entrants irrespective of the level of the aggregate demand. Since the quality of the signal during expansionary and recessionary periods are not significantly different in the case with $\tau = 0$, the decline of entry cost implies that potential entrants that responds to the policy are not qualitatively different after entry.

8.3 Anticipated decline in the fixed cost of entry

Persistent signal implies that potential entrants respond not only to contemporaneous changes but also to anticipated changes in the fixed entry cost that is going to materialize in the future.

To understand the mechanism consider the following exercise. At time zero the economy starts with aggregate demand level z. In the beginning of period zero agents also learn that the fixed entry cost will decline by 0.1 unit in the beginning $N(> 0)$ periods later. As we saw earlier, a permanent decrease in the fixed entry cost in period N is going to decrease the threshold signal, and hence is going to increase the number of entrants in all aggregate states.

In contrast to the actual decline in entry cost the news about the permanent decline in the fixed entry cost increases threshold signal which decreases number of entrants at time zero. The reason is as follows. Actual decrease in the entry cost increases net expected value of being an incumbent (gross value minus fixed entry cost) and as a result increases value of entry in the beginning of period N. Increase in entry value in the beginning of period N

translates back and (weakly) increases the option value of delay at time zero. The news and the actual decline in the entry cost does not effect incumbent firms decisions and value functions, so it does not affect potential entrants net value of being an incumbent at time zero. As a result, at time zero, the news encourages some of the potential entrants to differ entry decision up to the period N through indirectly increasing option value of delay.

Figure 33 depicts the response of the potential entrants at time zero to the anticipated decline in fixed entry cost for $N = 1, 2$, and 5.[98] The figure shows that for any N, threshold signal in news scenario is weakly higher compared to the baseline (no-news) scenario in all aggregate states. The figure also shows that the magnitude of the change depends on the distance between today and the actual time of the policy.

The closer the distance between today and the actual time of the policy (small N), the higher the effect of the future increase in entry value on today's option value of delay and the higher the response of potential entrants in all aggregate states. In particular, Figure 33 shows that if $N = 1$, threshold potential entrants always prefer to delay entry, irrespective of the aggregate state. However, when $N = 5$, the news has a weaker effect on the potential entrants entry decision and the magnitude of the effect depends on the aggregate demand level at time zero. For example, the news has no effect on the threshold potential entrant during recessionary periods. Those potential entrants are relatively high productive entrants which does not find it optimal to wait for five periods to get advantage of lower fixed entry cost. When $N > 1$, potential entrants do not find it optimal to wait during expansionary periods too, since threshold potential entrants are low-productive ones and their entry decision significantly depends on the aggregate demand level. As a result, waiting more than one period increases the risk that they might loose the entry value today and not be able to advantage the lower entry cost in the future if the aggregate demand level decreases. Potential entrants around the steady state, who can be counted as a medium range of signals are the ones again responding the policy in most of the cases.

Interestingly, if the time of the actual decrease in entry cost is close enough (small N), the

[98] Higher the decline in fixed entry cost in the future, higher the effect of the news on potential entrants entry decision and bigger the number of entrants that decide to differ entry today.

indirect effect of the news that decreases entrants today is quantitatively more significant than the increase in number of entrants at time N in response of the lower fixed entry cost.

8.3.1 News effect and aggregate dynamics

Change in number/composition of entrants directly affects the dynamics of total number of firms, aggregate employment and output. As a result, the model produces aggregate co-movement in response to news about change in fixed entry cost in the future, even without changing the aggregate demand level.

Figure 34 describes response of the baseline economy to an announcement, timing of which is as follows. The economy starts at the stochastic steady state at time zero. At time one, unanticipated announcement happens. Potential entrants learn that there will be 0.05 unit decline in the fixed entry cost beginning five periods later.[99] The level of the fixed entry cost is described in Figure 36(a). The aggregate demand series is described in Figure 36(b). Note that aggregate demand level is fixed at $z = 1$, which lets us to exclusively identify the effect of the news shock on the dynamics of aggregate variables.

Consistently with the previous discussion, as a response of the news, potential entrants start postponing entry the market starting from period one. The evolution of the number of entrant firms is given in Figure 35(c). Starting from period one up to period 4 the number of entrants decreases and the magnitude of the decrease increases as the economy gets closer to the actual time of the decline of the entry cost. At period 5 number of entrants increases since entry into the market is less expensive. But the increase in number of entrants is significantly smaller than the total number of entrants that delayed entry as a response of the news. If we allow return of the accumulated potential entrants that delayed entry than number of entrant increases significantly in period 5. At period 6, increased number of entrants only reflects potential entrants that decided to enter into the market due to lower entry cost.

The change in the total number of firms as a response of the news is given in Figure 35(d). Since aggregate demand is fixed at the steady state level change in the total number of firms

[99]The magnitude of the decline is chosen to result around 3.0 percent decline in entry rate, highest decline that was observed during the Great Recession.

reflects changes in the entrant firms. Starting from period one, the number of firms decreases. At period 5, the number of firms increases due to increase in the number of entrants that decide to enter into the market after the decline in entry cost, and also due to the entrants that delayed entry into the market and returned in period five. Starting from period 5, the number of entrants slowly goes back to the stochastic steady state level.

Finally, Figure 35(e) and Figure 35(f) show the dynamics of aggregate employment and output, which echoes dynamics of the total number of firms. The decrease in total number of firms from period one to 4 decreases employment as well as output up to period 5 and creates effect of recession. Starting from period 5, with the increase in number of firms, employment and output increases which in turn creates the effect of a boom. Moreover, if we allow return of the entrants that delayed entry in previous periods we see that after the boom in period 5 the economy starts to go back in new steady state and which again creates the effect of a recession.

To sum up, the model implies that even in the stochastic steady state, news shock can create co-movements in aggregate employment and output through the option value of delay.

9 Conclusions

In the dissertation I investigate what accounts for the documented significant effect of the initial entry conditions on the selection of entrants over the cycles, and how does the observed life cycle demographics of entrants shape the aggregate fluctuations.

In the first part of the dissertation, I show that potential entrants ability to delay entry, previously ignored by the existing firm-dynamics literature, accounts for the significant effect of the aggregate conditions on selection of entrants. Procyclical variation in the expected survival rates moderate the relationship: during recessions increased risk of post-entry failure creates positive *value of waiting* and increases relative cost of entry today on top of fixed entry cost. The mechanism generates group of potential entrants who decide to postpone entry even if entry promises positive life-time value. I show that for reasonable parameter values (i) the total cost of entry during recessions increases by 7% for medium productivity entrants, (ii) the seemingly small increase in the entry cost generates delays that could last from 1 to 8 years. The expected duration of delay is negatively correlated with entrants' productivity level. At an aggregate level the mechanism produces a countercyclical cost of entry which leads to the documented significant selection of entrants over the cycles. I show that without the ability to delay entry, existing firm dynamics models that rely on traditional entry decision rule and fixed entry cost require counterfactually large variance of the aggregate demand shock process to reconcile the observed facts.

In the second part of the dissertation, I investigate what accounts for the observed characteristics of cohorts over the cycles. I calibrate the model developed in previous section to U.S. establishment level data. The calibrated model matches average cohorts' post-entry survival, growth and employment share. The model is able to account for the observed dynamics in entry rate and the documented persistent and significant differences in cohorts' life-cycle characteristics. I find that the option value of delay channel is quantitatively and qualitatively important for generating the results. Due to the option value of delay the opportunity cost of entry almost doubles during recessions, generating cohorts of entrants with countercyclical average productivity. The medium productivity entrants who delay entry contribute to persistent procyclical variation in cohort level employment, since they represent high-growth and high-survival firms. I find that the ability to delay entry allows potential entrants to choose time of entry that leads to higher profits for longer periods of time, leading to countercyclical survival rate. A model without the option value of delay implies acyclical average survival rate.

In the third part of the dissertation, I quantify the role the observed demographics of entrants play in shaping aggregate fluctuations. I find that a calibrated model that accounts for the life-cycle dynamics of U.S. establishments produces business-cycle fluctuations in aggregate employment and output that are similar to those observed in the data. Moreover, I show that the variation in the number and the composition of entrants at entry, rather than the post-entry shocks, is responsible for generating the observed persistence and variance of aggregate variables. Using the model, I re-examine the causal relationship between the persistent and significant drop in the number of entrants and the slow recovery observed after the Great

Recession. A counterfactual exercise shows that if the entry rate had stayed at the pre-crisis level, the drop in aggregate employment would have been 45 percent lower, and the economy would have recovered two times faster.

In the final part of the dissertation, I show that accounting for the option to delay entry has important policy implications and could significantly alter the existing firm-dynamics models' predictions about the response of potential entrants to various policies. In a model with the option value, a full effect of a policy depends on its effect on entrants' relative cost of entry today versus tomorrow, while the traditional entry decision rule accounts only for the direct effect of the policy. Depending on the magnitude of an indirect effect, these two specification could imply quantitatively and qualitatively different responses of potential entrants depending on the timing, duration, and the magnitude of policies. For example, I show that existing models predict that a temporary or a permanent decrease in entry cost have a same effect on the number of entrants, whereas a model with the option value predicts that temporary decrease in entry cost is more effective in increasing the number of entrants during recessions since it decreases relative cost of entry more than the permanent policy. The main takeaway is that while designing policies that intend to affect potential entrants one needs to take into account very carefully indirect effects implied by potential entrants ability to delay entry.

References

[1] Arrow, K., J., 1968. Optimal capital policy with irreversible investment. Institute for Mathematical Studies In the Social Sciences, Stanford, California. Technical report no. 146, 1–19

[2] Barlevy, G., 2001. The sullying effect of recessions. Review of Economic Studies (2002) 69, 65-96

[3] Bayard, K., Dinlersoz, E., Dunne, T., Haltiwanger, J., Miranda, J., and Stevens, J., 2018. Early-stage business formation: an analysis of applications for employer identification numbers.

[4] Bernanke, B., S., 1993. Irreversibility, Uncertainity and Cyclical Investment. The Quarterly Journal of Economics, 98 (1), 85–106,

[5] Bilbiie, F., Ghironi, F., and Melitz, M., 2012. Endogenous Entry, Product Variety, and Business Cycles. Journal of Political Economy, 120(2), 304-345.

[6] Clementi, G., Khan, A., Palazzo, B., Thomas, J., 2014. Entry, exit and the shape of aggregate fluctuations in a general equilibrium model with capital heterogeneity.

[7] Clementi, G., Palazzo, B., 2016. Entry, exit, firm dynamics, and aggregate fluctuations. The American Economic Journal: Macroeconomics, 8(3), 1-41

[8] Cogley, T., Nason, J. M., 1995. Output dynamics in real-business-cycle models. The American Economic Review, 85(3), 492-511.

[9] Constantini, J., Melitz, M., 2007. The dynamics of firm-level adjustment to trade liberalization.

[10] Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J., 2014. The role of entrepreneurship in US job creation and economic dynamism. Journal of Economic Perspectives, 28(3), 3-24

[11] Dixit, A., Pindyck R., (1994). Investment under uncertainty. Princeton University Press, Book.

[12] Fajgelbaum, P., Schaal, E., and Taschereau-Dumouche, M., 2017. Uncertainty traps. The Quarterly Journal of Economics, 132(4), 1641–1692

[13] Folta, T., B., and O'Brien J., P., 2004. Entry in the presence of dueling options. Strategic Management Journal, 25, 121-138

[14] Fort, T., Haltiwanger, J., Jarmin, R., and Miranda, J., 2013. How firms respond to business cycles: the role of firm age and firm size. IMF Economic Review. Vol 61(2)

[15] Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, firm turnover, and efficiency: selection on productivity or profitability? American Economic Review, 98 (1), 394–425

[16] Foster, L., Haltiwanger, J., C., Syverson, C., 2016. The slow growth of new plants: Learning about demand. Economica, 83(329), 91-129

[17] Geroski., P., E., 1995. What do we know about entry? International Journal of Industrial Organization, 13, 421-440

[18] Gourio, F., Messer, T., Siemer, M., 2016. Firm Entry and macroeconomic dynamics: a state-level analysis. The American Economic Review, 106(5), 214-218

[19] Gourio, F., Messer, T., Siemer, M., 2015. A missing generation of firms? Aggregate effects of the decline in new business formation.

[20] Haltiwanger, J., Jarmin, R., and Miranda, J. 2013. Who creates jobs? Small versus large versus young. The Review of Economics and Statistics, 95(2), 347-261.

[21] Hong, S., 2018. Price markups for small and large firms over the business cycle. Federal Reserve Bank of St. Louis Economic Synopses, 2018, No. 17

[22] Hopenhayn, H., A., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica, 1127−1150

[23] Hopenhayn, H., Rogerson, R., 1993. Job turnover and policy evaluation: a general equilibrium analysis. Journal of Political Economy. 101, 915–938.

[24] Jarmin, R., and Miranda, J., 2002. The Longitudinal Business Database. CES Working Paper, US Census Bureau.

[25] Jovanovic, B., 2008. Investment options and business cycle. Journal of Economic Theory, 144, 2247-2265

[26] King, R. G., Rebelo, S. T., 1999. Resuscitating real business cycles. Handbook of macroeconomics 1, 927-1007.

[27] Lee, Y., Mukoyama, T., 2015. Entry and exit of manufacturing plants over the business cycle. European Economic Review, 77, 20-27

[28] Lee, Y., Mukoyama, T., 2008. Entry, exit and plant-level dynamics over the business cycle.

[29] Lee, Y., Mukoyama, T., 2018. A model of entry, exit, and plant level dynamics over the business cycle. Journal of Economic Dynamics & Control.

[30] McDonald, R., and Siegel, D., 1986. The Quarterly Journal of Economics, 101(4), 707-728

[31] Moreira, S., 2015. Firm dynamics, persistent effects of entry conditions and business cycles.

[32] O'Brien, J., P., Folta., T., B., Johnson D., R. 2003. A real options perspective on entrepreneurial entry in the face of uncertainty. Managerial and Decision Economics, Vol. 24, No.8, pp. 525-533

[33] Pindyck, R., S., 1991. Irreversibility, uncertainty, and investment. Journal of Economic Literature, 29(3), 1110-1148

[34] Pindyck, R., S., 2009. Sunk costs and risk-based barriers to entry. Draft: February 13, 2009

[35] Pugsley, B., Sedlaćek, P., Sterk, V., 2017. The Nature of Firm Growth.

[36] Samaniego, R., M., 2008. Entry, exit and business cycles in a general equilibrium model. Review of Economic Dynamics, 11, 529–541

[37] Schoar, A., 2010. The divide between subsistence and transformation entrepreneurship. Innovation policy and the economy. National Bureau of Economic Research, 10, 57-81.

[38] Sedlaćek, P., Sterk, V., 2017. The Growth Potential of Startups over the Business Cycle. American Economic Review, 107(10): 3182-3210

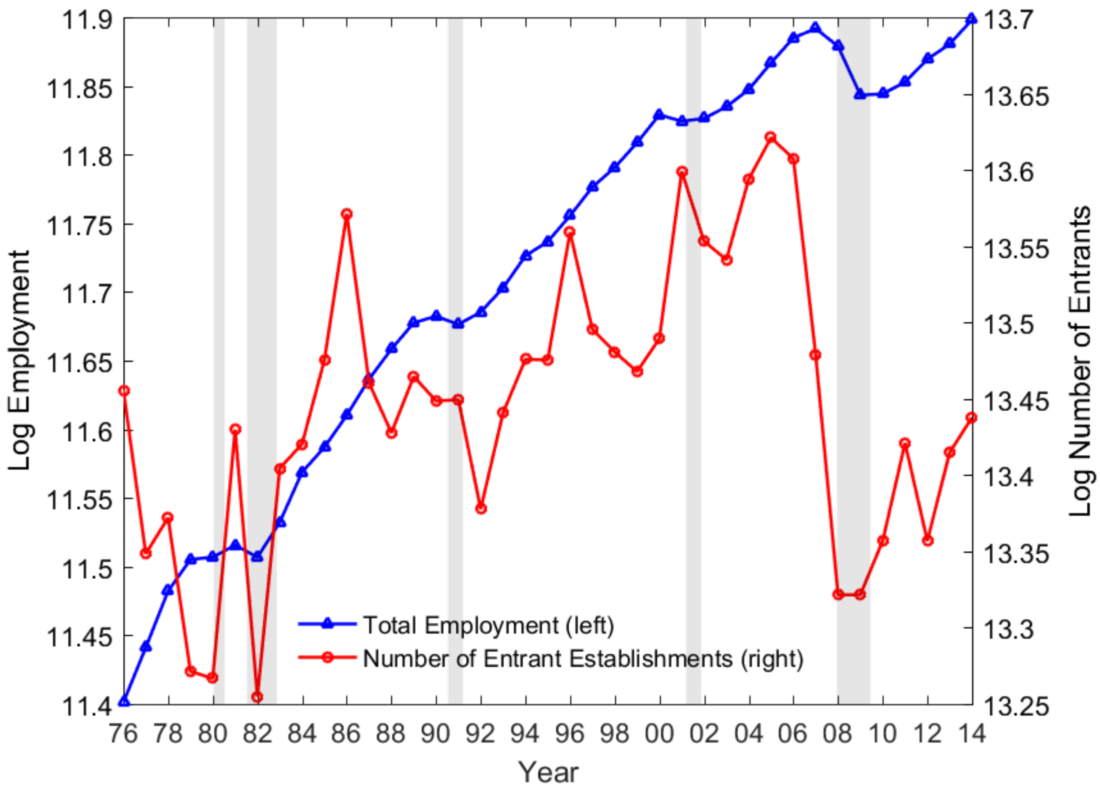
[39] Siemer M., 2016. Firm entry and employment dynamics during the Great Recession.

[40] Veracierto, M., L., 2002. Plant-level irreversible investment and equilibrium business cycles. American Economic Review, 92, 181–197

[41] Young., E., 2010. Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. Journal of Economic Dynamics and Control, 34(1)

10 Tables and Figures

Figure 1: Dynamics of the Number of Entrant Establishments and Aggregate Employment



Note: Figure 1 uses data from the Longitudinal Business Database provided in Business Dynamic Survey dataset. Establishment entry is defined by the existence of March 12 employment. Monthly aggregate employment data comes from the Federal Reserve Economic Dataset. The construction of annual aggregate employment is described in Appendix C.2. Period: 1977-2014.

Figure 2: **Average Survival Rate over the Business Cycles**. [1] Indicator: A period is recessionary if the cycle component of the log real GDP is negative (log real GDP is below trend)

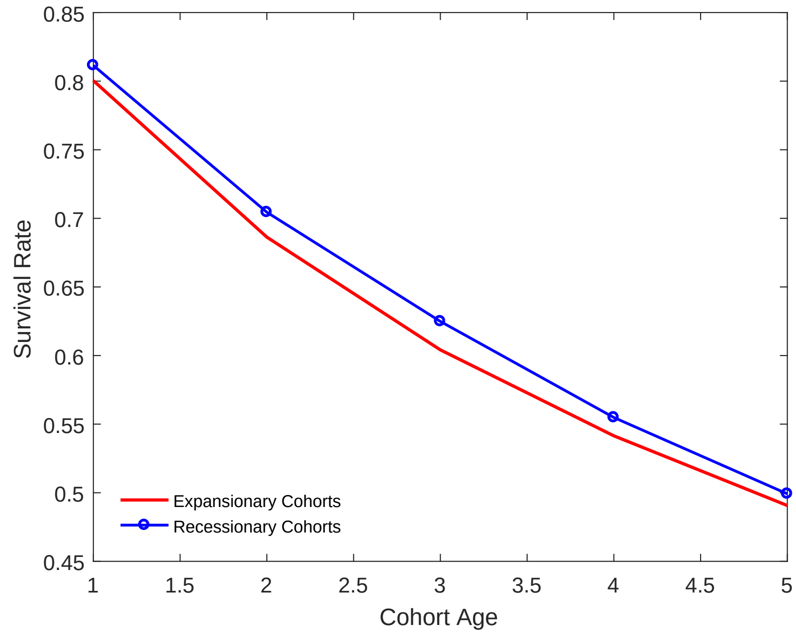
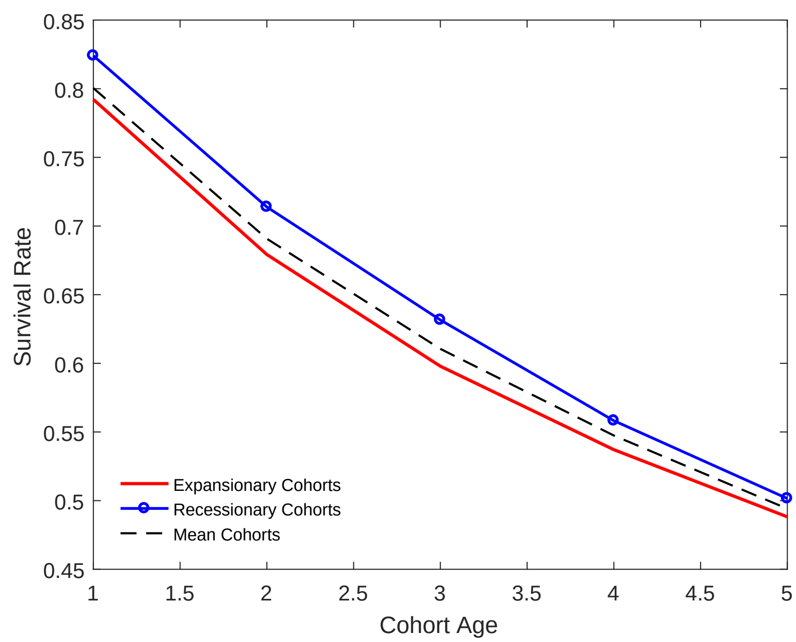


Figure 3: **Average Survival Rate over the Business Cycles**. [2] Indicator: A period is referred as recession if the log real GDP is below trend by 1 percent and expansion if the log real GDP is above trend by 1 percent. Cohorts born in years when the absolute deviation of the log real GDP from the trend is less than 1 percent are referred as mean cohorts.



Source: BDS, Establishment by Age. 1977-2015.

Figure 4: **Average Survival Rate over the Business Cycles**. [3] Indicator: NBER based Recession indicator for the United States from the period Following the Peak through the Trough.

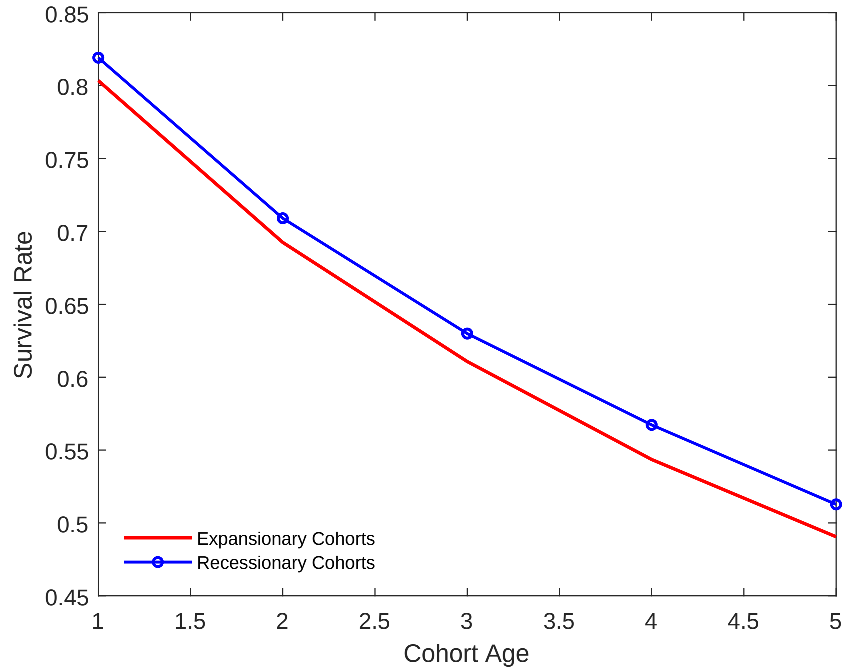


Table 1: **Correlations between the average survival rate and the recession indicators.**

	[1] Indicator	[2] Indicator	Cycle of log real GDP	[3] Indicator
Survival up to age 1	-0.16	-0.37**	-0.24	-0.16
Survival up to age 2	-0.29*	-0.45***	-0.34**	-0.19
Survival up to age 3	-0.33**	-0.44***	-0.32**	-0.23
Survival up to age 4	-0.25	-0.33**	-0.23	-0.34**
Survival up to age 5	-0.19	-0.24	-0.12	-0.37**

Note. Cycle of log real GDP refers to a cyclical component of the HP filtered log real GDP. Smoothing parameter equals to 100. [1] Indicator: A period is referred as recession if the cycle component of the log real GDP is negative (log real GDP is below trend). The indicator takes value 1 if a year is defined as recession and 0 if a year is defined as expansion. [2] Indicator: A period is referred as recession (expansion) if the log real GDP is below (above) trend by 1 percent. Cohorts born in years when the absolute deviation of the log real GDP from the trend is less than 1 percent are referred as mean cohorts. The indicator takes value 1 if a year is defined as recession, 0 if year is defined as mean year and -1 if a year is defined as expansion. The choice of the magnitude of the deviation equally divides 39 observation into three groups. [3] Indicator: NBER based Recession Indicators for the United States from the Period following the Peak through the Trough.

Figure 5: **Incumbent Firm's Timing**

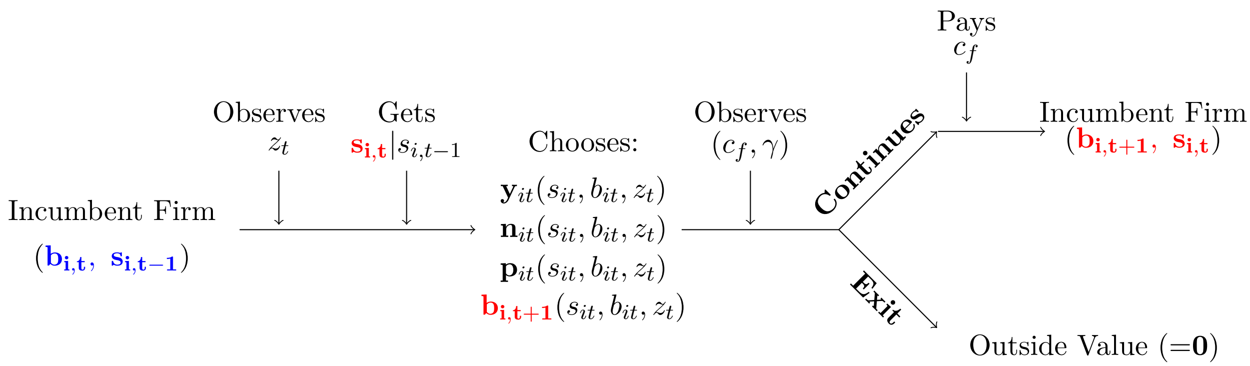


Figure 6: **Potential Entrant Firm's Timing**

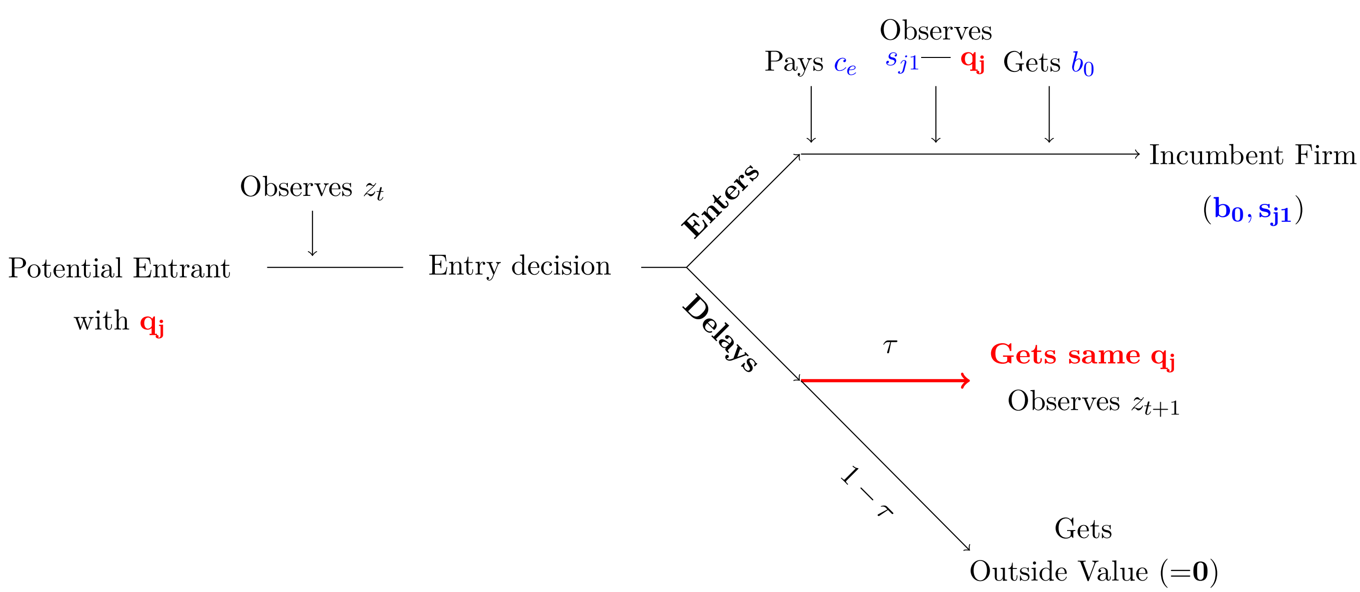


Figure 7: *Option Value of Delay* (q, z)

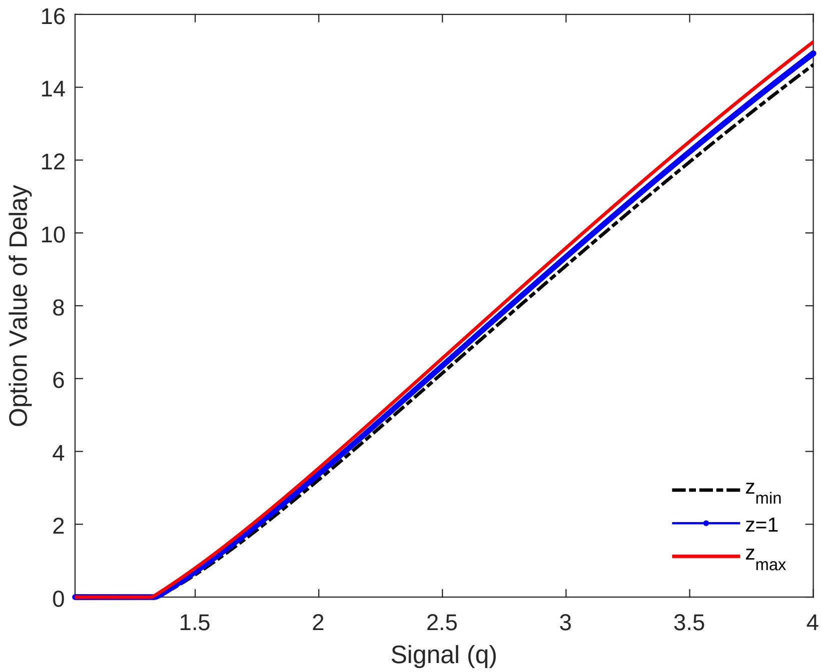
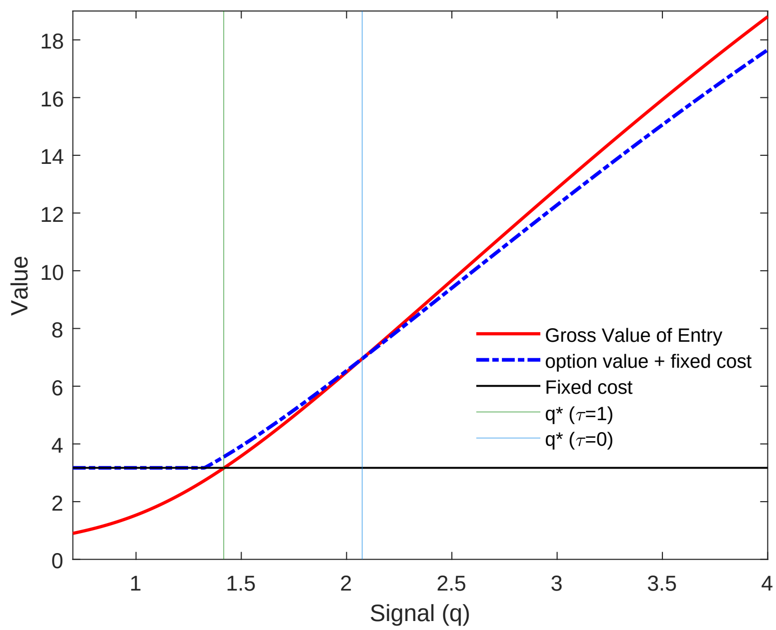
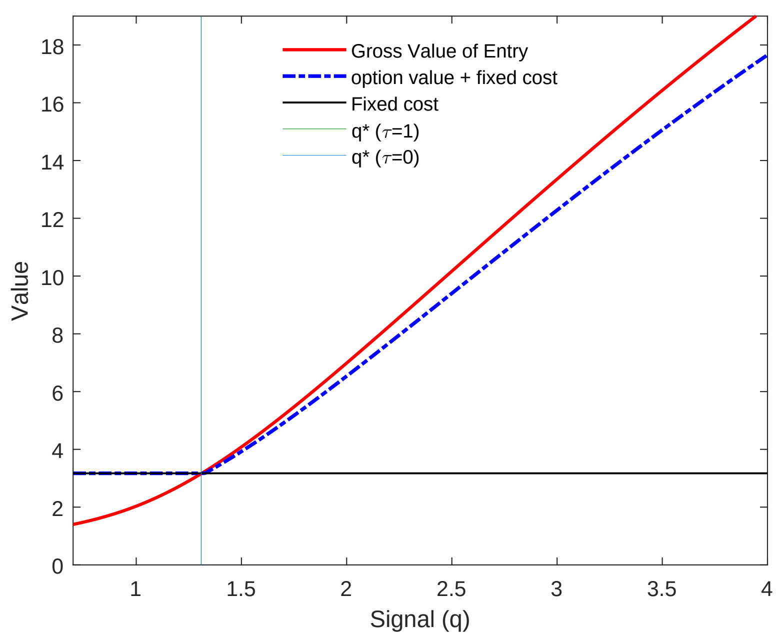


Figure 8: Entry decision for each signal in the lowest ($z = z_{min}$) and in the highest ($z = z_{max}$) aggregate demand levels (Option value of delay is constructed for illustrational purposes. The figure captures key features of the selection. Actual option value of delay is given in previous figure).



(a) Selection when $z = z_{min}$



(b) Selection when $z = z_{max}$

Figure 9: Equilibrium threshold signal

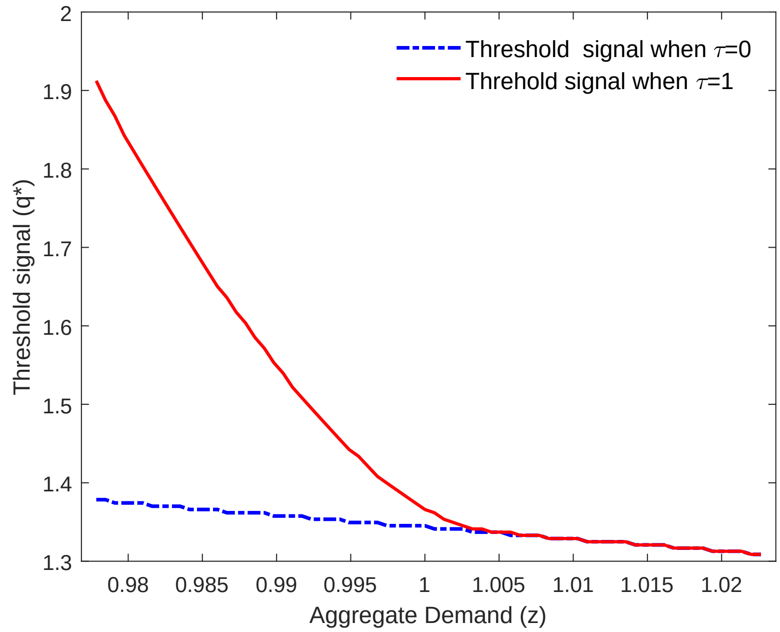


Figure 10: Equilibrium opportunity cost of entry

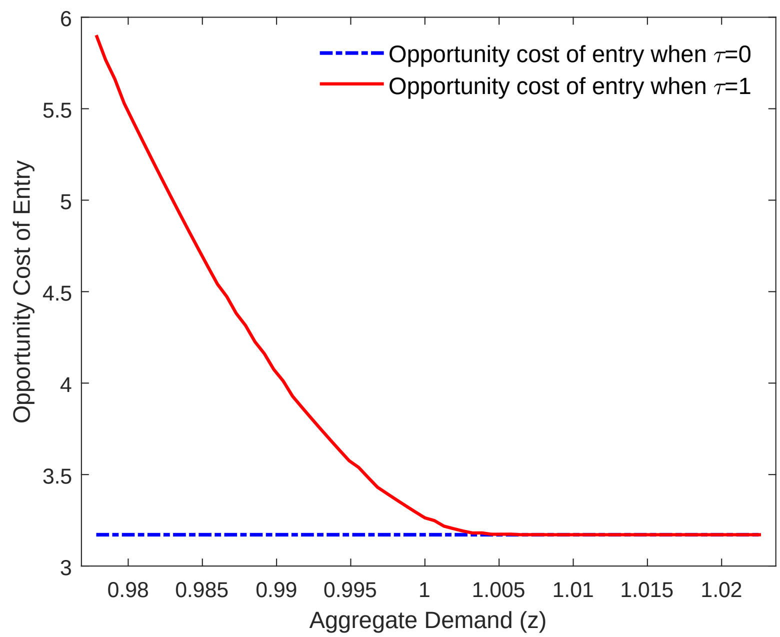


Figure 11: Value of waiting

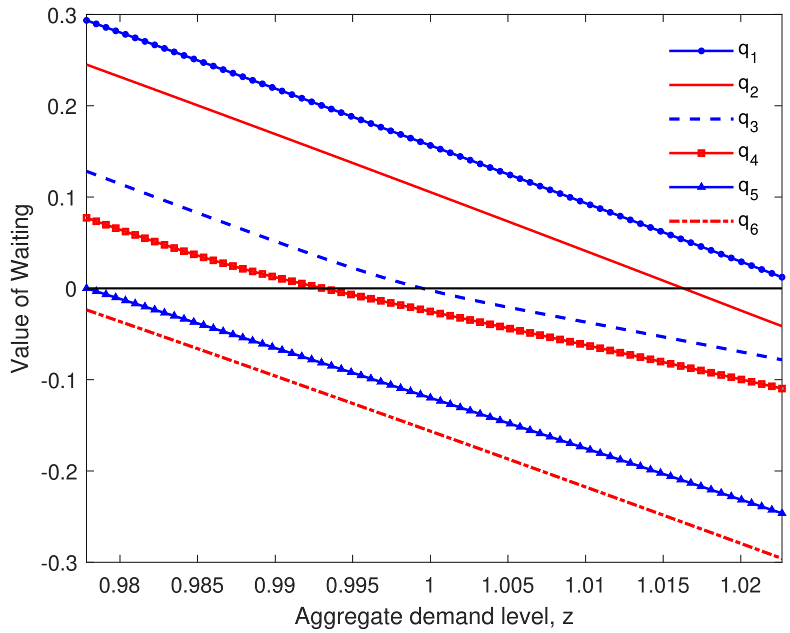


Figure 12: Total cost of entry

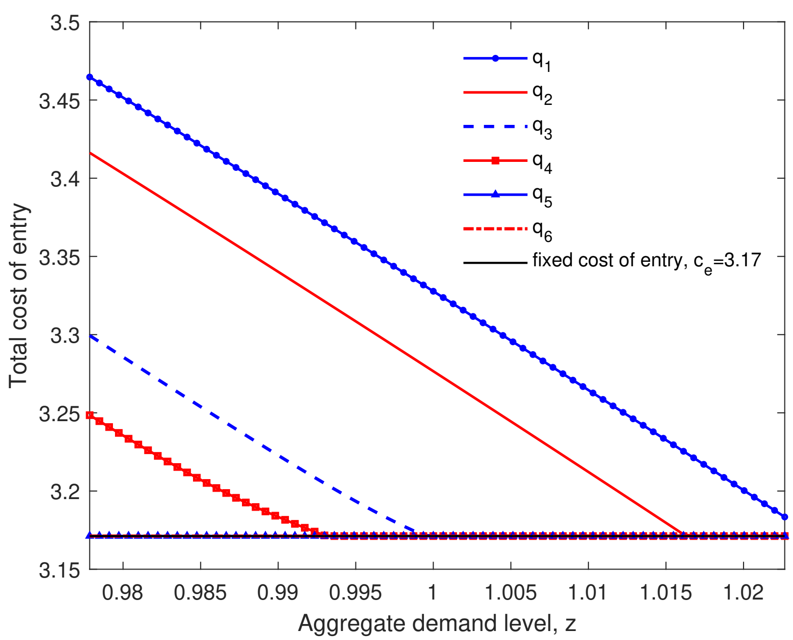
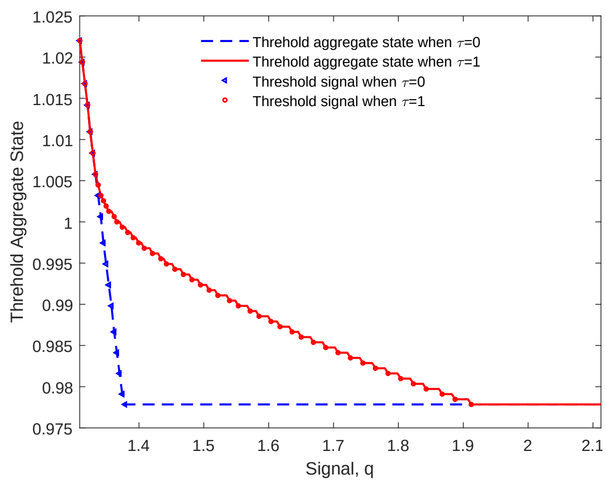
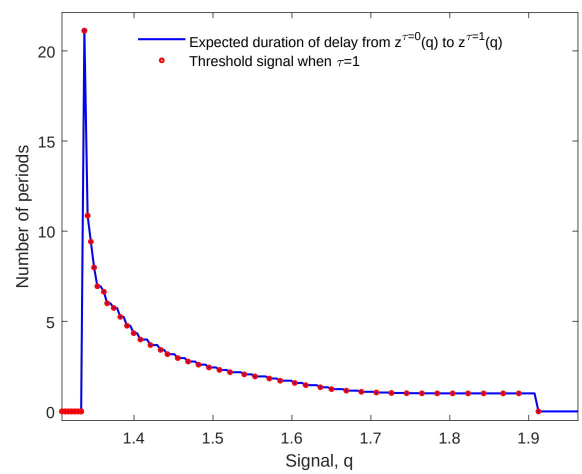


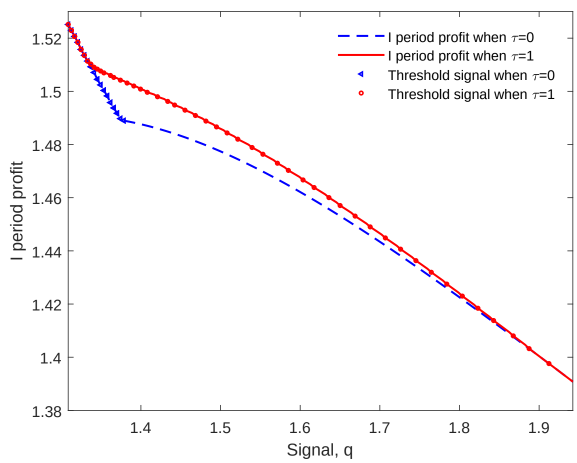
Figure 13: Decomposition of the potential entrants' expected gross value of entry at the threshold aggregate state (minimum aggregate state that the entrant with signal q is ready to enter) between expected first period profit and expected long run profit. The figures include values for the case when the signal is perfectly persistent ($\tau = 1$) and when the signal has zero persistence ($\tau = 0$).



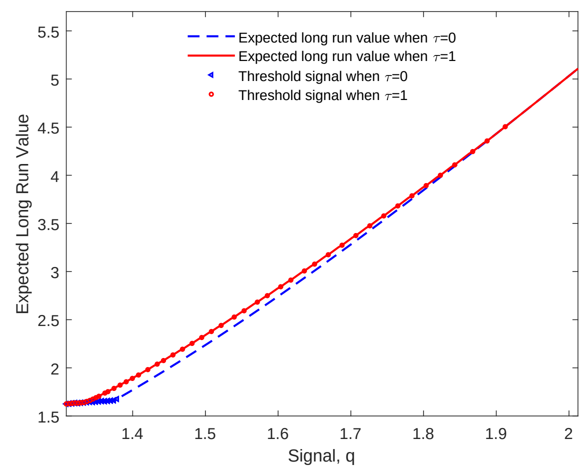
(a) Threhold aggregate state $\underline{z}^{\tau}(q)$



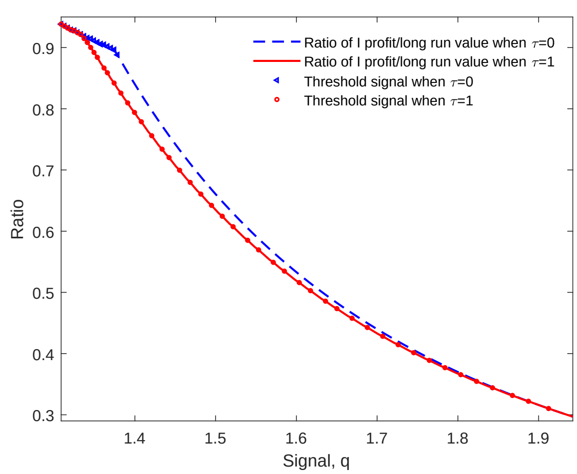
(b) Expected duration



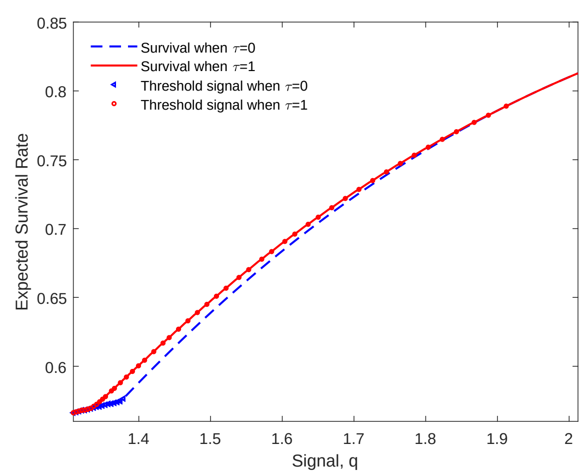
(c) Expected first period profit



(d) Expected long-run profit

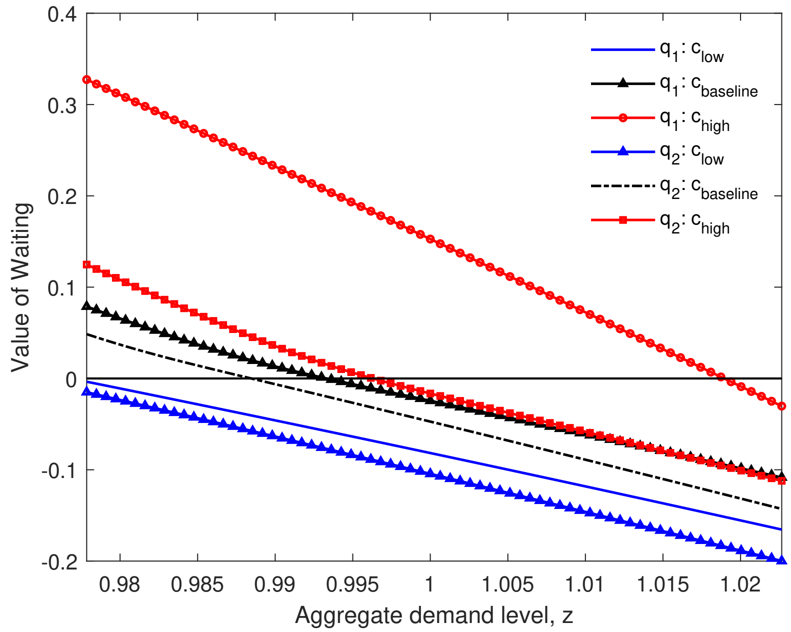


(e) Ratio of first period profit to long-run value

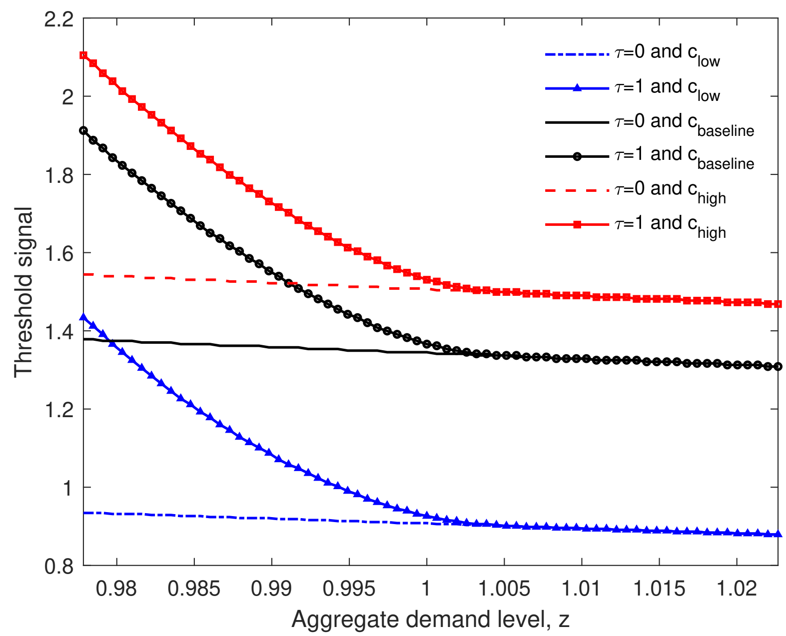


(f) Expected survival rate

Figure 14: Comparative statics: different fixed cost of entry

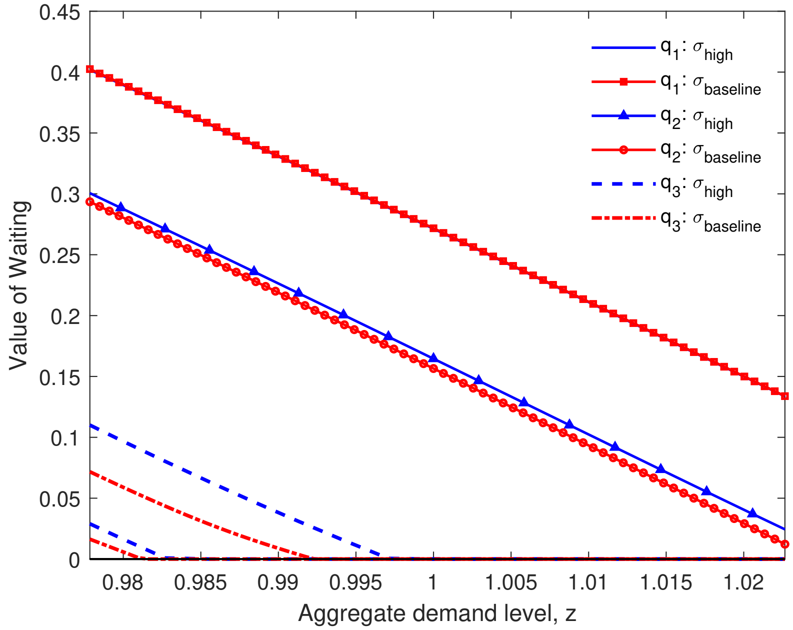


(a) Value of waiting

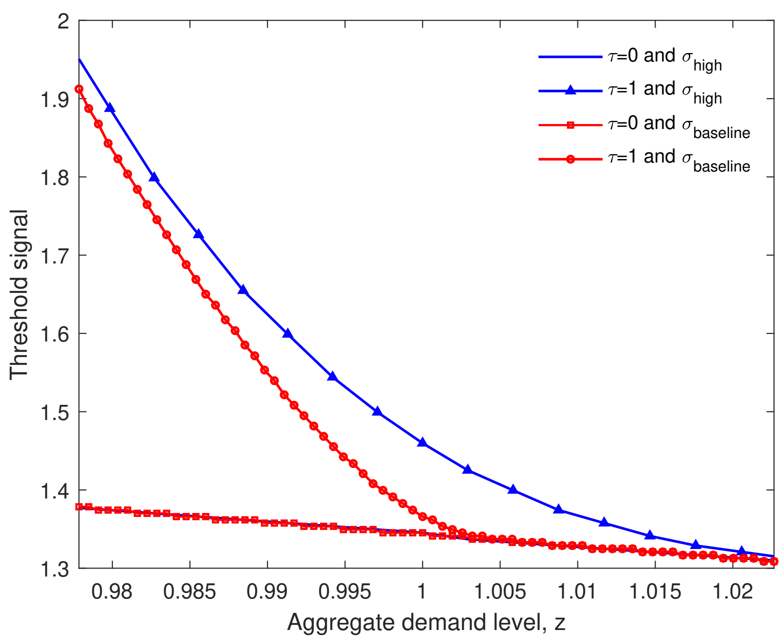


(b) Threshold signal

Figure 15: Comparative statics: different σ

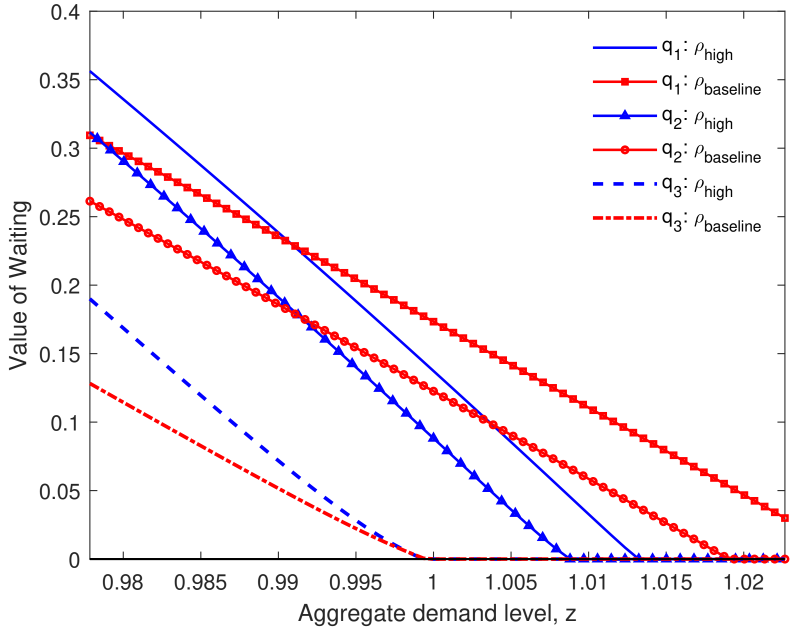


(a) Additional cost due to value of waiting

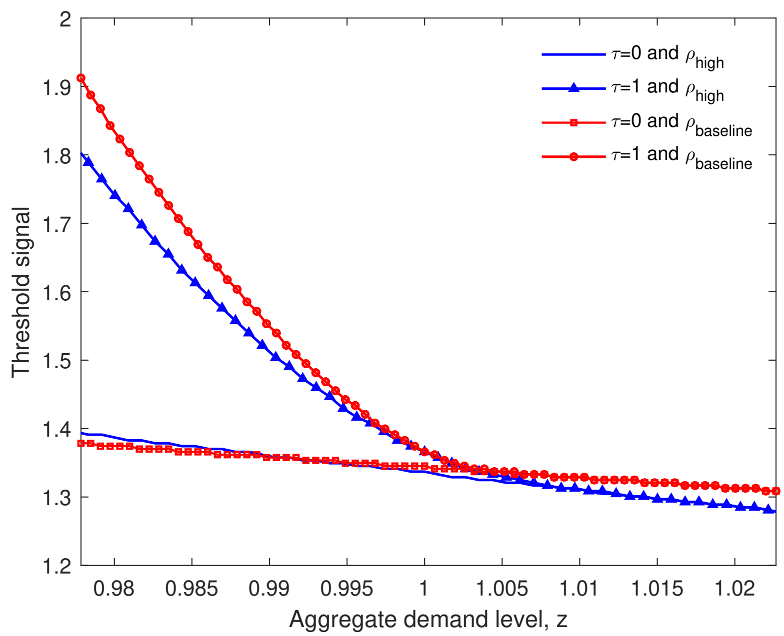


(b) Threshold signal

Figure 16: Comparative statics: different ρ

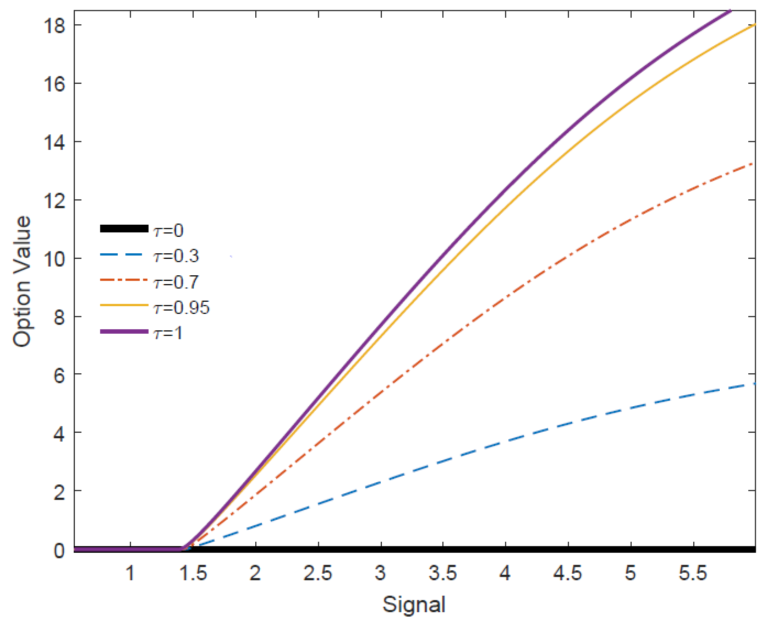


(a) Additional cost due to value of waiting

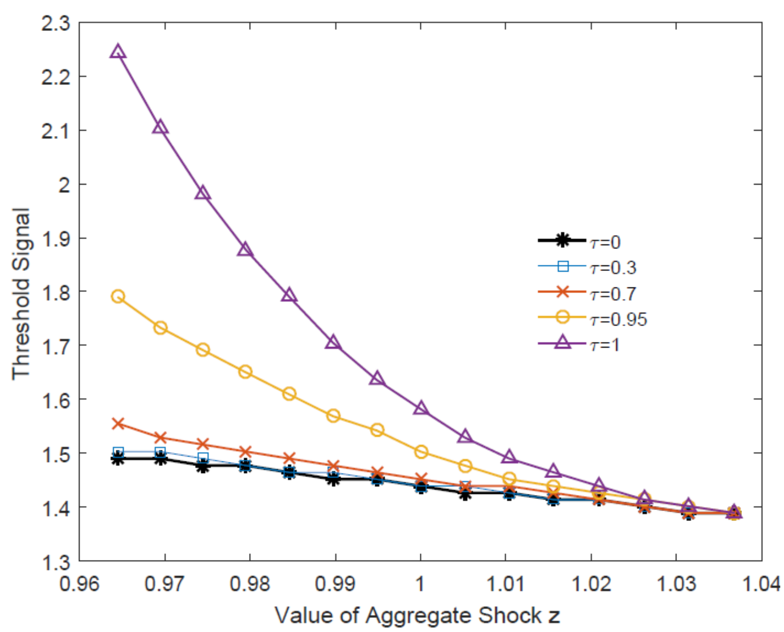


(b) Threshold signal

Figure 17: Comparative statics: different τ.



(a) The option value of delay



(b) Threshold signal

Figure 18: Share of the high propensity business applications that becomes employer business within 4 and 8 quarters. *Source: The Business Formation Statistics (BFS), the U.S. Census Bureau.*

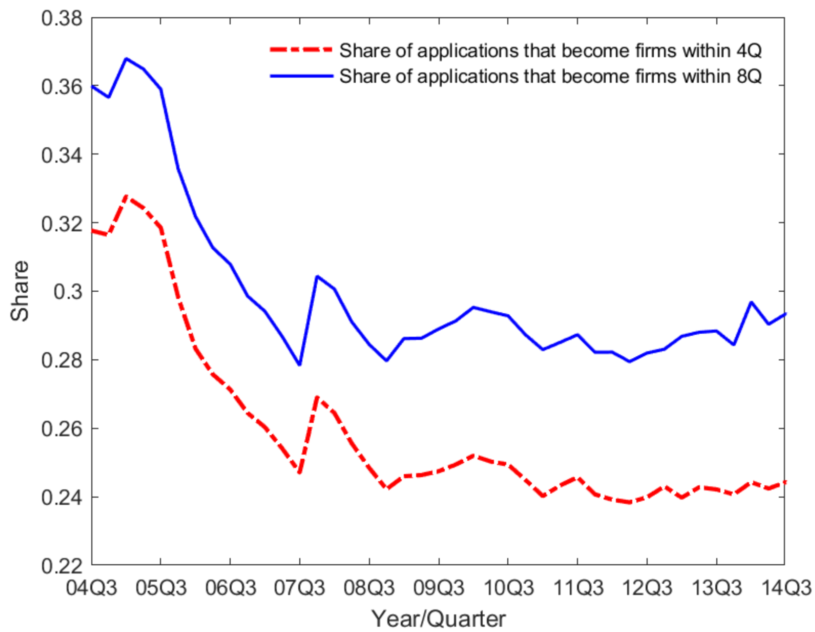


Figure 19: Average duration (in quarters) from business application to formation within 4 and 8 quarters. *Source: The Business Formation Statistics (BFS), the U.S. Census Bureau.*

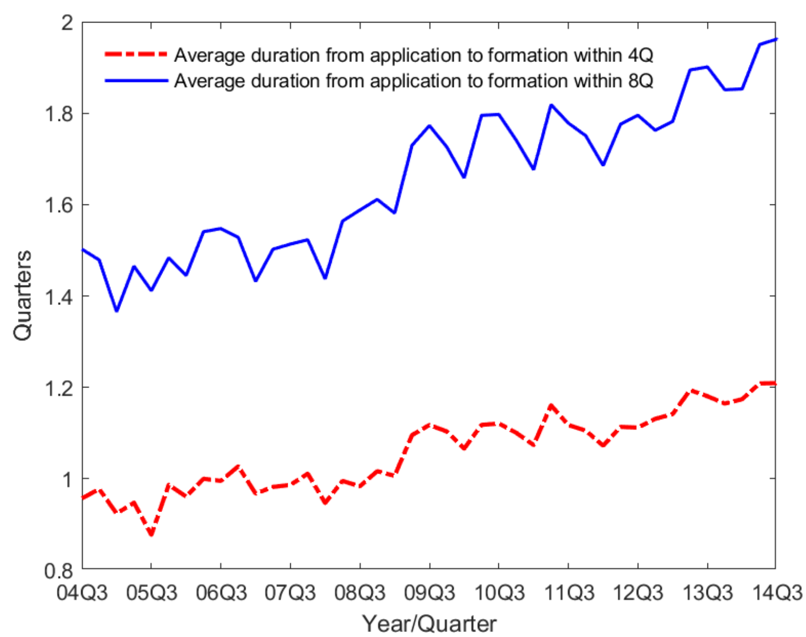


Table 2: Calibration

Symbol	Description	Value	Calibration Targets/Source
τ	Signal persistence	1.0	
β	Discount rate	0.960	Annual riskless interest rate. $R = 1.04$
ρ_z	Persistence of aggregate shock	0.570	Persistence of entry rate
σ_z	Std. Dev. aggregate shock	0.002	Standard deviation of entry rate
ρ_s	Persistence of idiosync. prod. shock	0.814	Foster et al. (2008)
ρ	Price elasticity of demand	1.622	Foster et al. (2016)
η	Elasticity of demand to capital	0.919	Foster et al. (2016)
δ	Depreciation rate of reputation	0.188	Foster et al. (2016)
b_o	Initial customer capital	12.0	Establishment level moments
σ_s	Std. Dev. idiosyncratic shock	0.16	Establishment level moments
σ_s^e	Std. Dev. Initial productivity	0.26	Establishment level moments
α	Demand shifter	0.26	Establishment level moments
$\underline{q}$	Pareto location	0.70	Establishment level moments
ξ	Pareto exponent	3.98	Establishment level moments
μ_f	Mean log operational cost	0.62	Establishment level moments
σ_f	Std. Dev. log operational cost	0.41	Establishment level moments
γ	Exit shock	0.07	Establishment level moments
c_e	Entry cost	3.98	Establishment level moments

Table 3: Calibration targets for establishment level charachteristics

Statistics	**Data**	**Model**
Average entry rate (1991-2006) (%)	12.1	12.1
Average size of all establishments	17.0	16.3
Entrant employment share in total employment (%)	5.9	6.4
Cohort employment share in total employment at age 5 (%)	4.2	4.2
Average size of entrants (age 0)	8.7	8.5
Average size of cohort at age 5	13.9	14.1
Average size of cohort between 21 and 25 years	21.4	22.4
Survival until 5 years old	0.48	0.41
Survival between 21 and 25 years	0.15	0.10
Establishments exit rate at age 5	0.12	0.09

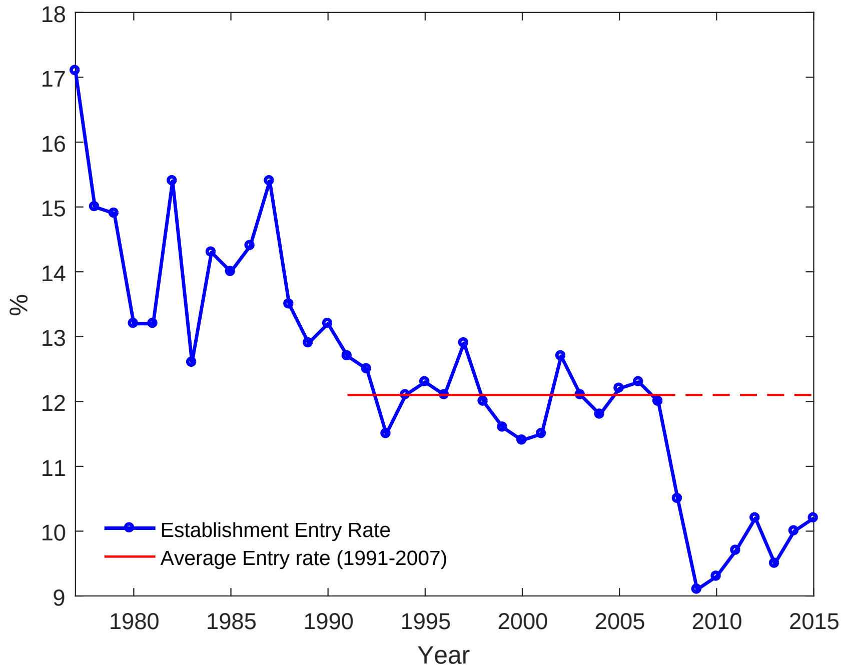
Notes. The moments are calculated from the BDS dataset covering the economy-wide establishment level data over the period 1977-2015. Construction of the empirical moments is described in details in Appendix D.1.1

Table 4: Calibration targets for aggregate demand shock process

Statistics	**Data**	**Model**
Autocorrelation of the cycle component of entry rate	0.25	0.25
Standard Deviation of the cycle component of entry rate	0.06	0.06

Notes. Entry Rate comes from the BDS and covers period from 1977 to 2015. The cyclical component of the log entry rate is calculated using the HP filter with smoothing parameter 100. I simulated the process of the entry rate over 10000 periods. To be consistent with the data moments I apply the same detreding method to the simulated entry rate series in the model.

Figure 20: Entry Rate



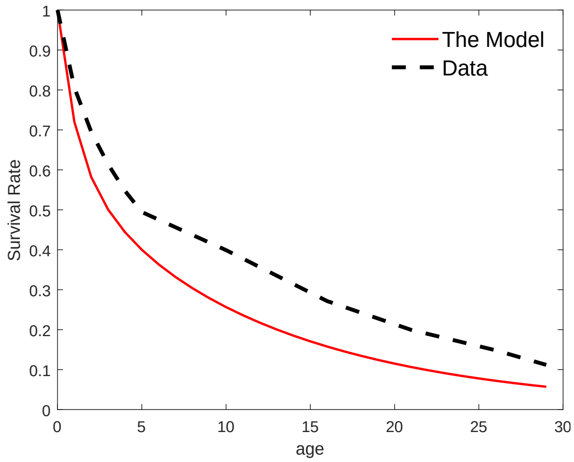
Source: BDS, Economy wide establishment level data. 1977-2015.

Table 5: Calibration of alternative scenarios

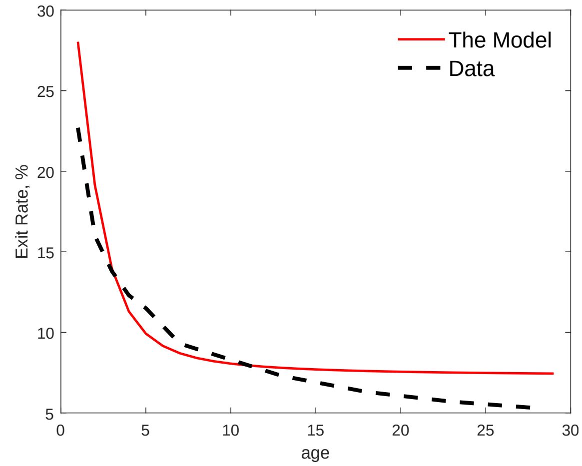
Description	**Symbol**	**Baseline** (a)	$\boldsymbol{\tau = 0}$ (b)	**Without** (c)	**Calib** τ (d)
Persistence of signal	τ	1.0	0.0*	0.0*	0.965*
Discount factor	β	0.96	0.96	0.96	0.96
Persistence of agg. shock	ρ_z	0.57	0.57	0.57	0.57
St.Dev. agg. shock	σ_z	0.0022	0.0022	0.016*	0.0038*
Persistence of product.	ρ_s	0.814	0.814	0.814	0.814
Price elasticity	ρ	1.622	1.622	1.622	1.622
Capital elasticity	η	0.919	0.919	0.919	0.919
Capital Depreciation	δ	0.188	0.188	0.188	0.188
Initial customer capital	b_o	12.0	12.0	12.0	12.0
Std. dev. prod.	σ_s	0.16	0.16	0.16	0.16
Std. dev. initial prod.	σ_s^e	0.26	0.26	0.26	0.26
Demand shift.	α	0.26	0.26	0.26	0.26
Pareto Location	$\underline{q}$	0.70	0.70	0.70	0.70
Pareto Exponent	ξ	3.98	4.41	4.41	4.41
Mean c_f	μ_f	0.62	0.62	0.62	0.62
Std. Dev. c_f	σ_f	0.41	0.41	0.41	0.41
Exit shock	γ	0.07	0.07	0.07	0.07
Entry cost	c_e	3.17	3.26*	3.26*	3.17*

Note. 'Baseline' refers to the baseline scenario. '$\tau = 0$' refers to a case which sets $\tau = 0$ in the baseline scenario. 'Without' refers to a model with $\tau = 0$ and everything is re-calibrated to match the same set of moments as in the baseline scenario. 'Calib τ' refers to a scenario where the baseline model is re-calibrated to additionally match the variation in cyclical component of employment in the model to the data counterpart, using τ as free parameter. The value indicated with (*) highlights the parameters that is used as free parameters in contrast to the baseline scenario.
Fixed entry cost in the case with $\tau = 0$: $3.26 = 3.17(c_e) + 0.9$ (Option Value of delay in the stochastic steady state)

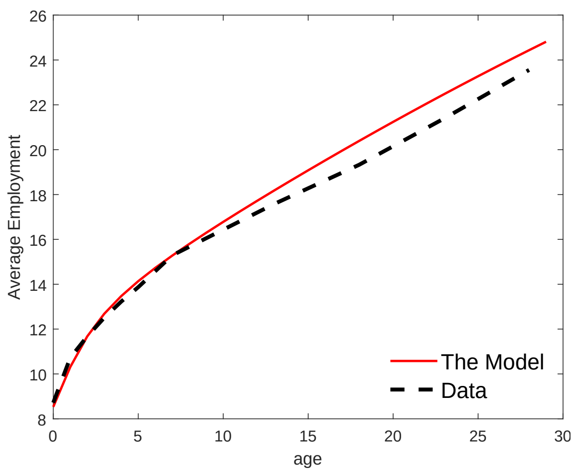
Figure 21: Average Cohorts Characteristics: Data, Model



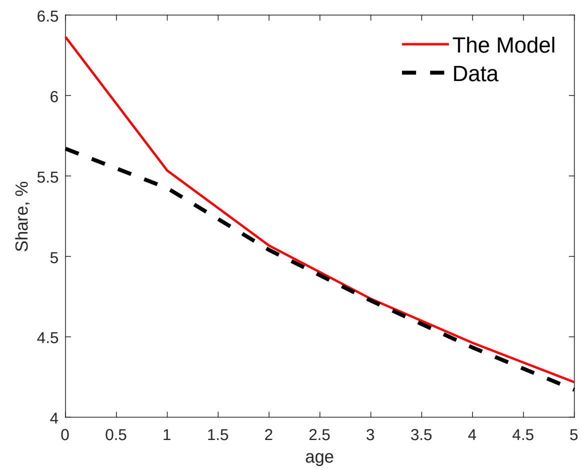
(a) Average Survival Rate



(b) Exit rate by age

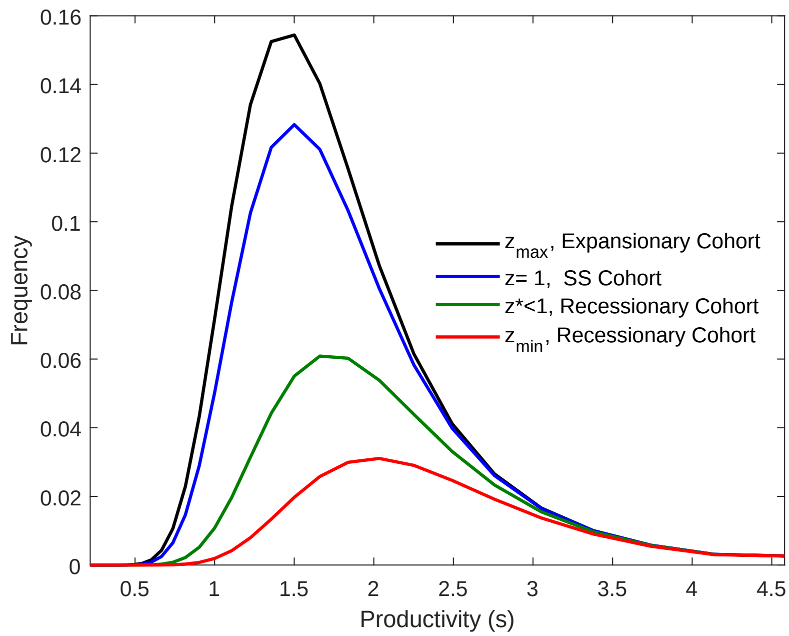


(c) Average Size

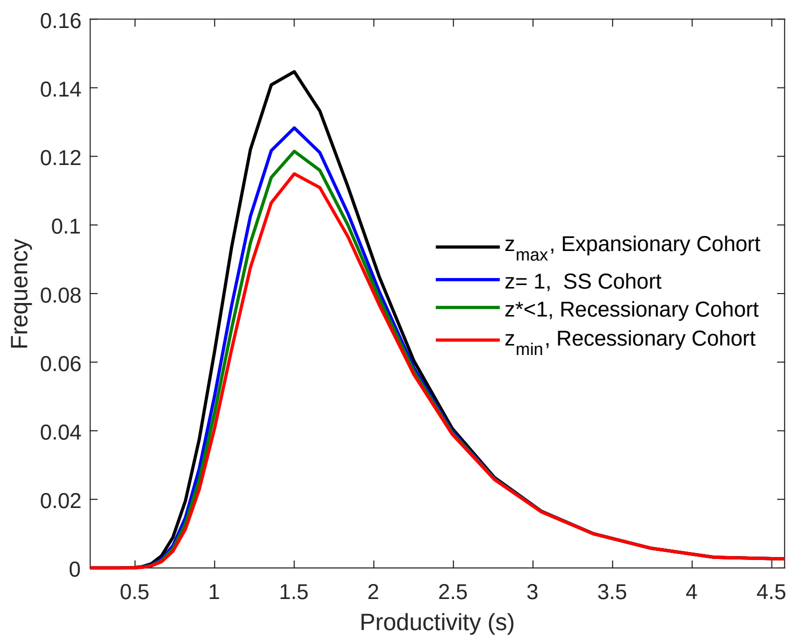


(d) Share of cohort employment in total employment

Figure 22: Cohorts over BC: Productivity



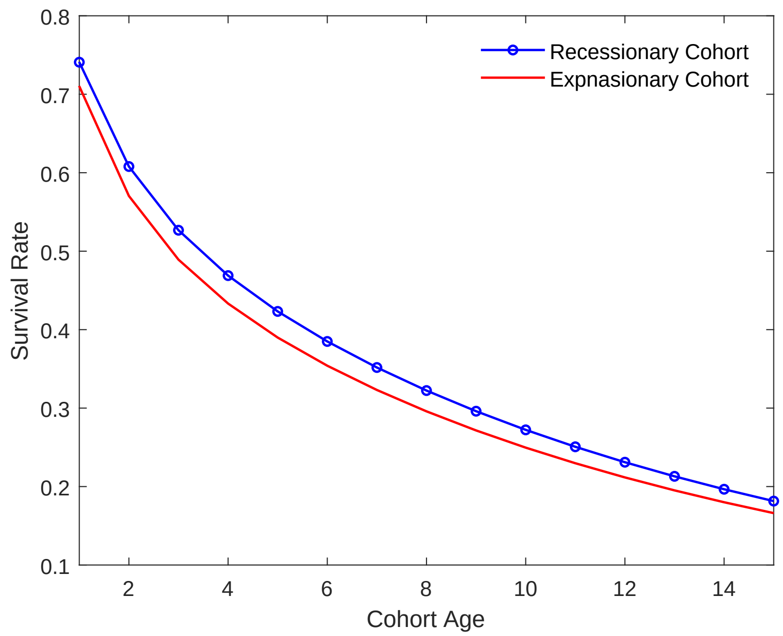
(a) Baseline model



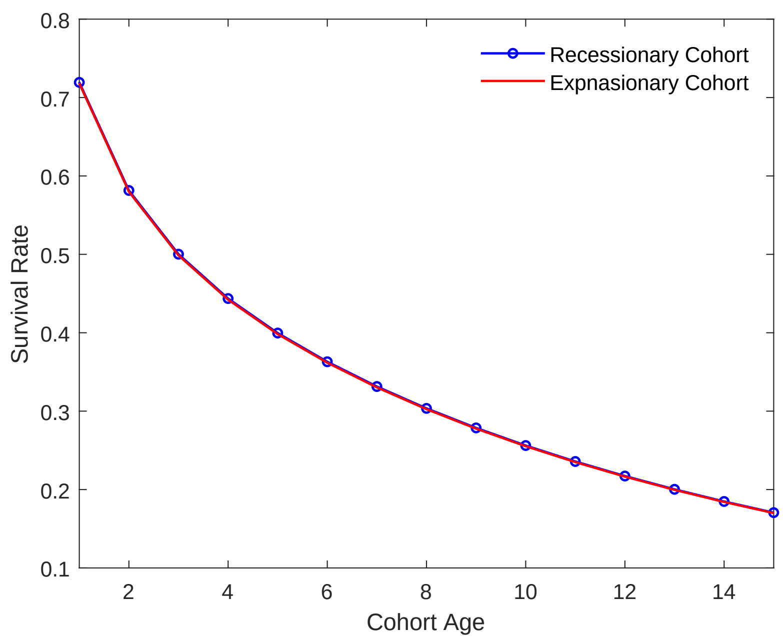
(b) The case with $\tau = 0$

Note. Simulation over 1200 periods. Recessionary cohorts are defined as cohorts born during negative aggregate demand shock.

Figure 23: Cohorts over BC: Survival Rate



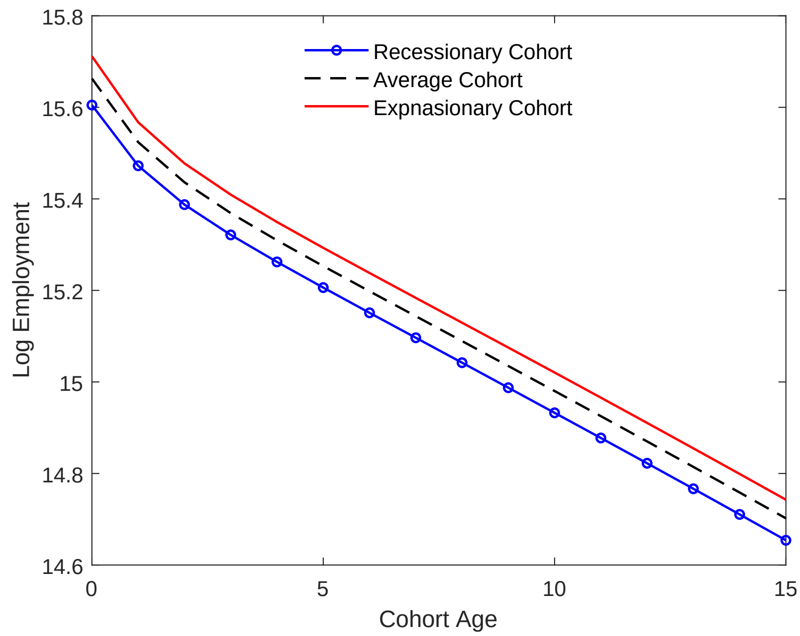
(a) Baseline model



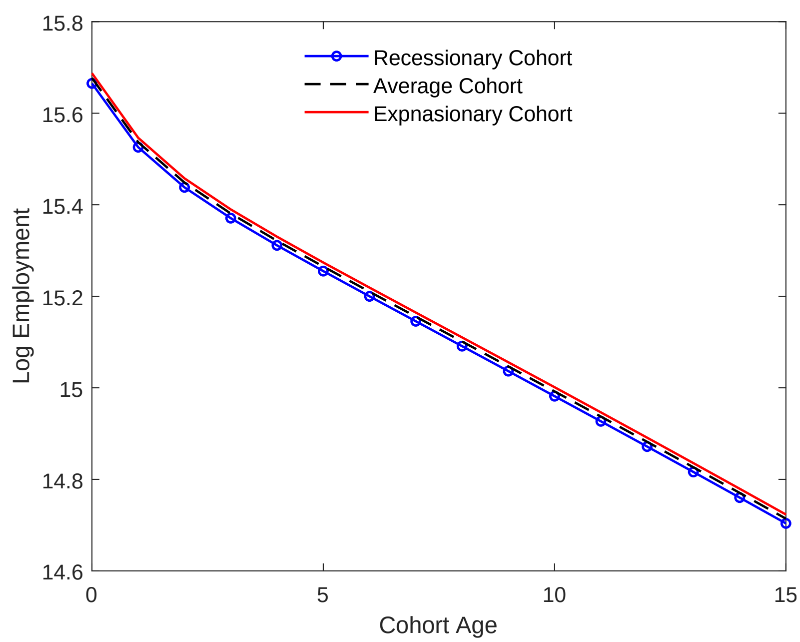
(b) The case with $\tau = 0$

Note. Recessionary cohorts are defined as cohorts born during the states when the aggregate demand level is 1% below the steady state level. Expansionary cohorts are defined as cohorts that enter the market when the aggregate demand level is 1% above the steady state level. I simulate the economy over 1200 periods.

Figure 24: Cohorts over BC: Cohort employment



(a) Benchmark Model



(b) Baseline model

Note. Recessionary (expansionary) cohorts are defined as cohorts that enter the market when the aggregate demand level is below (above) zero. I simulate the economy over 1200 periods.

Table 6: Cohort employment over cycles: The basleline model and the counterfactual scenarios.

		Recessionary Cohorts			Expansionary Cohorts		
		At entry % dev.	Age 5 % dev.	Age 15 % dev.	At entry % dev.	Age 5 % dev.	Age 15 % dev.
(a)	Baseline	-5.7	-4.7	-4.8	5.0	4.0	4.1
(b)	Baseline, no selection	-0.4	-0.3	-0.3	0.4	0.3	0.3
(c)	$\tau = 0$	-1.2	-1.0	-1.0	1.0	0.9	1.0
(d)	$\tau = 0$, adjust lowest s	-3.4	-1.4	-1.5	2.6	1.2	1.3
(e)	$\tau = 0$, adjust highest s	-12.5	-14.1	-13.3	10.0	11.2	10.6

Note: The numbers describe percentage deviations (% dev.) of the recessionary (expansionary) cohorts employment from the average cohort employment. 'Baseline' refers to the economy with the baseline specification. 'Baseline, no selection' refers to the baseline scenario while the number and the composition of entrants are fixed at the stochastic steady state level. '$\tau = 0$' refers the case with $\tau = 0$. '$\tau = 0$, adjust lowest s ' ('$\tau = 0$, adjust highest s') refers to the scenario when the distribution of entrants in the case $\tau = 0$ is adjusted using lowest (highest) productive entrants to generate data-conforming variation in the number of entrants. 'Recessionary' ('Expansionary') cohorts refer to the group of firms that started operation when $z < 1$ ($z > 1$).

Table 7: Business Cycle Moments: Data, the baseline model, and the counterfactual scenarios.

		Data (a)	Baseline (b)	Baseline, only selection (c)	Case $\tau = 0$ (d)	$\tau = 0$, adjust lowest s (e)	$\tau = 0$, adjust highest s (f)
No. of firms	ρ	0.640	0.619	0.607	0.661	0.481	0.680
	σ	0.012	0.010	0.010	0.002	0.009	0.076
Employment	ρ	0.610	0.574	0.622	0.432	0.457	0.657
	σ	0.015	0.012	0.010	0.004	0.009	0.126
Output	ρ	0.630	0.576	0.620	0.440	0.459	0.677
	σ	0.020	0.012	0.010	0.004	0.009	0.115
Entry Rate	ρ	0.250	**0.253**	0.252	0.222	0.254	0.404
	σ	0.062	**0.065**	0.065	0.010	0.065	0.103
No. of Entrants	ρ	0.311	0.278	0.278	0.245	**0.278**	**0.278**
	σ	0.066	0.073	0.073	0.011	**0.073**	**0.073**

Notes. The numbers that are in bold refer to the targeted model statistics. All other values indicate untargeted model statistics and their empirical counterparts. 'Baseline' refers to the economy with the baseline specification. 'Baseline, only selection' refers to a counterfactual scenario where the aggregate demand shocks affect only selection of firms at entry (same way as in the baseline model) and has no affect on the firms post-entry decisions. '$\tau = 0$, adjust lowest s ' ('$\tau = 0$, adjust highest s') refers to the scenario where the distribution of entrants in the case with $\tau = 0$ is adjusted using lowest (highest) productive entrants to generate data-conforming variation in the entry rate. The time-series about the entry rate and the number of entrants come from the BDS and covers 1977-2015 period. I take time-series about real GDP and total employment from the FRED. The cyclical component of each of the data series are calculated using the HP filter with smoothing parameter 100.

Table 8: Business Cycle Moments: Data, Model

		Data	Baseline	The model **without** persistent signal	
				selection and post-entry shocks	selection but no post-entry shocks
		(a)	(b)	(c)	(d)
No. of firms	ρ	0.640	0.619	0.684	0.605
	σ	0.012	0.010	0.011	0.010
Employment	ρ	0.610	0.574	0.439	0.620
	σ	0.015	0.012	0.025	0.011
Output	ρ	0.630	0.576	0.448	0.623
	σ	0.020	0.012	0.024	0.011
Entry Rate	ρ	0.250	**0.253**	**0.272**	0.265
	σ	0.062	**0.065**	**0.064**	0.063
No. of Entrants	ρ	0.311	0.278	0.294	**0.294**
	σ	0.066	0.073	0.071	**0.071**

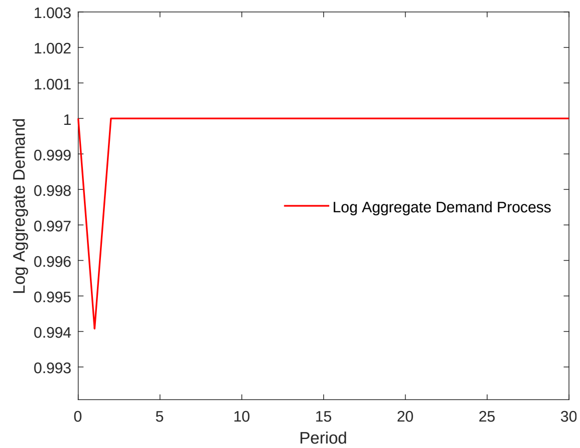
Notes. The numbers that are in bold refer to the targeted model statistics. All other values indicate untargeted model statistics and their empirical counterparts. The time-series about the entry rate and the number of entrants come from the BDS and covers 1977-2015 period. I take time-series about real GDP and total employment from the FRED. The cyclical component of each of the data series are calculated using the HP filter with smoothing parameter 100.

Table 9: Propagation of the shocks

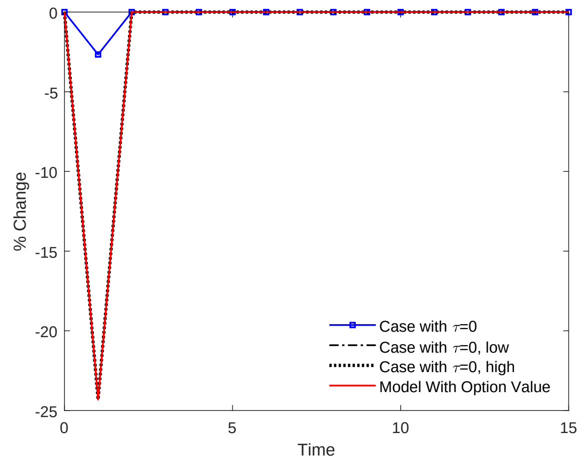
		1-time shock			Persistent Shock		
		Full Model	Fixed entry		Full Model	Fixed entry	
				z_{high}			z_{high}
		(a)	(b)	(c)	(d)	(e)	(f)
Depth (%)	Employment	-1.83	-0.72	-1.90	-2.0	-0.72	-2.1
	No. of Firms	-2.93	-0.07	-0.18	-3.04	-0.14	-0.43
50% Recovery	Employment	3	2	2	16	7	6
	No. of Firms	3	15	14	9	17	18
75% Recovery	Employment	15	2	2	28	14	15
	No. of Firms	8	23	23	17	26	27

Note: 'The model' refers to the baseline model. 'CF: Fix entry' refers to a case where the entry rate is fixed at the stochastic steady state and adjustment happens at the intensive margin. z_{high} refers to a case where the magnitude of the shock is chosen such that to produce a drop in employment as in 'The model'. 'Depth' refers to the highest deviation of the time series from trend. 50% Recovery (75% Recovery) describes number of periods (years) starting from period 1 after which economy recovers 50% (75%) from the 'depth'. The magnitude of the generates 25% of the drop in the number of entrant firms similar that was observed during the Great Recession.

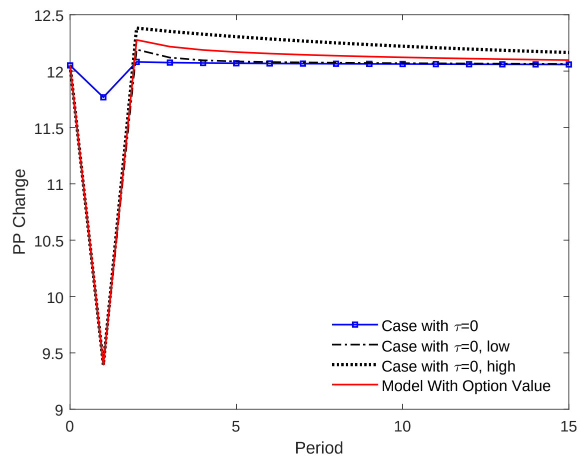
Figure 25: Impulse response to 1-time aggregate demand shock



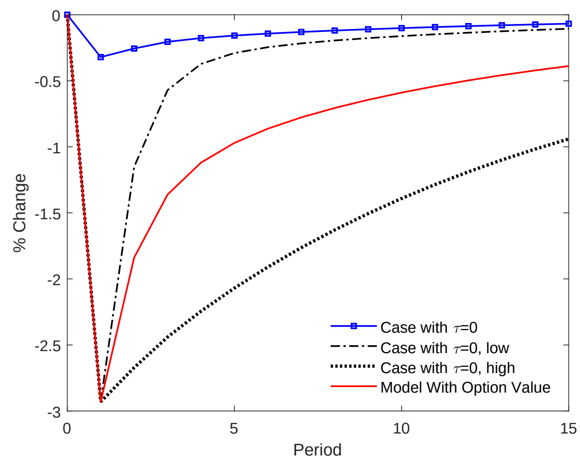
(a) Shock Process



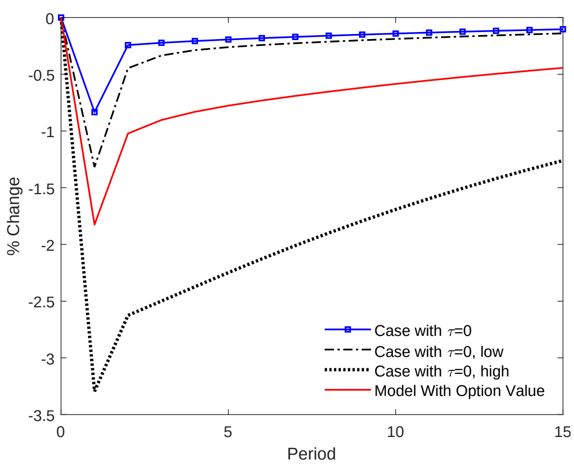
(b) Number of Entrants



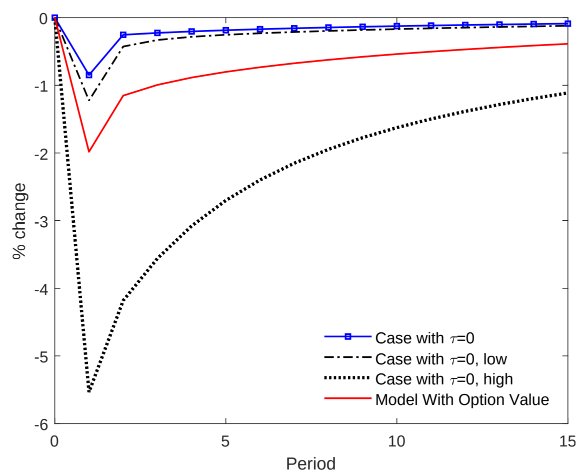
(c) Entry Rate



(d) Number of Firms



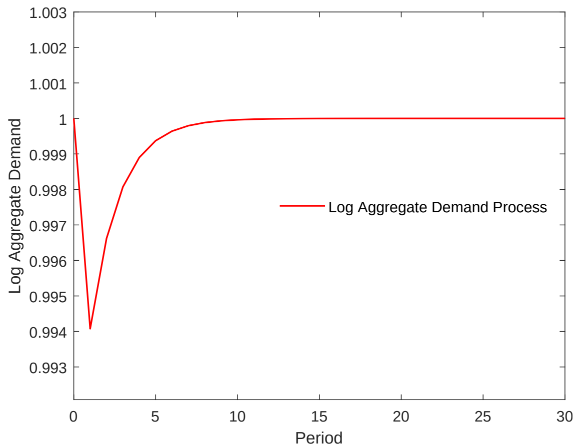
(e) Employment



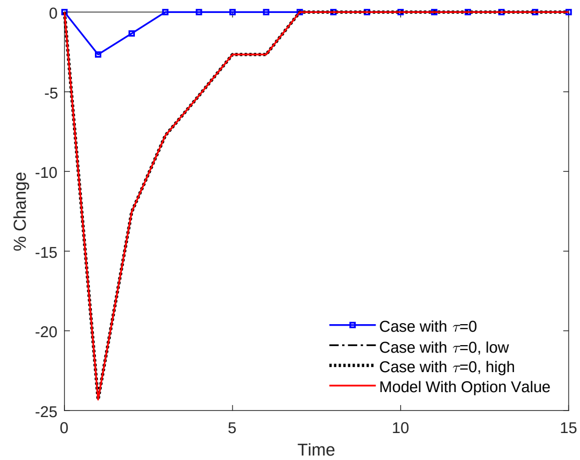
(f) Output

Notes: The figure shows the response of the baseline economy to a one time negative aggregate demand shock. The magnitde of the shock is chosen to result 3% decline in entry rate, observed during the Great Recession. 'Model with Option Value' refers to the baseline model. 'Case with $\tau = 0$' refers to case when in the baseline model $\tau = 0$. The 'Case with $\tau = 0$, low' ('Case with $\tau = 0$, high') refers to a case when in 'Case with $\tau = 0$' lowest (highest) productive entrants are adjusted to match the number as in the model with option value.

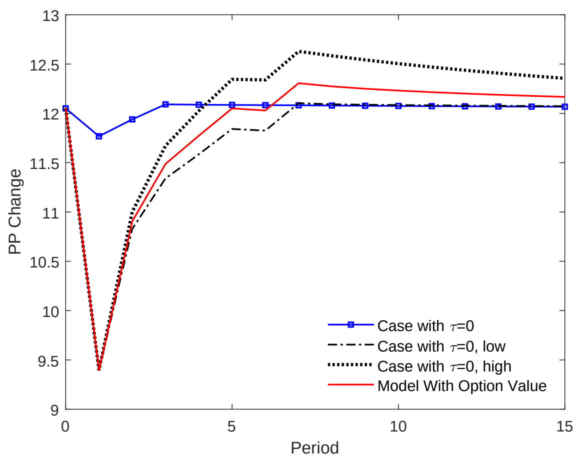
Figure 26: Impulse response to persistent aggregate demand shock



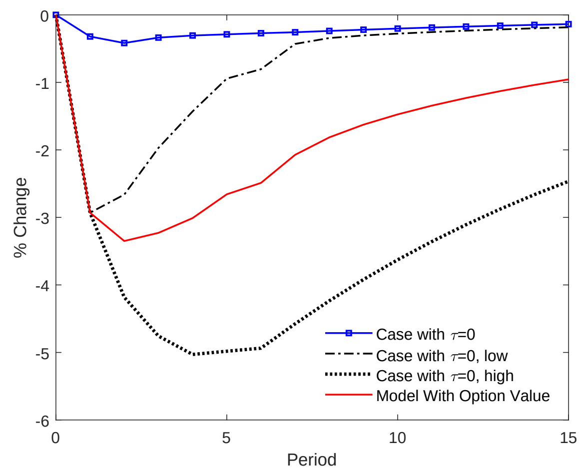
(a) Shock Process



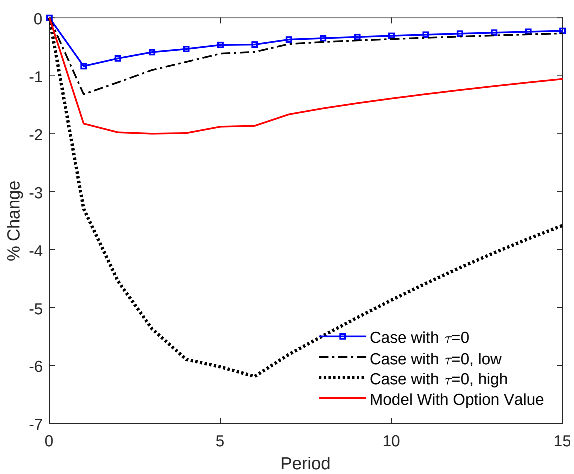
(b) Number of Entrants



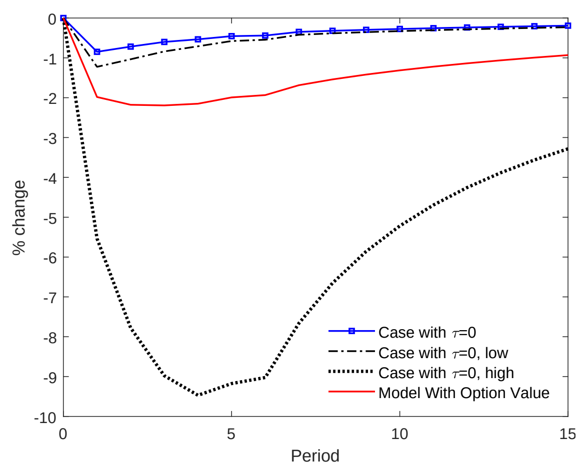
(c) Entry Rate



(d) Number of Firms



(e) Employment



(f) Output

Notes: The figure shows the response of the economy to a persistent decline in aggregate demand level. The magnitde of the shock is chosen to result 3% decline in entry rate, observed during the Great Recession. 'Model with Option Value' refers to a baseline scenario. 'Case with $\tau = 0$' refers to a case with $\tau = 0$. The 'Case with $\tau = 0$, low' ('Case with $\tau = 0$, high') refers to a case when in 'Case with $\tau = 0$' lowest (highest) productive entrants are adjusted to match the variation in the entrants to the baseline scenario

Table 10: Propagation: 1-time shock

		The Model (a)	$\tau = 0$ (b)	No. of entrant as in The Model: lowest productive (c)	No. of entrant as in The Model: highest productive (d)
Depth (%)	Employment	-1.83	-0.83	-1.32	-3.3
	No. of Firms	-2.93	-0.32	-2.93	-2.93
50% Rec.	Employment	3	2	2	11
	No. of Firms	3	4	3	10
75% Rec.	Employment	16	5	5	23
	No. of Firms	9	14	4	19

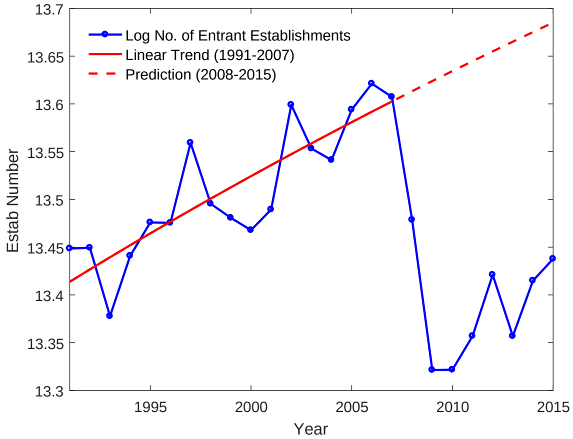
Note: 'The model' refers to a baseline scenario. '$\tau = 0$' refers to the Case with $\tau = 0$. And the 'CF' refers to the Case with $\tau = 0$ but adjusted number of entrants to match 'The model'. 'Depth' refers to the magnitude of the drop in variables during the 'Trough'. 50% Recovery (75% Recovery) describes number of periods (in years) after which economy recovers up to 50% (75%) from the 'depth'. The magnitude of the shock is chosen to generate 25% decline in the number of entrant firms, similar to the one observed during the Great Recession.

Table 11: Propagation: persistent shock

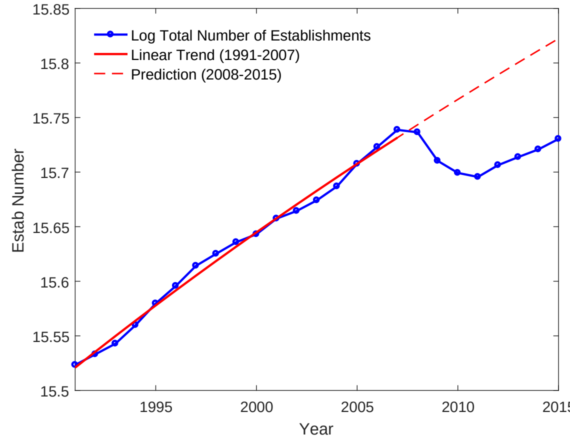
		The model (a)	$\tau = 0$ (b)	No. of entrant as in The Model: lowest productive (c)	No. of entrant as in The Model: highest productive (d)
Depth (%)	Employment	-2.00	-0.83	-1.31	-5.03
	No. of Firms	-3.05	-0.42	-2.93	-9.5
50% Rec.	Employment	16	7	5	18
	No. of Firms	9	10	4	11
75% Rec.	Employment	28	17	12	29
	No. of Firms	17	19	7	24

Note: 'The model' refers to a baseline scenario. '$\tau = 0$' refers to the Case with $\tau = 0$. And the 'CF' refers to the Case with $\tau = 0$ but adjusted number of entrants to match 'The model'. 'Depth' refers to the magnitude of the drop in variables during the 'Trough'. 50% Recovery (75% Recovery) describes number of periods (in years) after which economy recovers up to 50% (75%) from the 'depth'. The magnitude of the shock is chosen to generate 25% decline in the number of entrant firms, similar to the one observed during the Great Recession.

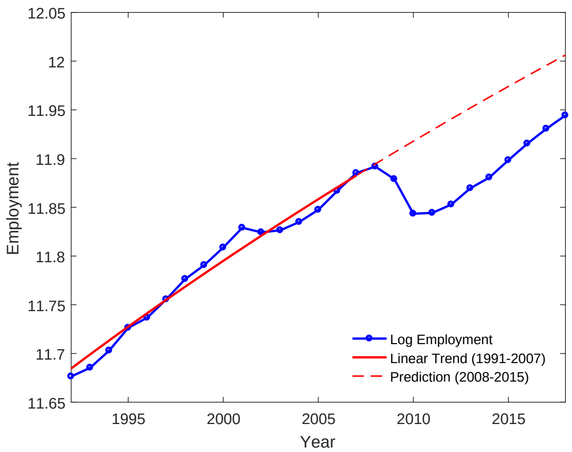
Figure 27: Deviations from the pre-crisis trend



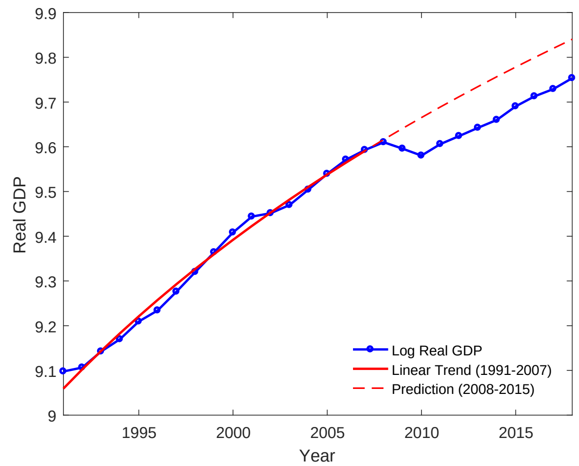
(a) No. of entrant estabs.



(b) Total no. of estabs.



(c) Total employment



(d) Real GDP

Figure 28: The number of entrant establishments across sectors and over time relative to the number of entrant establishments in respective sector in year 2007.

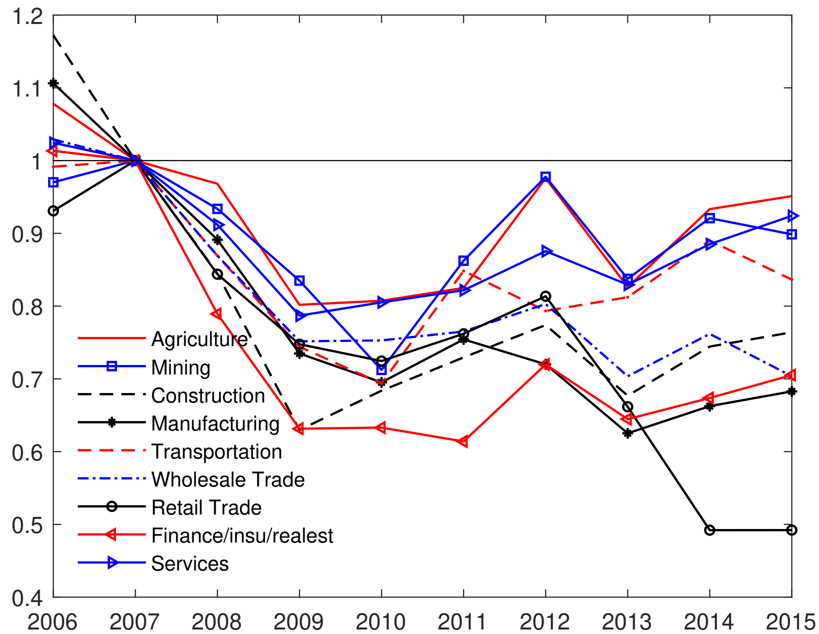
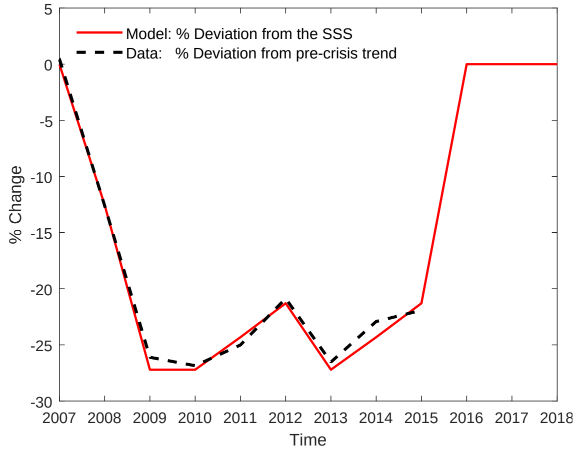
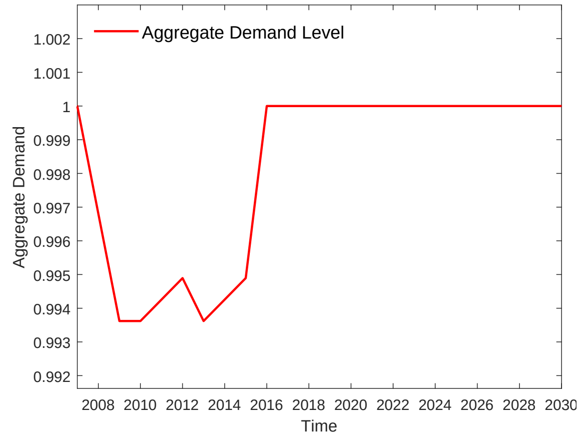


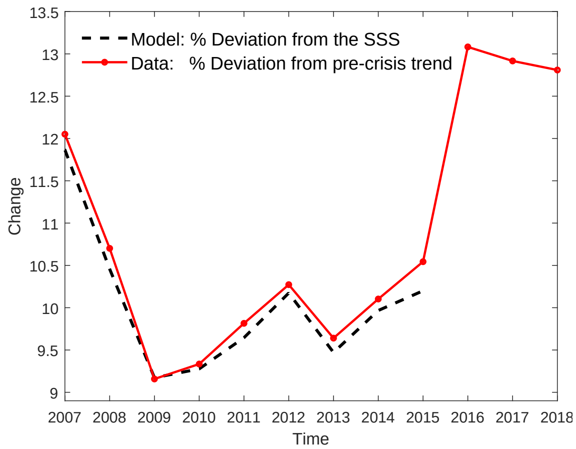
Figure 29: The Great Recession



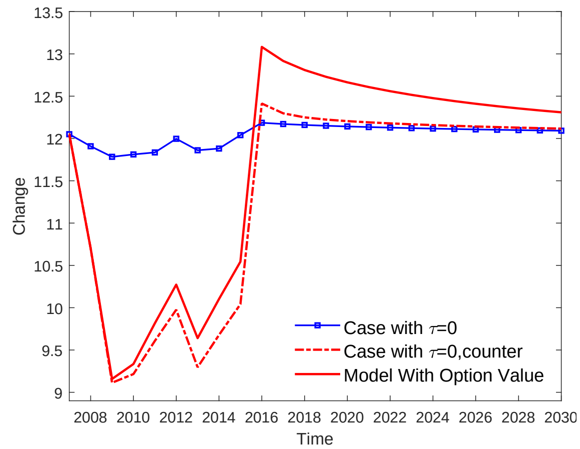
(a) No. of Entrantsnt



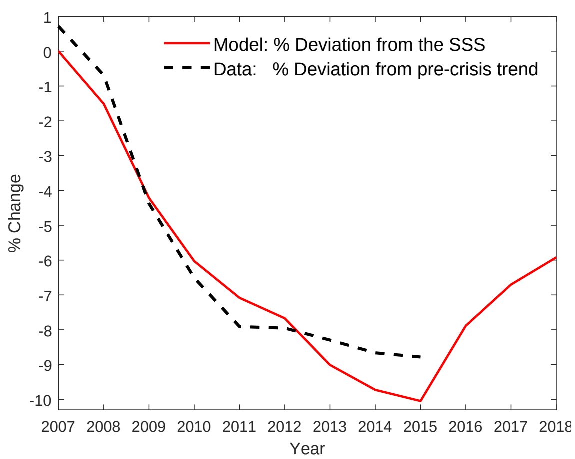
(b) Constructed Shock Process



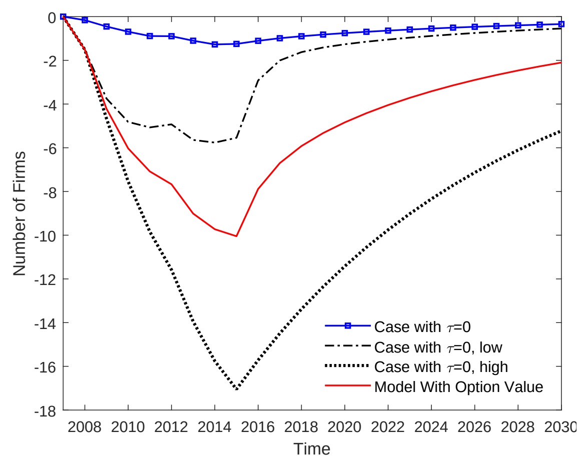
(c) SR: Entry Rate



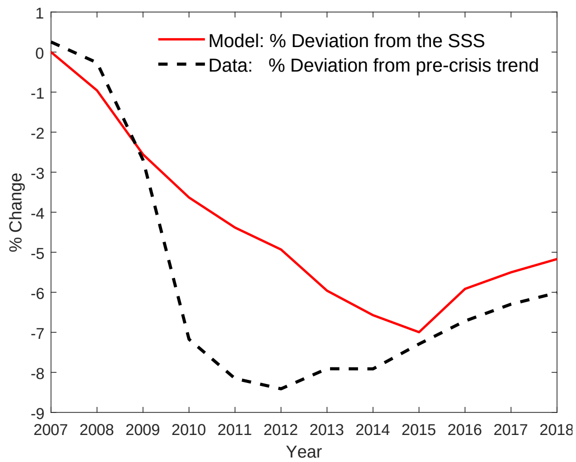
(d) LR: Entry Rate



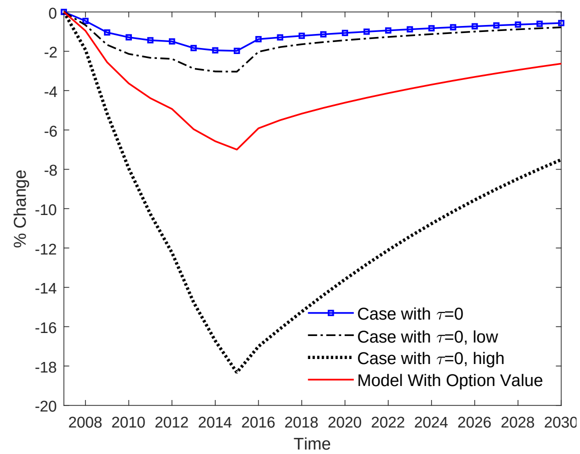
(e) SR:Establishments



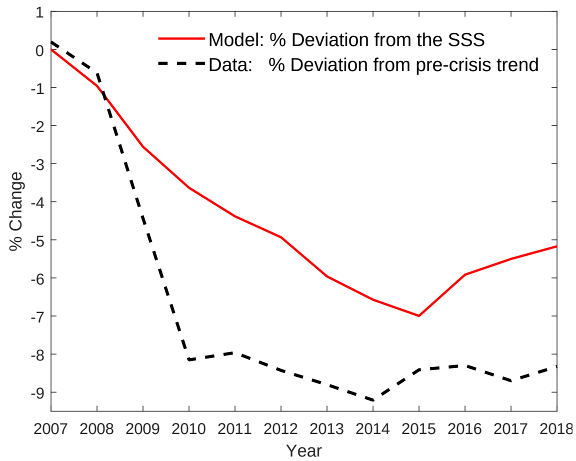
(f) LR: Establishments



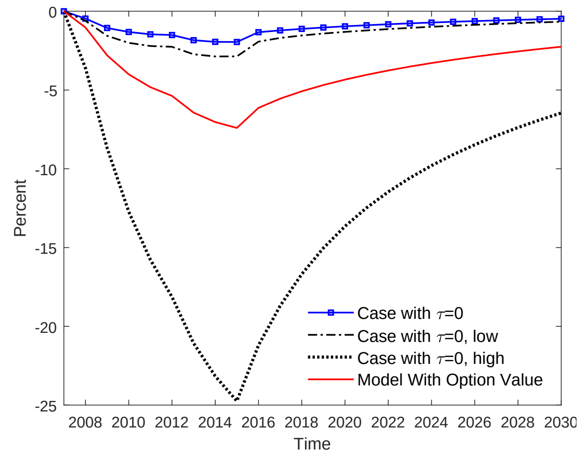
(g) SR: Employment



(h) LR: Employment

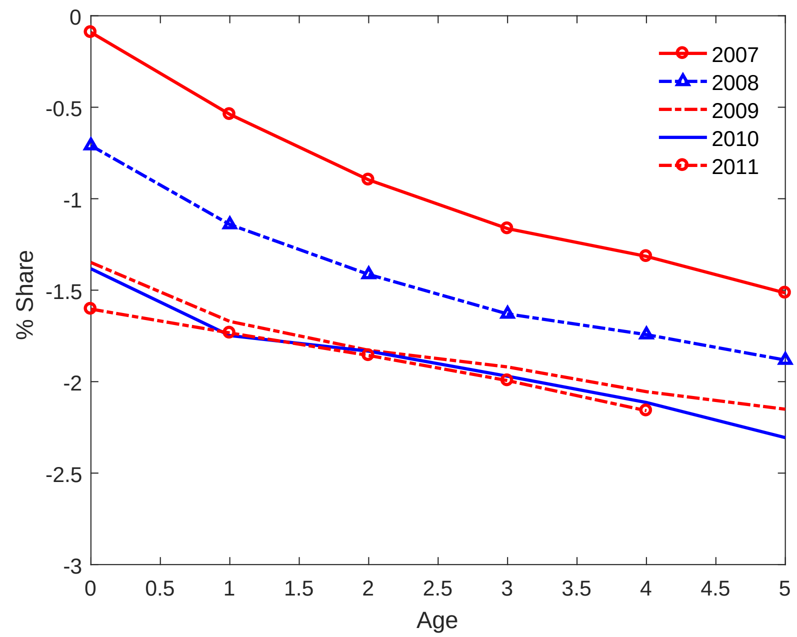


(i) SR: Change in real GDP

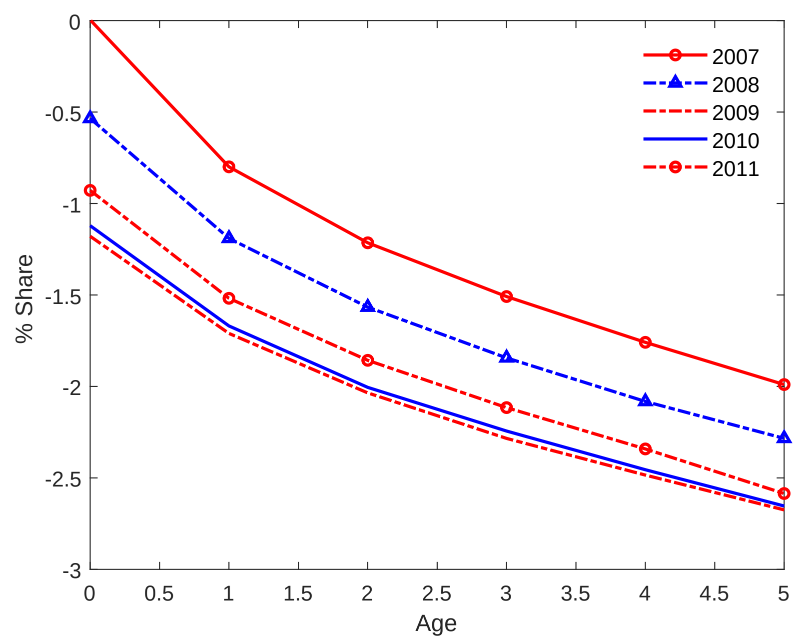


(j) LR: Change in Output

Figure 30: Cohort Employment Share in Total Employment

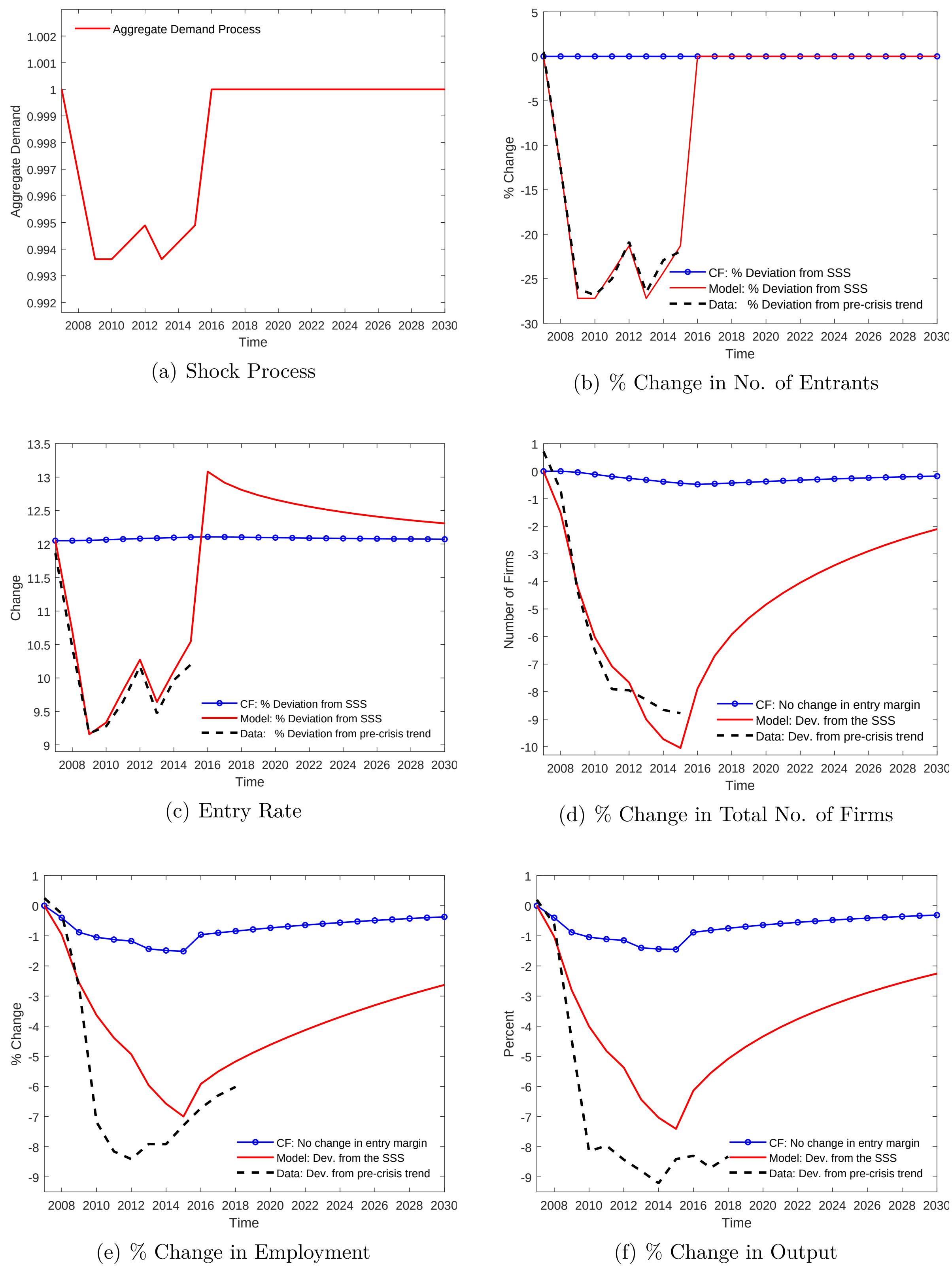


(a) Data



(b) The Model

Figure 31: Counterfactual: the effect of the persistent decline in number of entrant establishments.



(a) Shock Process

(b) % Change in No. of Entrants

(c) Entry Rate

(d) % Change in Total No. of Firms

(e) % Change in Employment

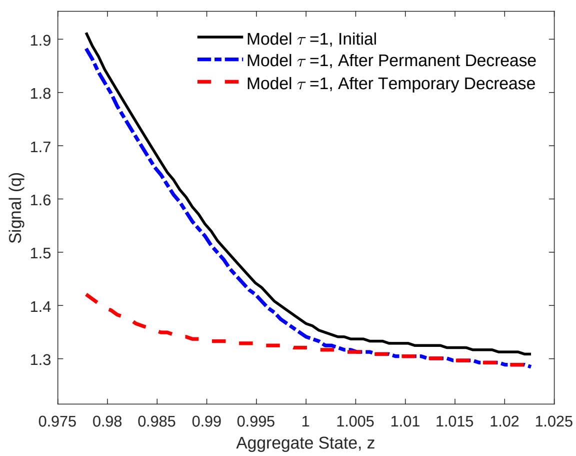
(f) % Change in Output

Table 12: The Great Recession

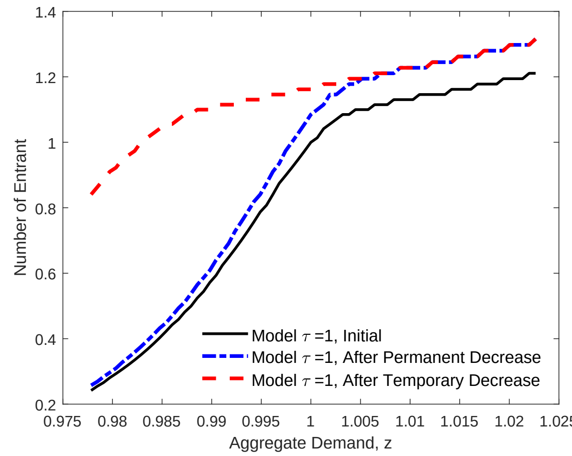
		2007	2012	2015	2018	Recovery (year) 50%	75%
No. of Entrants	The Data	+2.98	-25.01	-22.92	-	-	-
	The model	0	-25.01	-22.92	0	2016	2016
	CF: fixed entry	0	0	0	0		
No. of Firms	The Data	+0.72	-7.95	-8.78*	-	-	-
	The model	0	-7.67	-10.0*	-5.92	2019	2027
	CF: fixed entry	0	-0.26	-0.43*	-0.43		
Employment	The Data	+0.25	-8.41*	-7.29	-6.01	-	-
	The model	0	-4.93	-7.00*	-5.17	2024	2035
	CF: fixed entry	0	-1.17	-1.51*	-0.84		
Output	The Data	+0.20	-8.43*	-8.41	-8.33	-	-
	The model	0	-5.37	-7.41*	-5.08	2022	2033
	CF: fixed entry	0	-1.15	-1.45*	-0.75		

Note: The values indicate (a) % deviation of the time-series of the respective variables from the pre-crisis trend. (b) % deviation of the simulated aggregate variables from the stochastic steady state levels as a response of the constructed shock series. Shock affects the economy starting from year 2008. 'The data' refers to the statistics from the data. 'The model' refers to the statistics from the baseline scenario. 'CF: fixed entry' refers to the case when the composition and number of entrants are fixed at the stochastic steady state level. The numbers in bold refer to the periods when the variables reached the 'depth'. 50% Recovery (75% Recovery) describes number of periods (in years) after which economy recovers up to 50% (75%) from the 'depth'.

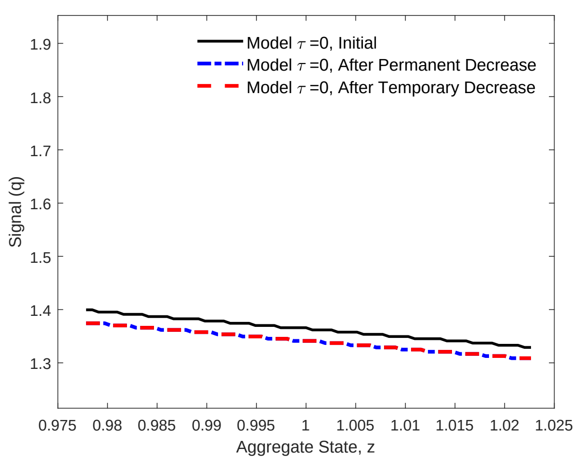
Figure 32: Change in threshold signal and number of entrant firms as a response of the permanent/temporary decline in fixed entry cost across aggregate states.



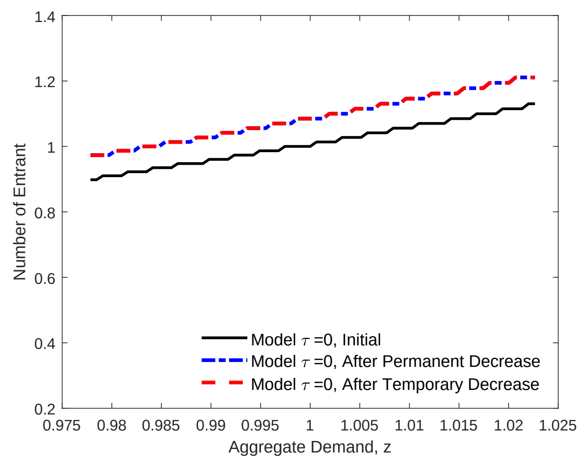
(a) Threshold signal, $\tau = 1$



(b) Number of firms, $\tau = 1$



(c) Threshold signal, $\tau = 0$



(d) Number of firms, $\tau = 0$

Figure 33: Change in the threshold signal over time as a response of an anticipated decline in entry barriers that is going to take place after five years from the announcement.

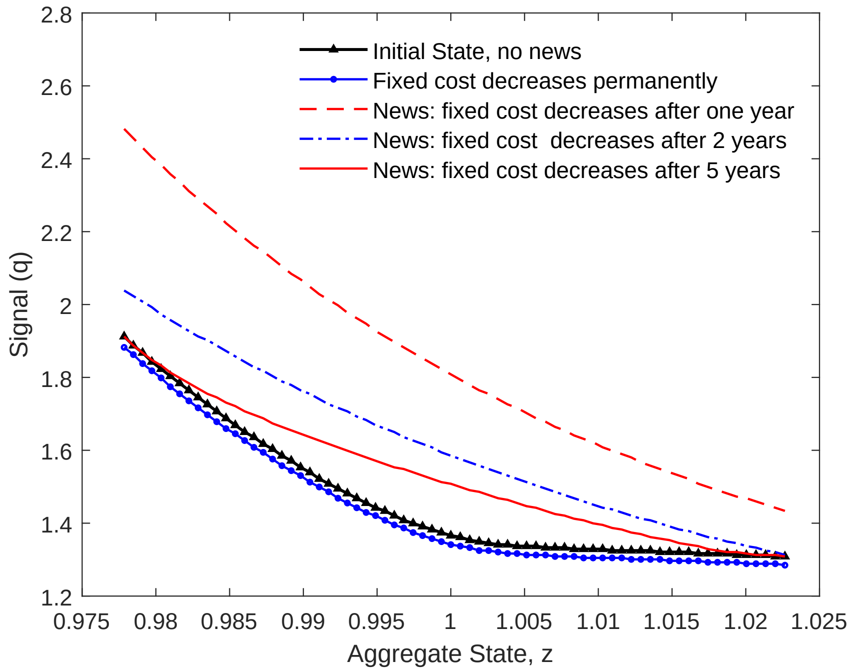
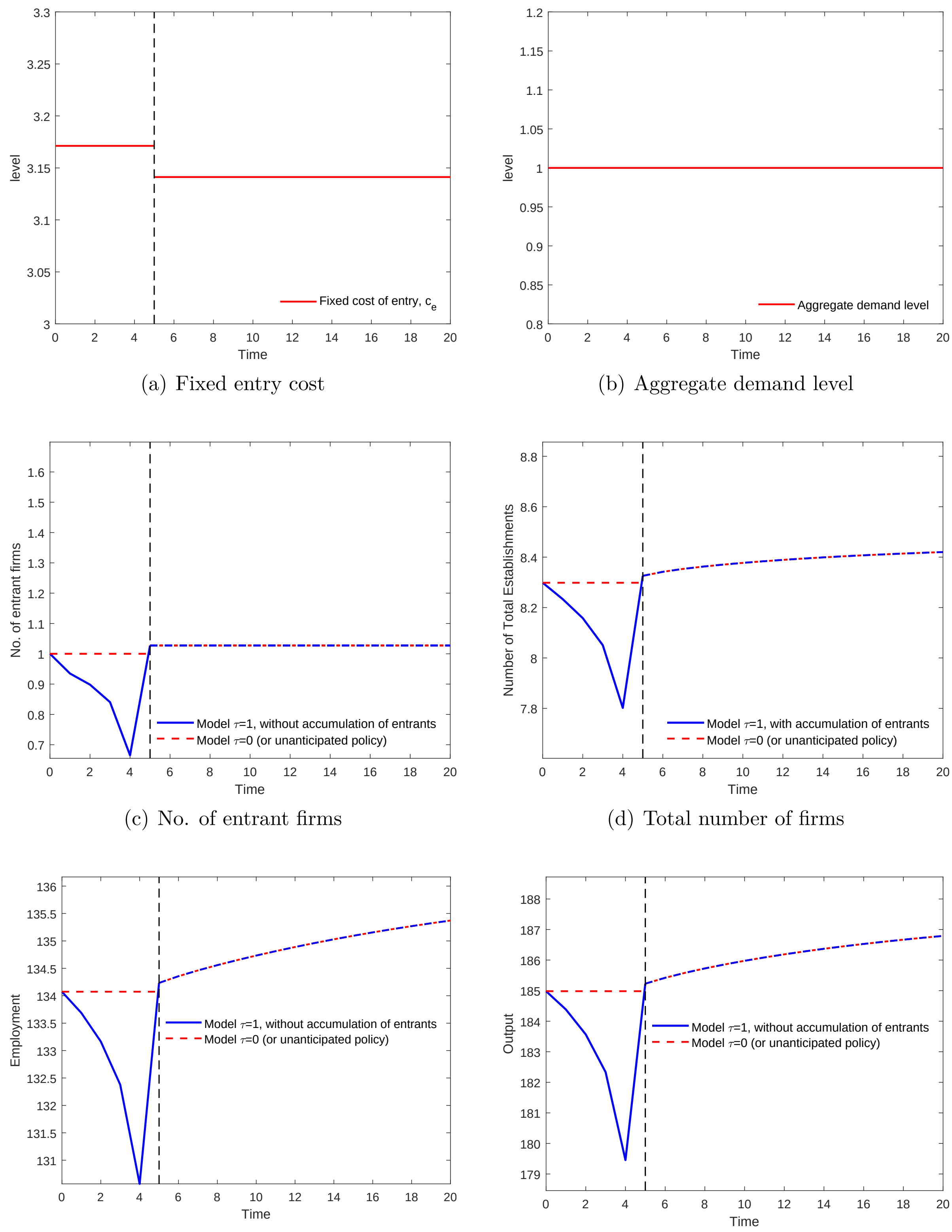


Figure 34: Response of the aggregate variables to an anticipated decline in entry barriers that is going to take place after five years from the announcement (without signal accumulation).



(a) Fixed entry cost

(b) Aggregate demand level

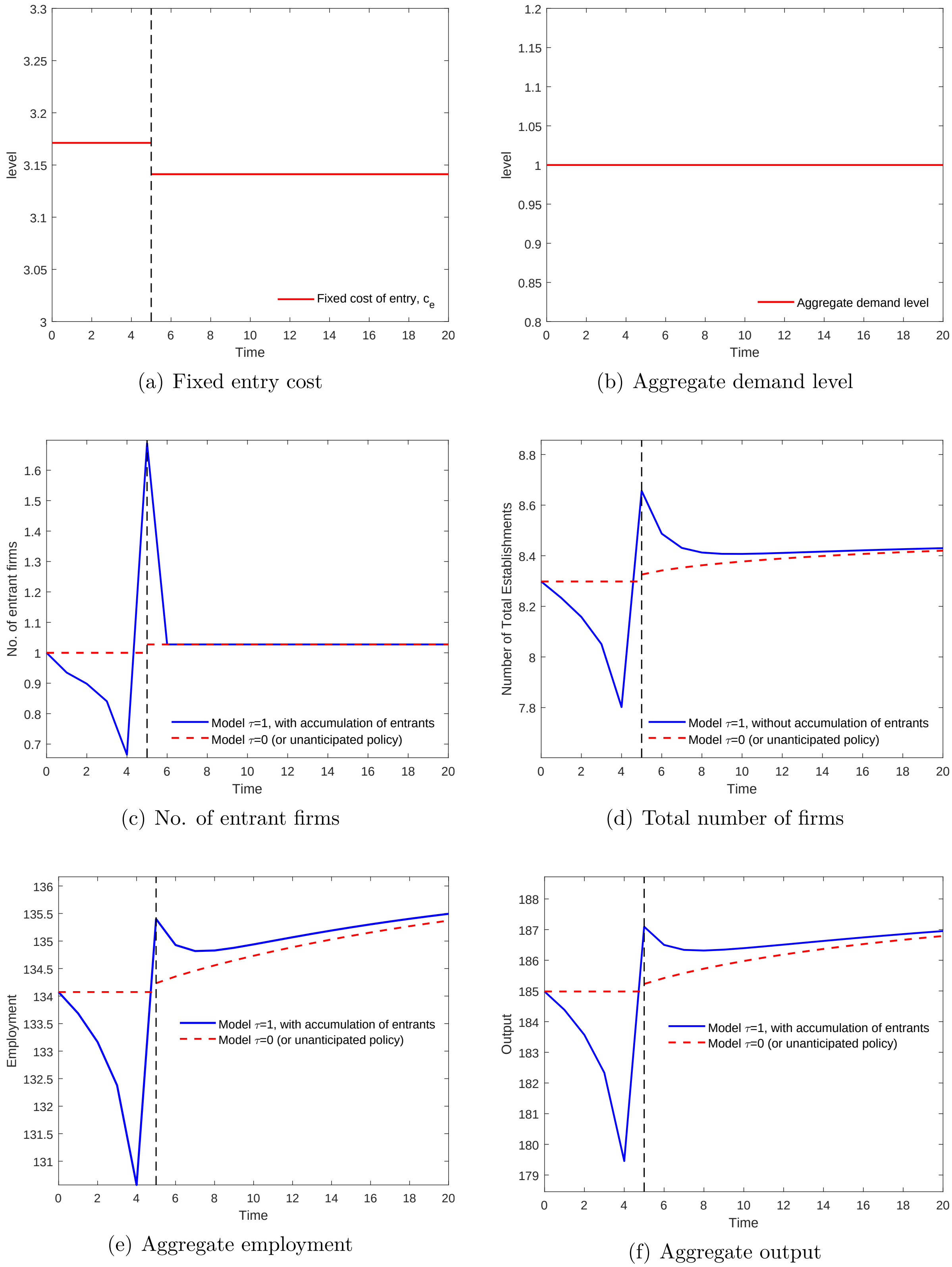
(c) No. of entrant firms

(d) Total number of firms

(e) Aggregate employment

(f) Aggregate output

Figure 35: Response of the aggregate variables to an anticipated decline in entry barriers that is going to take place after five years from the announcement (with signal accumulation).



(a) Fixed entry cost

(b) Aggregate demand level

(c) No. of entrant firms

(d) Total number of firms

(e) Aggregate employment

(f) Aggregate output

Appendices

A Model appendix

A.1 Extensions

A.1.1 Two-stage entry phase

In the section I provide an extended description of the entry phase that justifies the assumption about the constant mass of potential entrants.

Every period there is a limited mass of heterogeneous business opportunities that can be implemented into the market by potential entrants. Each of the business opportunities is characterized with a signal q, which describes the productivity of a business opportunity after it is implemented in the market. For a given signal q the distribution of the initial period productivity is given by $H_e(s|q)$. The higher the signal, the higher the expected first period productivity of a business opportunity. The distribution of business opportunities over the signal is time-invariant and is given by $q \sim W(q)$.[100]

Entry phase consists of two stages. During the first stage an infinite mass of individuals make decisions whether to compete or not for the available business opportunities. To participate in the competition individuals need to pay fixed cost c_q, after which they are free to direct their search for a particular group of business opportunities characterized with a signal q. Since there is a limited number of business opportunities within each signal category, not all aspiring startups end up receiving a signal. During the second stage, those aspiring startups that receive a signal about the business opportunities become potential entrants and make entry decisions. The signal is persistent over time, which gives a potential entrant the ability to exercise the business opportunity in the future instead of today. If a potential entrant with a signal q postpones entry to the next period, the potential entrant gets the same signal tomorrow with a probability $\tau \in [0, 1]$. With a probability $(1 - \tau)$ the potential entrant loses the signal and drops out from the pool of potential entrants.

[100] The distribution is such that the mass of business opportunities with signal q decreases with q.

After describing the entry phase the decision process for each stage is as follows.

Stage 1. The expected value of attempting to seize a business opportunity with a signal q equals to

$$V^o(q,z) = \frac{B_t(q)}{n_t(q)} V^e(q,z_t) - c_q \tag{7}$$

where $B_t(q)$ is a mass of available business opportunities with quality q at time t.[101] The total mass of available business opportunities equal to the total number of business opportunities within each signal category $W(q)$ minus the mass of business opportunities that is held by the group of potential entrants that delayed entry in the previous periods. $n_t(q)$ refers to a number of aspiring startups competing for the business opportunities with signal q. The ratio in the equation represents a probability by which an individual receives a signal q and becomes a potential entrant.[102] $V^e(q,z_t)$ is a value of a potential entrant with signal q at time t.

If $V^e(q,z_t) < c_q$ individuals do not compete for the business opportunities with signal q. A positive mass of individuals decide to pay fixed cost c_q and compete for a signal q if $V^e(q,z_t) > c_q$. Due to the free entry the number of individuals $n_t(q)$ competing for each signal is such that $V^o(q,z) = c_q$.

Denote $\underline{q_t}$ a signal at time t that satisfies $V^e(\underline{q_t}, z_t) = c_q$. Since the value of entry increases with a signal level, aspiring startups choose to compete for business opportunities with signal level $q > \underline{q_t}$. The total number of individuals attempting to get the business opportunities equals to

$$N_{t,\text{aspiring startups}} = \int_{\underline{q_t}} n_t(q) dq$$

Stage 2. Stage 2, where potential entrants make entry decisions, is same as in the main description of entrants.

While modeling the entry phase I use information from the newly developed Business Forma-

[101] $0 < B_t(q) < W(q)$

[102] $0 \leq \dfrac{B_t(q)}{n_t(q)} \leq 1$.

tion Statistics dataset that collects information about business applications and formation. Business application data is based on applications for Employer Identification Number (EINs) filed in the United States. Using the dataset Bayard et al. (2018) shows a significant lag between the dates of the applications and the dates the applications transition into employer businesses. More specifically, from the applications that are referred as 'high propensity business applications' only 13% transitions into employer businesses within the first four quarter window and 15% in eight quarter window.

In the entry phase described above the number of applications is associated with the number of aspiring start ups. I choose c_q, the fixed cost that individuals need to pay to become aspiring startups, so that the share of the actual entrants in total number of aspiring startups is 13%. The value corresponds to $c_q = 0.022$. The way of modeling explains the low transition rates from the business applications to employer businesses partly due to the restricted number of actual business ideas.

The data also indicates that only additional 2% of the applications transitions into employer businesses in the following year. In terms of the model setup the fact implies that $B(q)$ is close to $W(q)$; only few potential entrants that decide to delay entry enters the market next period.[103]

To conclude, the restriction on the number of available business opportunities implies that the aggregate distribution of potential entrants are constant over time, and accumulation of the entrants happens at aspiring start up level.

[103]The ability to delay entry is an option for a potential entrant and does not require the potential entrant to enter the market in the future; Explaining the reasons behind what makes potential entrants actually come back or not come back in the market after delaying entry is beyond the scope of this paper and is left for the future research.

A.1.2 Accumulation of potential entrants

In this section, I relax the assumption that keeps the aggregate distribution of potential entrants constant in the baseline model. I investigate how the accumulation of potential entrants, that decide to delay entry, modifies entrants' characteristics over the cycles and affects the dynamics of aggregate variables. I find that cohorts that enter during different aggregate economic conditions have significantly and persistently different characteristics even after allowing accumulation of potential entrants over time.

Figure 36: New potential entrants, $W(q)$

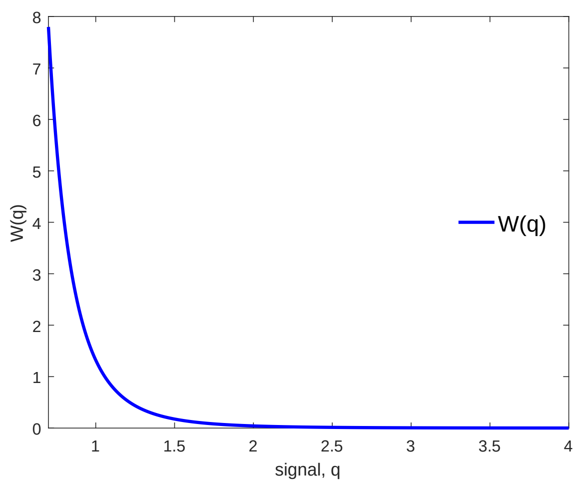
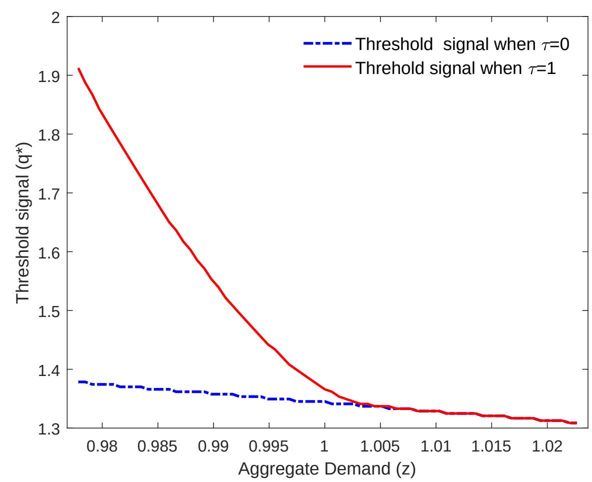


Figure 37: Threshold signal



In the baseline model, in every period, the distribution of new potential entrants, that make entry decisions for the first time, equal to the distribution of potential entrants that entered into the market in the previous period. The assumption ensures that the number of potential entrants is constant over time, and the aggregate distribution of potential entrants over the signal is time-invariant and is given by $W(q)$. In this section, I relax the assumption in the following way. In the beginning of every period constant mass of new potential entrants are born and make entry decisions. The distribution of new potential entrants over the signal is given by $W(q)$, see Figure 36. In addition to the new potential entrants, the aggregate distribution of potential entrants consists of old potential entrants. Old potential entrants are those that decided to delay entry in the previous periods, while their expected value of being an incumbent was more than zero.[104] Figure 37 displays the threshold signal, $q^*_\tau(z)$ for

[104] Consistent to the baseline model I keep $\tau = 1$, which means that potential entrants that delay entry are

each aggregate state when $\tau = 0$ (blue-dashed line) and $\tau = 1$ (red solid line). For given z, potential entrants that decide to delay entry hold signals in between $[q^*_{\tau=0}(z) \quad q^*_{\tau=1}(z)]$.

The distribution of old potential entrants evolves endogenously and depends on the realization of the aggregate states in the previous periods. Denote the mass of old potential entrants with signal q at the beginning of period t with $\Lambda_t^{\text{old entrants}}(q)$.

$$\Lambda_t^{\text{old entrants}}(q) = \sum_{k=0}^{t} W(q)\ 1\left\{q^*_{\tau=0}(z_k) \leq q < q^*_{\tau=1}(z_k)\right\} + \Lambda_0^{\text{old entrants}}(q)$$

Where $\Lambda_0^{\text{old entrants}}(q)$ denote the distribution of old potential entrants at time 0.

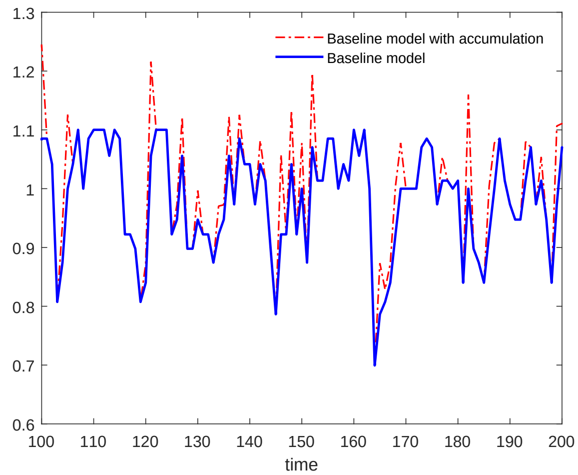
Then, the total mass of potential entrants with signal q at the beginning of period t, $\Psi_t(q)$ is given by

$$\Psi_t(q) = W(q) + \Lambda_t^{\text{old entrants}}(q)$$

Figure 38 compares the dynamics of the entrants in the baseline model to a model that allows accumulation of potential entrants. Note that when the aggregate demand decreases from z_{t-1} to z_t then the distribution/number of entrants in the baseline model and in the model with signal accumulation coincides with each other. If potential entrants delayed entry when the aggregate state was z_{t-1}, nobody from these old potential entrants is going to enter in an aggregate state z'_t ($< z_{t-1}$). As a result, selection of potential entrants at entry happens only from the distribution of new potential entrants, like in the baseline model. However, if the aggregate demand level increases from period $t-1$ to period t in addition to new potential entrants, some of the old potential entrants also decide to enter into the market, resulting higher number of entrants in the model with signal accumulation compared to the baseline model.

able to keep the signal forever.

Figure 38: Entrant firms



It turns out that the increase in the number of entrants during expansions outweighs increase in number of entrants during recessions and extending the baseline model to account for the accumulation of potential entrants increase procyclical variation in the entry rate. Moreover, the differences in cohorts' characteristics that start operating during different aggregate economic conditions increases after allowing for the accumulation of potential entrants. The latter feature modifies the distribution of the entrants over the cycles in the following way. Now, during recessions, defined as periods when $log(z) < 0$, potential entrants that enter the market hold lower signals on average compared to the baseline scenario. Consequently, as shown in Figure 39 and Figure 40, average productivity and survival rates of the cohorts that enter the market during recessionary periods decrease compared to the baseline scenario. The decrease is minor, since the option value of delay mechanism is quite strong during recessionary periods and potential entrants that decide to stop waiting hold the highest signal levels. Those accumulated group of old potential entrants that decide to delay entry until the expansions hold on average less productive signals and most of them ends up to be low productivity firms after entering the market. Consequently, average productivity and survival rate decreases significantly during expansionary periods compared to the baseline scenario. Altogether, the extension produces not only countercyclical average productivity and survival rates, but also the difference between the cohorts that enter the market during recessionary and expansionary periods increases compared to the baseline model.

Figure 39: Average productivity

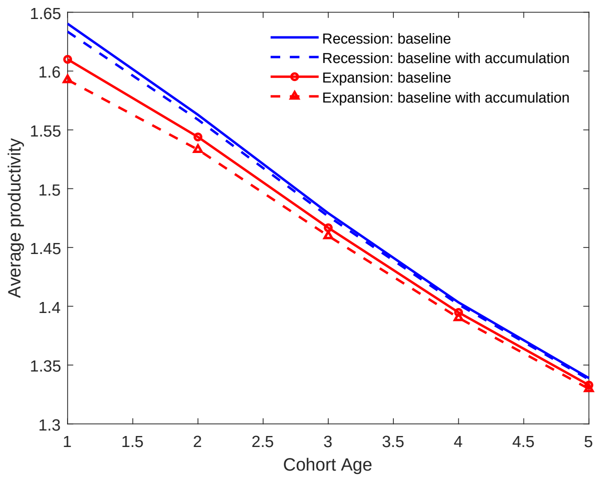


Figure 40: Average survival rate

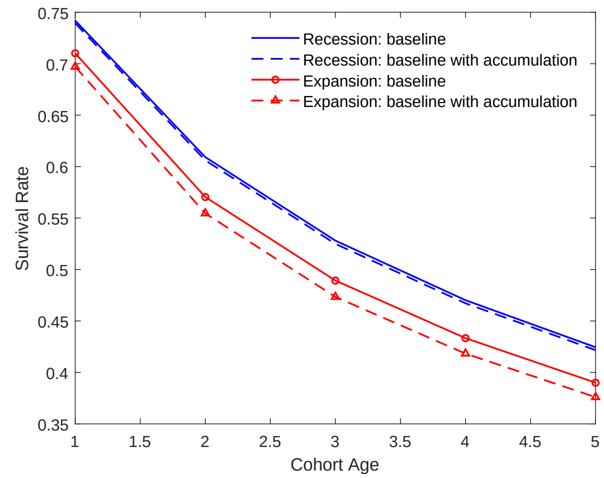
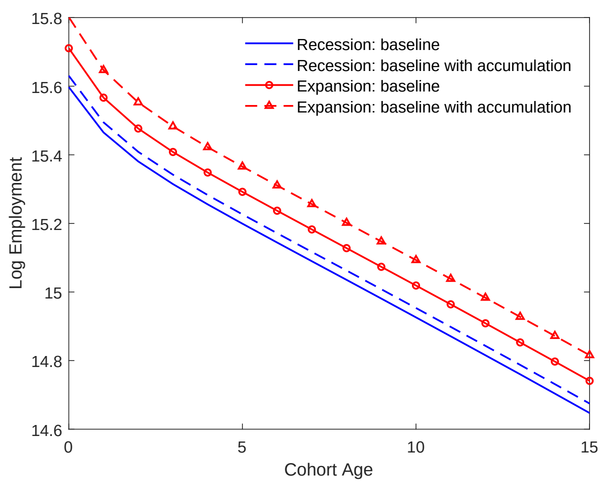


Figure 41: Cohort-level employment



Allowing accumulation of potential entrants over time increases recessionary as well as expansionary cohorts employment compared to the baseline model. However, since the number of entrants significantly increases during expansionary periods the difference between recessionary and expansionary cohorts employment increases compared to the baseline scenario.

A.2 General equilibrium

A.2.1 Consumers

The economy is populated by a unit mass of atomistic, identical household. At time t, the household consumes the basket of goods C_t, defined over a continuum of goods Ω. At any given time t, only subset of goods $\Omega_t \subset \Omega$ is available. Let $p_t(\omega)$ denote the nominal price of a good $\omega \in \Omega_t$.

First layer maximization:

$$\max_{\left(C_t, L_t, (c_t(\omega))_{\omega \in \Omega_t}\right)_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma} - 1}{1-\sigma} - \chi(L_t) \right]$$

s.t.

$$P_t C_t = P_t w_t L_t + \Pi_t$$

Second layer maximization:

$$\max_{(c_t(\omega))_{\omega \in \Omega_t}} C_t = \left(\int\limits_{\omega \in \Omega_t} (\alpha z_t)^{\frac{1}{\rho}} b_t(\omega)^{\frac{\gamma}{\rho}} c_t(\omega)^{\frac{\rho-1}{\rho}} d\omega \right)^{\frac{\rho}{\rho-1}}$$

s.t.

$$\int\limits_{\omega \in \Omega_t} p_t(\omega) c_t(\omega) d\omega \leq P_t C_t$$

A.2.2 Households' budget constraint

Households hold shares in a mutual fund of firms.

The value of mutual fund, Λ_t at the beginning of time t, after entry and exit has occurred:

$$\Lambda_t = \int_s \int_b V(s,b,z)\Omega(b,s,z)dsdb + \int_{q_*}^{\infty} \int_s V(b_o,s,z)H(s|q)W(q)dq$$

Denote $N_{e,t}$ be the number of entrants in period t, than: $N_{e,t} = \int_{q_*}^{\infty} W(q)dq$

At the end of the period value of mutual fund is

$$\Lambda_t' = \Pi - N_{e,t}c_e + (\Lambda_t - \Pi)$$

Let $x_t \in [0,1]$ was the share household decides to hold of the mutual fund in period t. Then, household budget constraint will be

$$\Lambda_t x_t + C_t \leq [\Pi - N_{e,t}c_e + (\Lambda_t - \Pi)]\, x_t + L_t P_t w_t$$

The optimal solution implies that if $\Lambda_t \geq 0$ then $x_t = 1$

Which reduces HH budget constraint to

$$P_t C_t + P_t N_e c_e = P_t w_t L_t + \Pi_t$$

A.2.3 Incumbent Firms

Incumbent firms are distributed over consumer capital (b) and productivity (s). The distribution given by $\Omega_t(s,b)$.

At time t when aggregate state equals to z, an incumbent firm characterized by (s,b) takes as given real wage w, aggregate price index P and solves the following functional equation:

$$V^I(b,s,z) = \max_{y,p,b} \quad py - P\frac{w}{s} + \int max\left\{0, -Pc_f + \tilde{\beta}(1-\gamma)E[V^I(b',s',z')|s,z]\right\} dG(f)$$

s.t.

Production technology:	$y_t^s = s_t n_t$
Demand Constraint:	$y_t^d = \alpha z_t b_t^\eta \left(\frac{p_t}{P_t}\right)^{-\rho} Y_t$
Motion of Consumer Capital:	$b_{t+1} = (1-\delta)\,(\,b_t + y_t p_t\,)$
Fixed production cost :	$c_f \sim G(f)$. c_f is in consumption units.
Motion of id. Productivity :	$log(s_{it}) = \rho_s log(s_{it-1}) + \sigma_s \varepsilon_{it}$
Motion of agg. demand shock:	$log(z_t) = \rho_z log(z_{t-1}) + \sigma_z \epsilon_t$

A.2.4 Potential Entrants

Potential entrants are characterized with signal, q about their initial productivity. At any t, density of potential entrants over q is constant and is given by $W(q)$.

To enter into the market the potential entrant needs to pay fixed entry cost in consumption units c_e (value $P_t c_e$). Upon entry the potential entrant observes actual idiosyncratic productivity (s), receives fixed initial capital stock (b_0) and behaves like an incumbent with (b_0, s).

At time t when aggregate state equals to z, a potential entrant with signal q takes correctly anticipates aggregate price index P and solves the following problem:

$$V^e(b_0, q, z) = max\left\{\tau\ \tilde{\beta} E[V^e(b_0, q, z')|z],\quad -Pc_e + \int_s V^I(b_0, s, z) dH_e(s|q)\right\}$$

A.3 Numerical approximation

The following section describes numerical solution algorithm used to solve the model.

A.3.1 Incumbent's value function

1. Define grid points for the state variables s, z, and b.

The grids and the transition matrices for the idiosyncratic productivity shock s and the aggregate demand shock z are constructed following the Rouwenhorst's method. Denote the number of grid points as I_s and I_z, and the probability transition matrices as $P^s(s'|s)$ and $P^z(z'|z)$, respectively. To construct grid points for the customer capital I use equally distributed grid points on a logarithmic scale on the interval $[b_0,\ \ b_{max}]$. b_0 is chosen to match entrants' average size. b_{max} is chosen so that employment by large firms is more than 1000+ , the size of establishments in the highest size bin in the BDS dataset. Denote number of customer capital grid points as I_b.

2. For all grid points (b, s, z) guess values for the incumbent firm's value function $V_0^I(b, s, z)$.

3. Construct a revised guess value function $V_1^I(b, s, z)$ by solving:

$$V_1^I(b, s, z) = \max_b \left\{ \Pi(b, s, z) + G(c_f^*) \left(\beta(1-\gamma)E[V_0^I(b', s', z')|s, z] - E[c_f|c_f < c_f^*] \right) \right\}$$

subject to

$$\Pi(b, s, z) = \left(\frac{b'}{1-\delta} - b \right) - \frac{w}{s} \left(\frac{b'}{1-\delta} - b \right)^{\frac{\rho}{\rho-1}} b^{\frac{-\eta}{\rho-1}} \ (\alpha z)^{\frac{-1}{\rho-1}}$$

$$E[V_0^I(b', s', z')|s, z] = \sum_i \sum_j V_0^I(b', s_i, z_j) P^z(z_j|z) P^s(s_j|s)$$

Where $P^z(z_j|z)$ and $P^s(s_j|s)$ represents probabilities that next periods aggregate shock equals to z_j and idiosyncratic shock equals s_j.

Where c_f^* is the value of fixed cost which equals to incumbent's expected continuation value

$\beta(1-\gamma)E[V^{I*}(b', s', z')|s, z]$. In other words, when an incumbent firm receives c_f^*, the incumbent firm is indifferent between staying or exiting from the market.

4. Stop criteria: $|\frac{V_{n+1}^I(b,s,z) - V_n^I(b,s,z)}{V_n^I(b,s,z)}| \leq 10.0^{-8}$

A.3.2 The signal

Define grid points for the signal.

I use Guass-Legendre quadrature method over the $[\underline{q},\ q_{max}]$ interval with I_q number of nodes to generate grid points q and weights w_q for the signal.

1. The aggregate signal distribution $W(q)$ has Pareto Distribution with location Parameter $\underline{q}$ and Pareto exponent ξ. I approximate the mass of potential entrants denoted by P_q, at each grid point of signal according to the following equation

$$P_q(q) = w_q(q)\ \xi \frac{\underline{q}^{\xi}}{q^{\xi+1}}$$

2. Next, initial idiosyncratic productivity $H(s|q)$ is constructed in the following way:

The idiosyncratic shock in the first period of operation follows Log-normal distribution. For each grid point $q_j \in I_q$ and $s_i \in I_s$ calculate $F(s_i|q_j)$, the probability that the entrant with signal q_j gets the initial productivity lower than s_i.

$$H(s_i|q_j) = \frac{1}{2}\left(F(s_i|q_j) - F(s_{i-1}|q_j)\right) + \frac{1}{2}\left(F(s_{i+1}|q_j) - F(s_i|q_j)\right)$$

For the initial and terminal grid points of the productivity I use:

$$H(s_1|q_j) = F(s_1|q_j) + \frac{1}{2}\left(F(s_2|q_j) - F(s_1|q_j)\right)$$

$$H(s_{I_s}|q_j) = \max\left(0, 1 - F(s_{I_s}|q_j)\right) + \frac{1}{2}\left(F(s_{I_s}|q_j) - F(s_{I_s-1}|q_j)\right)$$

Figure 42: $W(q)$

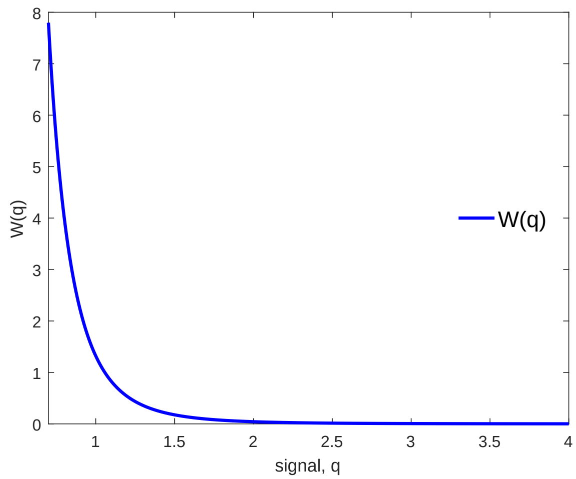
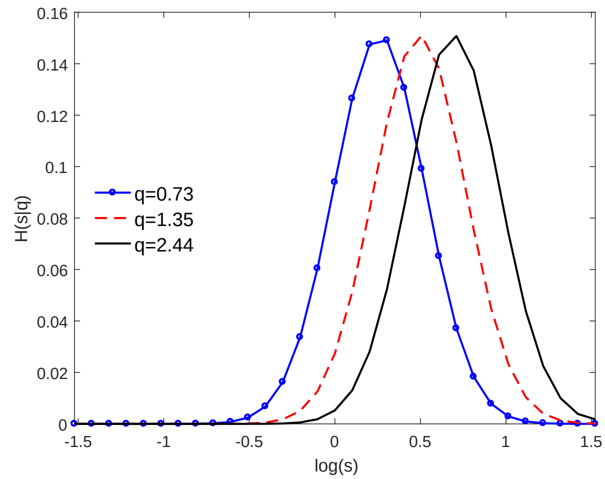


Figure 43: $H(s|q)$



A.3.3 Entrant's value function and the option value

1. For all grid points (q_j, z_k) calculate value of entering as

$$V^{gross}(b_o, q_j, z_k) = \sum_{i \in I_s} \Big[H(s_i|q_j) V^I(b_o, s_i, z_k) \Big]$$

2. To approximate value function of entry and option value of delay I use value function iteration.

a. Guess for the values of the entrant value function. $V_0^e(b_0, q, z)$

b. Given the guess find value of the option value of delay.

$$V^{Opt}(q, z) = \tau \;\beta E[V_0^e(b_0, q, z')|z] = \tau \;\beta \sum_{z_j \in I_z} V_0^e(b_0, q, z_j)$$

c. Update guess for value function of entry.

$$V_1^e(b_0, q, z) = \max \Big\{ V^{Opt}(q, z), V^e(b_0, q, z) - c \;\Big\}$$

d. Stopping criteria: $|\dfrac{V_{n+1}^e(b, s, z) - V_n^e(b, s, z)}{V_n^e(b, s, z)}| \leq 10.0^{-8}$

Denote the final value function of entry as: $V^e(b_0, q, z)$ and final option value function as $V^{Opt}(q, z)$.

3. Threshold value of signal $q^*(z)$ is defined as the value of signal, which makes potential entrants indifferent between entering and not entering into the market. So to find $q^*(z)$ one need to equate $V^e(b_0, q^*(z), z) = V^{Opt}(q^*(z), z)$

B Mathematical appendix

B.1 Proof of propositions

Proposition B.1.1. *(The properties of the net benefits of entry)*

(a) For given aggregate demand level z, NPV(q,z) strictly increases with the signal q.

(b) For given signal q, NPV(q,z) strictly increases with the aggregate demand level z.

Proof. $NPV(z,q) = V^{gross}(z,q) - c_e$, where $V^{gross}(z,q)$ equals to expected value of being an incumbent $\int_{s'} V^I(b_0, s', z) dH_e(s'|q)$ and c_e is entry cost that does not vary with the aggregate demand level.

(a) Expected value of being an incumbent is an increasing function of a signal q. Potential entrants first period distribution of the idiosyncratic productivity conditional on signal $H(s'|q)$ is a decreasing function of the signal q. Meaning that the higher the signal, the higher the expected first period productivity s. An incumbent's value function $V^I(b,s,z)$ is an increasing function of the idiosyncratic productivity shock s. Therefore, the expected value of being an incumbent and hence, $NPV(z,q)$ is a strictly increasing function of the signal q.

(b) $V^I(b,s,z)$ is an increasing function of the idiosyncratic productivity shock z. Therefore, for given signal q, expected value of being an incumbent, and hence $NPV(z,q)$ increases with the aggregate demand level z. □

Proposition B.1.2. *(The properties of the option value of delay)*

(a) Option Value of Delay(q,z) is non-negative for all q and z.

(b) For a given aggregate demand level z, Option Value of Delay(q,z) is a weakly increasing function of the signal q.

(c) For a given signal q, Option Value of Delay(q,z) weakly increases with the aggregate demand level z.

Proof. (a) Consider two groups of potential entrants which together span the space q. The first group of potential entrant which hold $q < \underline{q}$ do not enter the market for any z and always receive outside option value which equals to zero.[105] For these group of entrants *Option Value of Delay*(q, z) equals to zero for any z. Now, consider the rest of potential entrants that hold signals $q > \underline{q}$. These are the potential entrants that decide to enter at least for one of the aggregate demand level z. *Option Value of Delay*(q, z) is weighted sum of a potential entrant's entry value function $V^e(q, z)$. Hence, for the group of entrants *Option Value of Delay*(q, z) is greater than zero. Thus, *Option Value of Delay*(q, z) is non-negative.

(b) Entry value function $V^e(q, z)$ weakly increases with q, since expected value of entry strictly increases with the signal q (See Proposition 3.1.1(a)). As a result, *Option Value of Delay*(q, z) which equals to $\beta E\left(V^e(z', q)|z\right)$ weakly increases with the signal q.

(c) Expected value of aggregate demand level tomorrow z' increases with the aggregate demand level z today. Thus, higher the z today higher the expected aggregate demand level z' tomorrow. Besides, entry value function increases with the aggregate demand level, which comes out from Proposition 3.1.1(z). As a result, *Option Value of Delay*(q, z) which equals to $\beta E\left(V^e(z', q)|z\right)$ weakly increases with the aggregate demand level. $\square$

Figure 7 displays option value of delay across the signal q and for different aggregate state z. The figure illustrates above described features of the option value of delay.

[105]The signals satisfy the following inequality $\int\limits_{s'} V^I(b_0, s', z) dH_e(s'|q) - c_e < 0$ for all z.

C Data appendix

C.1 Aggregate data

To measure aggregate activity I use Real GDP and aggregate Employment.

Quarterly Real GDP data comes from the Federal Reserve Economic Data (FRED) and covers period from II quarter of 1976 to I quarter of 2015. Real Gross Domestic Product, Billions of Chained 2009 Dollars, Quarterly, Seasonally Adjusted Annual Rate.

Monthly data of *Civilian Employment* comes from the FRED and covers period from March of 1976 to February of 2015. (Civilian non institutional employment 16 over, yearly monthly, seasonally adjusted.)

I construct yearly aggregate data to match contemporaneous level of establishment activity. Since establishment level activity is given from March to March I define current year of the aggregate variable from second quarter of the previous year till the first quarter of current year. For example, since establishment level data on year 1977 gives establishment level activity from March 1976 to March 1977 the annual contemporaneous level Real GDP will be average real GDP from the second quarter of 1976 to first quarter of 1977. The annual contemporaneous level of employment will be average employment from March 1976 to March 1977.

C.2 Establishment-level data

Establishment level data comes from the publicly available Business Dynamic Statistics (BDS) dataset. The dataset covers U.S. economy wide active establishments over the 1977-2015 period.[106] Establishment is defined as a single physical location. Establishment activity is defined by the existence of March 12 employment. At year t the data describes establishment level activity from the second quarter of the year $t-1$ through the first quarter of the year t.

In the project I use the following data series:

[106] $https://www.census.gov/ces/dataproducts/bds/data\_estab.html$

1. Economy Wide Establishment Data: Gives information about Total number of establishments, Entry rate, Total non-farm employment over the 1977-2015 period.

2. Establishment by Age data: Number of establishment by age, Number of employment by establishment age. Establishment age is computed by taking the difference between the current year of operation and the birth year. Establishment Birth (age 0) is defined as the year when the establishment first reports positive employment in the Longitudinal Business Database (LBD).

Year t cohort is defined as the group of establishments who entered in year t. The data follows each cohort up to 5 years. After 5 years the dataset gives information in 5 year bins. More specifically, The data set characterizes cohorts within the following age groups $[0, 1, 2, 3, 4, 5, 6\text{-}10, 11\text{-}15, 16\text{-}20, 21\text{-}25, 26+]$.

Corrections: BDS database is assembled using various datasets. In every five years (on years ending 2 and 7) BDS is updated using the information from the Economic Census, which gives much more detailed information about the universe of employer establishments in the U.S. According to Jarmin and Miranda (2002) the update produces a 5-year cycle in the BDS data. To address the issue I creat dummy variable for the years ending with 2 and 7 to check the 5-year cycle trend for each establishment level data series. The trend was significant only for the number of entrant establishments (establishments atage 0) with p-value respectively 0.028. The trend coefficient for all other series were insignificant. As a result, I take out the 5-year cycle trend only from the number the entrant establishments.

The other problems connected to BDS dataset such as inaction periods (which might overestimate birth and death rates) is addressed during the construction of the Database. For more information see Jarmin and Miranda (2002).

D Calibration appendix

D.1 Construction of empirical moments

D.1.1 Moments at entry

The moments reported in Table 3 are constructed as follows:

1. Average size of active establishments. Defined as the arithmetic mean of the average establishment size over the 1977-2014 period. Average size of establishment at time t is defined as:

$$Average\ Size_t = \frac{Total\ Employment_t}{Total\ number\ of\ establishments_t}.$$

2. Average Relative Size of Entrants. Defined as the ratio of average size of entrants to average size of incumbent establishments. Time (t) varies from 1977 to 2014 year.

$$Average\ Entrant\ Size\ in\ t = \frac{Total\ Employment\ at\ Entry\ in\ t}{Number\ of\ Entrant\ Establishments\ in\ t}.$$

3. Entrant Employment Share in Total Employment. Defined as the arithmetic mean of the average entrant establishment employment share in total employment over the 1977-2014 period. Entrant Establishment Share in Total Employment at time t is defined as:

$$Share_t = \frac{Employment\ at\ Age\ 0_t}{Total\ Employment_t}.$$

4. The survival rate until age 5.

Survival rate at age $a = 5$ for each cohort born at t is defined as $S_{t,5} = \dfrac{N_{t,5}}{N_{t,0}}$, where $N_{t,5}$ is the number of establishments from cohort t in age 5 and $N_{t,0}$ is the number of entrant establishment in cohort t at entry (age 0).

To find out mean survival rate up to five years of operation I take simple average of the survival rate in each age group:

$$\overline{S}_5 = \frac{\sum\limits_{t=1977}^{2014-5} S_{t,5}}{2014-5-1977} \qquad a = 5$$

5. Average survival rate from ages 21 to 26.

BDS data reports cumulative number of establishments with age between 21 to 26. The average survival rate of the cohort born between [t-25; t-21] period is given according to the following formula.

$$Survival\ Rate_{[t-25;t-21]} = \frac{\sum_{i=t-25}^{t-21} Number\ entry\ establishment_i}{Number\ of\ Establishment_t}.$$

Average survival rate from 21 to 26 is calculated by taking arithmetic mean of the average survival rate of all cohorts who lived up to 21-25 years.

6. Average entry rate of establishments. Defined as the arithmetic mean of the Entry Rate reported in the BDS dataset over the period 1991-2006.

D.1.2 Survival rate over time

To construct survival rate I use economy-wide establishment level data from the BDS database over the 1977-2015 period.[107] For each cohort born at year t survival rate up to g years of operation is defined as $S_{t,g} = \dfrac{N_{t,g}}{N_{t,0}}$ where $g = 0, 1, 2, 3, 4, 5$. $N_{t,g}$ is number of establishments from cohort t in age g and $N_{t,0}$ is the number of entrant establishments in cohort t at birth (age 0).

To find out mean survival rate up to five years of operation I take sample average of the survival rates in each age group:

[107] The data is described in details in Appendix C.2.

$$\overline{S}_g = \frac{\sum_{t=1977}^{2014-g} S_{t,g}}{2014 - g - 1977} \qquad g = 0, 1, 2, 3, 4, 5$$

Above 5 years of operation data gives cumulative number of establishments in the following age intervals $a \in$ [6-10, 11-15, 16-20, 21-25, 26+]. I use the same formula to find survival rate for the cohorts within the age intervals.

D.1.3 Average size over time

Average size at age s for each cohort born at t is defined as total employment of cohort t in time $t+s$ over number of entrant establishments of cohort t at time $t+s$. To find out average size I take average of the size in each age group.

D.2 Possible calibration strategy for τ.

One might argue that the model without persistent signal undermines the volatility of the employment. In particular, standard deviation of the aggregate employment in the data is 0.015, while in the model it equals to 0.013. In the section, I show that manipulating with the persistence of the signal τ can allow the model to account observed volatility of the aggregate employment.

τ that is set to 1 in the baseline model, determines the magnitude of the option value of delay. Hence, it controls the magnitude of the indirect effect. Decreasing τ decreases option value of delay and decreases the response of a potential entrants over the cycles. As a result, using the τ and aggregate demand shocks I can generate the variance of the employment and output by keeping the previously targeted entry rate consistent to the data.

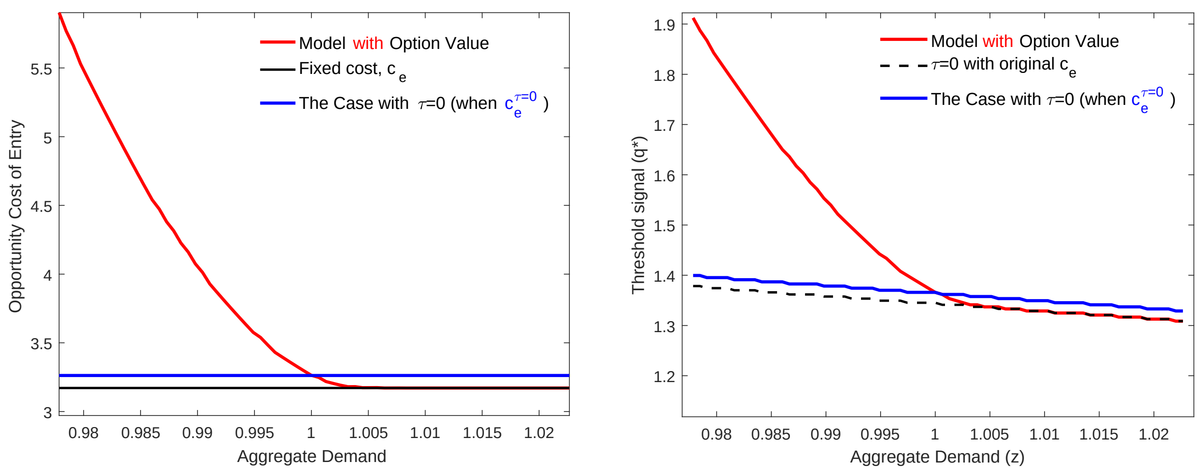
I use τ, ρ_z, σ_z and re-calibrate the model to generate match to the persistence and standard deviation of the entry rate and standard deviation of aggregate employment. Column (e) of table 5 summarizes parameter values that accomplish the goal. I find that persistence of the signal that accomplishes the goal is $\tau = 0.965$, close to the $\tau = 1$ that was considered throughout the paper and the variance of the aggregate demand shock process increased by 1.7 times.

E Different versions of the model

The section describes parametrization of different versions of the model, that are used to compare the results in the main text.

E.1 The case with $\tau = 0$ and lower number of entrants

Figure 44: The Case with $\tau = 0$



(a) Eq. Opportunity Cost of Entry (b) Threshold Signal

Change in aggregate demand level affects threshold signal directly through changing market profitability and indirectly through changing relative profitability of the market today. Option value of delay represents the indirect effect of the aggregate state. To quantify the effect of the option value of delay I set the persistence of the signal to zero $\tau = 0$, which shuts down the indirect effect of the aggregate state on the selection of entrants.

The fourth and the fifth column of table 2 contrasts parameter values used in the benchmark model and in the case with $\tau = 0$. Adjustment of the fixed entry cost ensures that the threshold quality of signal, hence the number and composition of entrant firms coincides in the stochastic steady state across cases. The difference between performance of the two cases come when I go beyond the stochastic steady state. The variation of the option value of delay generates different threshold signals, hence the different number and composition of potential entrants. Figure 45(a) summarizes the difference in equilibrium opportunity cost of entry across different aggregate states for these two cases. Note that with $\tau = 0$ opportunity cost of entry is constant over the cycles and equals the fixed entry cost. Selection mechanism when $\tau = 0$ is similar to the standard-firm dynamics model with fixed entry cost.

The difference between the benchmark model and the case explains the amplification of the shocks through option value of delay. The reader needs to note that increased threshold signal due to option value of delay will create two effects, first there will be less number of entrants in the market and second, the composition of the entrants that enter into the market changes.

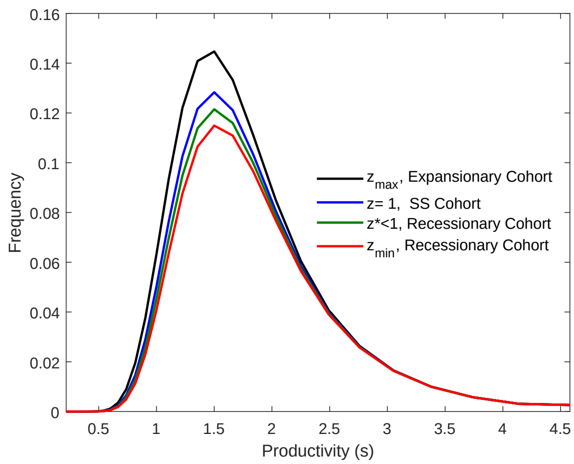
E.2 The case with $\tau = 0$ and the same number of entrants

The propagation of the entry rate dynamics comes through changing composition of the entrants by affecting the share of high productive/high growth/high survival rate firms in a cohort. The selection through the option value of delay plays an important role in composition selection of the cohorts. The following exercise emphasize the importance of the composition of the cohort and minor importance of the number in propagation of the shocks through entry rate.

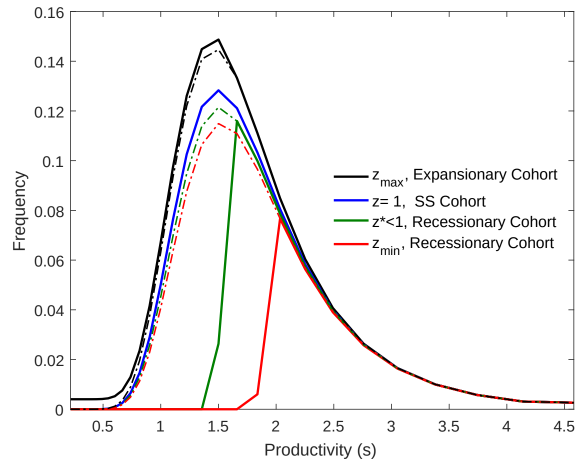
Consider the Case with $\tau = 0$. The standard deviation of the number of entrants is as low as 0.009 while the same statistics in the data is almost 7 times higher. Augmenting the model with option value of delay (setting $\tau = 1$) amplifies the effect of the aggregate shocks on the selection and as a result the volatility of the number of entrants over the business cycle increases up to 0.073.

Now, consider the following two alternative ways to generate the same observed volatility of number of entrants from the Case $\tau = 0$. In both of the scenario I take the distribution of the entrants over the idiosyncratic productivity across the aggregate sates generated in the Case $\tau = 0$ and systematically adjust number of entrants to generates the same number as in the benchmark model for each aggregate state. In particular, for the $z < 1$ aggregate states, number of entrants in the Case $\tau = 0$ is higher than in the benchmark model. To alter the number, for scenario 1 I cut the lower part of the distribution as shown in figure figure 46(b) and for scenario 2 I cut the highest productive firms as in figure 46(c). When $z > 1$ number of entrants in the Case $\tau = 0$ is lower compared to the benchmark model. In both of the scenarios I equally increase the number over the first 20 grid points of productivity.

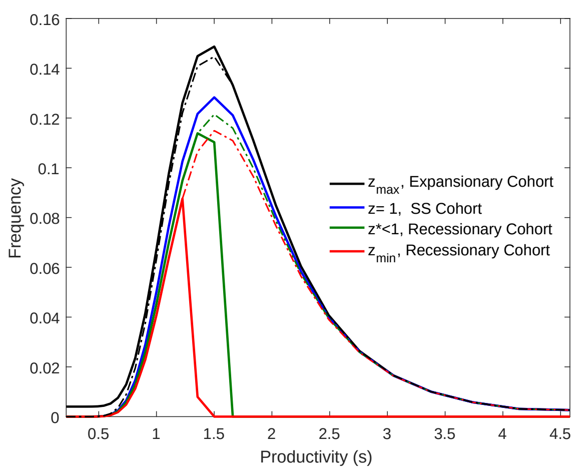
Figure 45: Selection of entrants over productivity



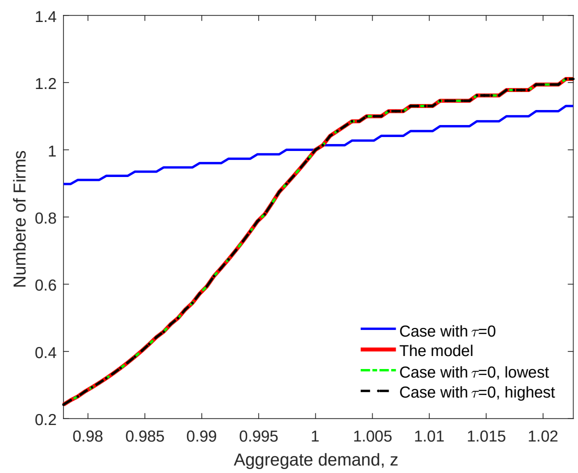
(a) The case with $\tau = 0$



(b) The case with $\tau = 0$, lowest



(c) The case with $\tau = 0$, highest



(d) Number of Entrants

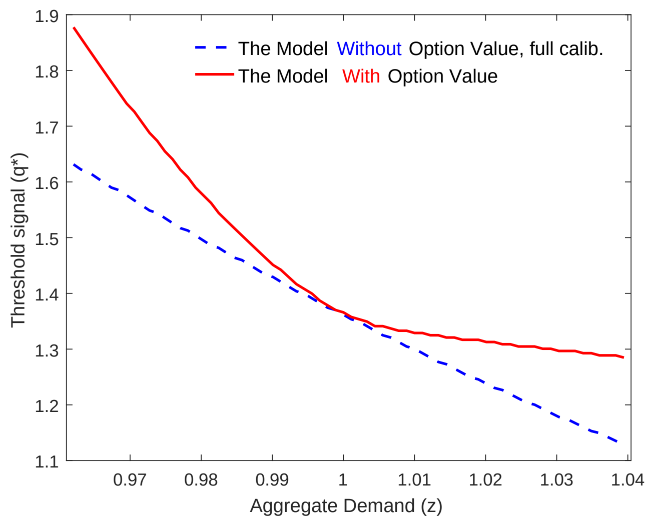
This modification ensures that these two scenarios generates same number of entrants across the different aggregate states as in the benchmark model, as shown in Figure 46(d)).

The difference between the counterfactual and the full model comes from the composition of the entrants that did not enter into the market. Examining the difference between this two case as a response of the shocks will roughly emphasize the importance of the composition of the firms who did not enter into the market.

tex

E.3 The model without the option value of delay

Figure 46: Threshold Signal



The section describes the parametrization of the Case with $\tau = 0$ and the model without option value of delay.

Note that in the stochastic steady state threshold signal uniquely determines distribution of entrants over initial productivity and customer capital, $\Omega^E_\tau(s, b_0)$. The mass of entrants with productivity s, varies only with the threshold signal and is determined based on the following function $\Omega^E(s, b_0) = \int_{q^*_\tau(z=1)}^{\overline{q}} H(s|q)dW(q)$. Once the entrant cohort is determined, it determines uniquely the invariant firm distribution in the stochastic steady state.[108]

Now, consider the model for different persistence of the signal $\tau \neq \tau'$ (in our case $\tau' = 0$. Equalizing threshold signal for these two different cases $q^*_\tau(z = 1) = q^*_{\tau'}(z = 1)$, leaving all other parameters unchanged, implies same distribution of entrants across productivity and customer capital, $\Omega^E_\tau(s, b_0) = \Omega^E_{\tau'}(s, b_0)$. Which also implies that the two cases generate same invariant distribution of firms.

Threshold signal depends on the total opportunity cost of entry, fixed entry cost plus option value of delay $c_{e,\tau} + \beta E[V^e(q^*_\tau(z), z')|z = 1]$. The value of the option is endogenously determined within the model. While fixed cost of entry $c_{e,\tau}$ can be modified to generate

[108]If the distribution of entrants are same, total distribution of firms in the stochastic steady state is generated by the many generation of same entrant firms.

desired level of the threshold signal.[109] Figure 46 summarizes the difference in fixed entry cost across models. As a result, in the Model Without Option Value setting fixed cost of entry $c_{e,\tau=0}$ to the total opportunity cost of entry from the Model With Option Value, $c_{e,\tau>0} + \beta E[V^e(q^*_{\tau>0}(z), z')|z = 1]$ implies same threshold signal in the stochastic steady state. Equalizing the threshold signal implies equalizing invariant distribution of firms and leaving all other parameters unchanged (parameters that drive the characteristics of the incumbent and entrant firms over time) implies that the two models generates same statistics in the stochastic steady state.

After calibrating the Model Without Option Value in the stochastic steady state, I find the process for the aggregate demand shock that matches the model simulated entry rate to the data counterpart.[110] The results are given in the fourth column of the table Table 7. The final values for the aggregate process are $\rho_{z,\tau=0} = 0.57$ and $\sigma_{z,\tau=0} = 0.015$.[111]

[109]Fixed cost of entry does not have an effect on the post-entry characteristics of the entrants.

[110]Since τ is not a parameter in the Model Without Option Value, I are not able to also match the model simulated variance of the employment to the data counterpart.

[111]Changing the process for the aggregate demand shock, specifically changing ρ_z and σ_z, has no effect on the statistics in the stochastic steady state. Since as discussed above the statistics is uniquely determined by the choice of the threshold signal, which in the Model Without Option Value of delay does not changes unless the fixed entry cost changes.The latter argument means that I can choose the process for the aggregate demand shock without changing the match to of the model in the the stochastic steady state. However, different $\tau > 0$ model would have been effected by the choice of the exogenous aggregate shock through changing the option value of delay. Higher variance of the aggregate demand shock increases option value of delay, while higher persistent decreases it.