The Effect of Technological Innovations on Economic Activity

by

Mykhaylo Oystrakh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Economics

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Thesis Abstract

In this PHD dissertation, the nature of technological shocks and their effect on economic activity are examined. The first chapter is dedicated to the analysis of general purpose technologies (GPTs) and their identification in the patent data. I argue that the previous literature has been identifying sub-technologies of a given GPT rather than the technology itself. Moreover, I argue that the quantity of active GPTs identified using the old methodology is substantially greater than theoretically possible. The first chapter of my thesis presents an alternative approach to the identification of a general purpose technology than the one used in the previous literature and provides an example of such a technology identified in the patent data. This technology is the microcomputer. The chapter examines its evolution, diffusion, and the effect it has on other patented technologies. The findings are in line with the theoretical GPT literature.

In the second chapter of my dissertation, I examine the effect of a positive technology shock on aggregate hours worked. Compared to the previous literature, the novelty of the approach proposed in the current study comes from two directions: the choice of variables and the technique used to identify the technological shock. A patent-based measure is the main measure used to approximate the unobserved aggregate technological process. The second important novelty of the study is the use of the sign restriction Vector Autoregression (VAR) shock identification technique that is believed to be more robust than the alternative identification techniques used in the literature. The sign restrictions are determined using a general equilibrium model with skilled and unskilled labour featuring skill-specific and general technology shocks. The analysis shows that aggregate hours increase following both kinds of technological shocks. The results obtained in the study are robust to the technology measure used. When a patent-based measure is replaced with a production function residual measure the effect of the shock still improves the aggregate hours. Moreover, the results are robust to the way aggregate hours themselves are specified. Finally the results are supported when more conventional long run restrictions are applied on the VAR. However, with the long run restrictions it does matter how the aggregate hours are specified the same way it mattered in the previous studies.

In the third chapter, I continue the analysis of the effect of a technological improvement on the labour market. In this study, I attempt to address the general criticism of the VAR methodology about the fact that only a limited number of variables can be processed. This limitation requires a researcher to make a choice in favour of certain variables and to justify this choice. Moreover, no matter which variables are chosen for the final modeling form, some information would still be excluded from the study. I overcome this limitation by incorporating factor analysis techniques into the VAR framework. I introduce Factor-to-factor VAR (F-FAVAR) that is an extension of the FAVAR approach that has already been succesfully used in the past.The F-FAVAR methodology allows an inclusion of a latent factor "impulse variable" besides the latent factor "response variables". As a result a large number of macroeconomic variables is examined and there is no necessity to exclude any of the variables. Moreover, there is no necessity to make a choice in favour of a particular measure of technology. All the relevant technological measures can be included into the model. As a result it is possible to examine the reaction of various economic and business variables to a technological shock. The reaction of the key economic variables to a technology shock is in accordance with the theory. The reaction of various labour measures to the shock was also examined. The results of the third chapter mainly support the findings of the second chapter about the positive effect of a technology shock on aggregate hours. In order to check the robustness of the results, instead of the F-FAVAR methodology, a simple FAVAR methodology was also used. For that methodology it was necessary to select a particular measure of technology to be included into the VAR model as well as to impose restrictions on the VAR. Several different technology measures with two alternative sets of restrictions were used. The results of the robustness analysis mainly support the findings of the F-Favar methodology.

KEYWORDS: Technology shocks,Vector Autoregression/ VAR, General Purpose Technology/GPT, Patents, FAVAR, Factor-to-Factor VAR/F-FAVAR, sign restrictions, tree structures, evolutionary path, aggregate technological process, impulse-response function/IRF, aggregate hours worked, skilled and unskilled labour

Acknowledgment

I would like to sincerely thank my thesis supervisor, Professor J. P. Lam for all the support, guidance, and help I received during my PhD career at the University of Waterloo. Whenever I found myself stuck with a problem, I knew that with Professor Lam's help a solution would be found. It was an enjoyable journey to work under his supervision. I am really thankful for the Macroeconomics reading class Professor Lam designed for me. Very advanced topics did not seem that hard after he explained them.

I would like to thank my PhD committee members Professor Lutz Busch and Professor Mathew Doyle. Their advisory help and general support are really appreciated. Classes they taught and further conversations during my PhD career helped form me into a researcher and a critical thinker.

I would like to thank Professor Anindya Sen for all the support I received from him when I was a student, but especially I would like to thank him for helping me radically change my life by recommending me for my current job at the Bank of the West in San Francisco, California.

I really appreciate all the support and help I received from Pat Shaw from the first day I arrived at the Economics department up until this day. I knew I could count on Pat in case of any problem or difficulty.

I am thankful to Professor Margaret Insley for believing in me and helping me receive my very first teaching assignment and further allowing me to run my classes to the best of my judgment.

I really appreciate the coding help I received from Guillermo Fuentes while working on the first chapter of my dissertation.

I would like to thank my undergraduate Professors from York University Glendon College Vincent Hildebrand and Ian McDonald. Professor Hildebrand is the person who introduced me to economic research and empirical analysis and who gave me my first research assistant assignment. Due to Professor McDonald's brilliant lectures and interesting office hours discussions I became interested in economic theory and decided to pursue this career path.

Last, but definitely not least, I would like to thank my wife, Alexsandra Oystrakh, for being patient, and supporting me during my PhD career, for believing in me, and also for proof-reading this thesis multiple times.

TABLE OF CONTENTS

Thesis Introduction

0.1 General Introduction

Technological progress is one of the major driving forces of an economy. The role of new technologies in economic development since the first industrial revolution is very significant. According to [Mokyr, 1990] out of three possible engines of economic growth: market expansion, capital accumulation, and technological innovation, only the latter is not subject to diminishing returns. Economic growth over the last two centuries has mainly been the product of technological innovations.

Technology in some form, either endogenous or exogenous, is present in many modern macroeconomic models of growth and business cycles. In such models, especially in those explaining business cycles, technology typically enters the production function in a form of a shock. This shock is a sudden change that improves factors' productivity. Understanding the nature of technology, how it evolves and how its evolution affects the economy is the key to our comprehension of how technology shocks influence economic activity in the short and long run. Although economists agree that technological shocks are very important for explaining economic fluctuations in the short run and long run, these shocks are difficult to identify. There is a large amount of literature on understanding the nature of technology shocks and the effects of technology shocks on aggregate economic activity. Despite this literature, economists disagree on the effects of technology on the economy, especially in the short run, and on how to identify these shocks.

The two most fundamental questions that apply to technological shocks are: "What is the nature of a technological shock?" and "What are the effects of a technological shock on economic variables such as productivity, output, or employment?" The goal of this dissertation is to increase our understanding of these questions.

0.2 Chapter 1: Evidence of General Purpose Technologies in the Patent Citation Data

In the first chapter of my dissertation, a permanent improvement of the aggregate technological process caused by a specific kind of technology called a "general purpose technology (GPT)" is identified using US patent data. The chapter includes a discussion of how to identify a GPT using the patent data. I argue that the previous literature has been identifying sub-technologies of a given GPT rather than the technology itself. Moreover, I argue that the quantity of active GPTs identified using the old methodology is substantially greater than theoretically possible. The first chapter of my thesis presents an alternative approach to the identification of a General Purpose Technology than the one used in the previous literature. It also provides an example of such a technology identified in the patent data.

The example of a general purpose technology examined in the first chapter is a microcomputer technology. In the previous studies that attempted to identify general purpose technologies in the patent data, the following approach was taken: certain patents' characteristics were examined. These studies then proceeded by trying to identify some outlying characteristics of a given patent and arguing that a certain combination of such outlying characteristics makes given patented technology a general purpose technology. The approach taken in the first chapter of my thesis is the opposite. Instead of looking at the patents first and trying to find the patents that look like general purpose technology patents, I look at the technology first and then allocate the patents associated with this technology. A technology that is commonly accepted as a general purpose technology is selected first. The main selection criterion is that the active development phase of the technology must coincide with the available time range of the patent citation data. Only after that the initial patents associated with the technology are allocated and examined.

In order to observe the subsequent evolution and diffusion of a general purpose technology starting with the identified initial patents, a structure linking patents by their citations called an evolutionary path is introduced. An evolutionary path is obtained by transforming the initial dataset into a tree looking structure with the initial patents at the root and assigning each patent within the evolutionary path a position level that identifies the largest quantity of citation links necessary to connect the patent with the root. The evolutionary paths can be constructed for any kind of technology, not just for a general purpose technology.

After obtaining the evolutionary path for the microcomputer technology we then analyze its evolution and diffusion across the technological fields. The following exercises were conducted over the microprocessor evolutionary path. First, it was noted that the patents contained in the evolutionary path are crucial in order to observe the patent growth spurt of the 1980s - early 1990s. If we subtract the patents connected by citation links to the initial microcomputer patents from the population of patents granted in the United States between 1976 and 2006, we do not observe the patent growth spurt in the aforementioned time period. One of the explanation of the patent growth spurt observed in the given time period is the presence of a fertile technology.[1]

The analysis of the microcomputer evolutionary path confirms that the computer technology boom experienced over the time period is at least partially responsible for the aforementioned increase in the patenting activity. Moreover, the technological history literature recorded similar and even stronger outbreaks in patenting activity.[2] Such outbreaks could be associated with the development of other important general purpose technologies such as electricity and internal combustion engines.

In order to link my work to the previous studies dedicated to GPT identification in the patent data, I conduct the following analytical exercise. I examine whether the patents with certain outlying characteristics have any relationships with active GPTs and thus acquired certain GPT features. Recall that the past studies would search for outliers in the patent data in order to identify GPTs. This analytical analysis is conducted the following way. First, a sample of one thousand patents was extracted from the population of patents granted in the United States in 1976 (the beginning of the dataset). Each of those patents was assigned to be an initial root patent from which an evolutionary path was constructed. As a result a sample of a thousand evolutionary paths was derived. Each of those paths was linked to the microcomputer evolutionary path by the number of coincident observations on each position level. A coincident observation is a patent that belongs to both an examined evolutionary path and the microcomputer evolutionary path. Subsequent regression analysis revealed that a closer connection to the general purpose technology evolutionary path makes any other evolutionary path bigger in terms of the number of patents it contains.

This observation confirms that a fertile general purpose technology is beneficial for other technologies and stimulates aggregate technological knowledge in general. Moreover, a diffusion of technologies across technological fields in an evolutionary path also benefits from a closer relation to the general purpose technology. Evolutionary paths that have more coincident observations with the microcomputer technology were found to have greater numbers of various technological fields and sub-fields covered.

Another set of regression equations was presented in order to analyze how a relation with the general purpose technology affects certain characteristics of an initial basement or root patent from the sample of one thousand patents. It was found that certain patent characteristics that were used by the previous literature to attribute a patent to a GPT actually depended on how close the given patent is to a GPT. In the general purpose technology theory dissertation appendix essay[3], I argue that general purpose technologies are too complex to be represented by a single patent. However a closer relation of a patent to an actively developed GPT is one of the key factors that make this patent look like an outlier by certain characteristics.

[1]See [Kortum & Lerner, 1999].

[2]See [Marco et al., 2015].

[3]In order to be able to identify general purpose technologies I conducted a detailed literature review on the nature and features of historically observed general purpose technologies. The results of this research instead of being included into the first chapter of my thesis are combined in a standalone essay located in the Appendix of the dissertation.

Therefore, many of the patents identified by the previous methodologies as GPT candidates are indeed related to GPT but each one of them does not represent a separate GPT.

As a result, in the first chapter I introduce a way of identifying, describing and, analyzing general purpose technologies using patent citation data. I argue that the previous literature has been identifying sub-technologies of a given general purpose technology rather than the technology itself. On the other hand, the approach introduced in the first chapter allows us to capture all the patented effects of a given general purpose technology and analyze it at a full scope. The identified patented microcomputer technology evolves and diffuses in agreement with the GPT literature.

0.3 Chapter 2: Effect of Technological Innovations on Hours Worked

In the second chapter of my dissertation, I examine the effect of a positive technology shock on aggregate hours worked. Neither theoretical not empirical studies give a determinate answer about the effect of technology on aggregate hours in the short run. On the theoretical side the debate is whether the income or the substitution effect dominates following a technology shock that improves aggregate factors' productivity. According to real business cycles models that feature flexible prices and wages, an improvement of productivity will cause an increase in aggregate hours due to the fact labour becomes more productive. According to the Neo-Keynesian literature, the effect is the opposite. In models with sticky prices, a positive technology shock causes a decrease in aggregate hours worked.

Empirical studies do not resolve the controversy. Empirical literature examining the effect of productivity on hours worked using time-series econometric analysis came to contradictory conclusions about the effect. The conclusions crucially depend on the method used to identify the technology shock and on the specification used.

Compared to the previous literature, the novelty of the approach proposed in the current study comes from two directions: the choice of variables and the technique used to identify the technological shock. I make two contributions to the literature. First, I choose a measure of technology shocks that is different from the one used in the literature. I use my results from chapter 1 to guide me in this task. A patent-based measure is the main measure used to approximate the unobserved aggregate technological process. In the comparative study based on literary sources I investigate the advantages and short-coming of this measure. The main advantage is that unlike a production function residual, the patent-based measure is a direct measure of innovations. Moreover, unlike the R&D expenditure accounts, every patent implies at least a minimum improvement of the pre-existing technological state. Nevertheless, I check the robustness of my results using a secondary measure of technology, the capacity utilization adjusted total factor productivity described in [Fernald, 2012].

The second novelty of the study is the use of the sign restriction technology shock identification technique that is believed to be more robust than the alternative identification techniques used in the literature.[4] Zero short and long run restriction methodologies used in the earlier studies require a researcher to have strong beliefs about effects that innovations to certain variables have on other variables. For example, a researcher may have to assume that innovations to certain variables have no immediate or no long-term effect on other variables. Sign restrictions are less demanding in terms of the assumptions. Instead of a no-effect assumption, an assumption about the sign of the effect is imposed.

In order to estimate the effect of a technology shock on aggregate hours worked a five-variate Vector Autoregressive (VAR) model with sign restrictions is used. Besides the aforementioned technology measure and the measure of the aggregate hours worked itself another three macroeconomic variables are included in order to identify the shock. Given that the sign of the response of aggregate hours to the

[4] See [Peersman & Straub, 2004] for the discussion of robustness of sign restriction identification technique.

shock is the main subject of the analysis, no sign restrictions are imposed on the aggregate hours worked measure itself. In order to uniquely identify a technology shock and to distinguish it from other common macroeconomic shocks, sign restrictions are imposed on the measure of technology as well as on the three other macroeconomic variables whose reaction to a technology shock is not subject of the investigation.

A model with skilled and unskilled labour is used to determine the sign restrictions that are imposed on the data. The model features two types of technology shocks: skill-specific shock that improves productivity of skilled workers only and skill-neutral shock that improves productivity of all workers in the economy. The model is constructed so that a technology shock can produce an increase or a decrease in aggregate hours worked. This flexibility allows us to examine the way aggregate hours react to a technology shock based purely on the data. The model allows for the aggregate hours to increase when the technology shock is of a general nature and to increase or decrease after a skill-specific shock depending on the model calibration and on the assumption regarding the ratio of skilled to unskilled workers.

The model, while providing flexibility about the reaction of aggregate hours to the technology provides a way of distinguishing a technology shock from other common economic shocks such as demand shock, or monetary shock. Therefore, when the sign restrictions are imposed on the data it is possible to argue that the shock being examined is a technology shock and not some other shock that could hypothetically affect aggregate hours worked. The model also allows us to distinguish a general technology shock from a skill-specific technology shock by imposing alternative sign restrictions on the variable that represents the difference between skilled and unskilled workers employment in the VAR.

From the empirical distributions of the impulse responses constructed in the study, it is apparent that aggregate hours increase following both kinds of technological shocks. Based on the model intuition it appears that the amount of skilled workers in the current economy is big enough to offset any negative effect of technology on unskilled labor and to drive up aggregate hours worked. The results received in the study are robust to the technology measure used. When a patent-based measure is replaced with a production function residual measure the effect of the shock on the aggregate hours is still positive. Moreover, the results are robust to the way aggregate hours themselves are specified. According to the previous studies, the way aggregate hours react to a technology shock empirically would depend on whether the aggregate hours are specified in levels or in first differences. In the current study the specification of the aggregate hours does not affect the conclusion. Finally the results are supported when more conventional long run restrictions are applied on the VAR. However, with the long run restrictions it does matter how the aggregate hours are specified the same way it mattered in the previous studies.

0.4 Chapter 3: Analysis of Aggregate Technology Shock with Factor-to-Factor VAR

In the third chapter, I continue the analysis of the effect of a technological improvement on the labour market. In this study I attempt to address the general criticism of the VAR methodology on the fact that only a limited number of variables can be processed. By including only a limited number of variables into the model a researcher has to make choices on which variables to keep. These choices are associated with the following problems. First, no matter which variables are chosen in the VAR, some potentially useful and relevant information will be omitted. Second, there often exist multiple alternative measures of a particular phenomenon. By limiting the number of variables a researcher has to decide and justify which of the available measures to use. Often each of the measures has its own advantages and disadvantages and there is no universally accepted "best" measure. Therefore, the results of the research immediately become a subject of criticism.

[Bernanke et al., 2004] introduced a Factor Augmented VAR (FAVAR) methodology to examine the effects of a monetary policy shock on macroeconomic data. The model presented in that study featured a monetary policy measure along with other few latent factor variables derived from a large macroeconomic dataset. It is necessary to note that in the case of the monetary policy, the choice of the shock variable is easy and uncontroversial. For the last decades, the Federal Funds Rate has been the main monetary policy instrument in the United States. If a researcher wanted to examine the effects of a technology shock using the FAVAR methodology, the researcher would still have to argue about the technological measure to be used as it is argued in the second chapter on my dissertation.

In the third chapter of the dissertation, I extended the FAVAR methodology to allow an inclusion of a latent factor "impulse variable" besides the latent factor "response variables". The methodology is called Factor-to-factor VAR (F-FAVAR) As a result a large number of macroeconomic variables is examined and there is no necessity to exclude any of the variables. Moreover, there is no necessity to make a choice in favour of a particular measure of technology. All the relevant technological measures can be included into the model.

The methodology works the following way. First, a large number of variables that include economic, technological, and business measures are combined into one dataset. The variables are sorted in terms of the category they represent. For example all the measures of output are grouped together and located at the beginning of the dataset, next all the labour market measures are located and so on. A small group of miscellaneous variables that could not be attributed to any of the groups is located at the bottom. The number of principal factors equal to the number of variable groups is extracted from the data. The matrix of loadings is rotated so that each of the factors has its maximum loadings in each of the particular groups and the loadings for all the other groups are close to zero. It is important to mention that only orthogonal rotations are used and therefore the independence of the latent factors is preserved.[5]

The rotated factors are combined into a VAR system. A VAR model is then solved. No restrictions to the VAR system is necessary since that the orthogonality of the factors is ensured. The responses of all the factors to the impulse of the technological factor were examined.[6] The impulse-response functions of the latent variables were projected to the initial variables. As a result, it was possible to examine the reaction of various economic and business variables to a technological shock. The reaction of the key economic variables to a technology shock is in accordance with the theory. The reaction of various labour measures to the shock was also examined. The results of the third chapter mainly support the findings of the second chapter about the positive effect of a technology shock on aggregate hours.

In order to check the robustness of the results, instead of the F-FAVAR methodology, a simple FAVAR methodology that was described in [Bernanke et al., 2004] was used. For that methodology it was necessary to select a particular measure of technology to be included into the VAR model as well as to impose restrictions on the VAR. Several different technology measures with two alternative sets of restrictions were used. The results of the robustness analysis mainly support the findings of the F-Favar methodology.

In general, the second and the third chapters of the dissertation explore different methodological approaches to examine the short run effects of technological innovations on macroeconomic variables. The effects on these variables are in accordance with the previous theory whenever there is a theoretical consensus on such effects. When such consensus does not exist, as it is in the case with hours worked, the findings of the both chapters indicate overall positive effect of technological innovations on aggregate hours worked.

[5]The rotation exercise allows to achieve two objectives. First of all each of the resulting factors has a certain economic meaning. Therefore, it is possible to allocate the factor that has high loadings for the numerous technological measures included in the data and low loadings for all the other variables. This latent factor is a latent technological factor that is later used in the VAR as a technological shock variable. The second objective of the rotation exercise is that, according to the factor analysis literature (See [Bai & Ng, 2012] for example). the rotation towards a block-diagonal matrix described above is sufficient to uniquely identify the factors and the loadings.

[6]Note that it is possible to examine the responses to the other meaningful factors in the system, for example of a factor with high loadings for the monetary variables' group, or with a fiscal variables' group. However, this was beyond the scope of the current study.

Chapter 1

Evidence of General Purpose Technologies in the Patent Citation Data

1.1 Introduction

In the last twenty years more and more economists have turned their attention to General Purpose Technologies (GPTs). Unlike a single purpose technology (SPT), a GPT is applied in many sectors and can serve various purposes. When a GPT arrives, it rejuvenates the economy by making production of existing commodities more efficient, and by introducing new commodities. Society has observed many GPTs. Examples of GPTs are "Electricity", "Computers", or "the Wheel".

The study of fundamental innovations is a relatively young topic in economics. The term GPT was first introduced in [Bresnahan & Trajtenberg, 1995]. However, one can find discussions about the phenomenon of fundamental innovations earlier in technological literature, for example in [David, 1990]. GPTs are sometimes contrasted in the literature with the Single Purpose Technologies or SPTs. One can also find several other terms that have meanings similar to those of the concepts of GPT and SPT. For example, [Mokyr, 1990] distinguishes between micro and macro-inventions, while [van Zon et al., 2003] use the terms core and periphery innovations, and [Aghion & Howitt, 1998] call GPTs fundamental innovations.[1] Throughout my dissertation, I will use the term GPT to mean a technology that can drastically affect the whole economy, improve production in existing industries, and give rise to new industries.

The main objective of the first chapter is to uncover GPTs using the US Patent Citation dataset. The "NBER US Patent Citation Data" was collected by Bronwyn Hall, Adam Jaffe, and Manuel Trajtenberg and presented in [Hall et al., 2001b]. This dataset includes all the patents granted in the United States between 1976 and 2006. There are 3210179 patents in total for that period. Besides listing the patents, the dataset provides information about the relationship between them. Patents in the data include citations to the "prior art", that is to the earlier inventions the current invention is based on. Thus the patent citation data provides a good way to examine links between various inventions, how these inventions diffuse over time and how they constitute a technology shock.

The first attempt to uncover GPTs in the patent data belongs to Bronwyn Hall and Manuel Trajtenberg. Their identification technique, described in [Hall & Trajtenberg, 2004], requires finding a patent that differs from the rest of the patents in the data by certain criteria and claiming that this patent may be

[1] Here and after I use the terms GPT and fundamental innovation interchangeably. Even though the term fundamental innovation describes the phenomenon much better, I will have to use the term GPT in order to be on the same page with the rest of the literature.

a potential GPT-carrier, that is a patent that would represent a GPT in the data. Their approach provided a list of 20 GPT candidate patents. In my research, I examine these candidate patents as well as other similar patents that I identify using the [Hall & Trajtenberg, 2004] methodology.

Based on the previous and new measures that I develop, as well as in accordance with the theoretical discussion, I argue that using the patent data, one cannot firmly establish or pre-identify a GPT in its early stages. Instead, I find that a particular GPT, that is well-established and universally agreed to be a GPT, can be described by a collection of patents. Growth of such a technology and its diffusion throughout the technological sectors can be examined using the patent citation data. I provide a technique of creating a subsample from the patent citation dataset that includes patents related to a particular GPT.

In order to describe GPTs in the data I introduce a tool that I call an evolutionary path. Evolutionary paths resemble tree-structures, they start with one node - the originating or level-zero patent, the first level includes patents that cite the original patents, the second level includes the patents that cite the first level patents and so on until the end of the observations in the patent citation dataset. These evolutionary paths allow one to observe how a particular patent affects the overall patent data, and in which technological fields the applications derived from this patent are found. In other words, evolutionary paths describe the contribution of a particular technology to the overall technological knowledge described by the patent data. Evolutionary paths serve as a useful tool for examining the GPT-candidate technologies given that these technologies naturally would involve patents from a greater variety of technological fields than the regular single purpose technologies. Note that evolutionary paths can be applied to describe any kind of technology within the data, not only a GPT. An evolutionary path is based on the citation feature of the patent data. Hence, evolutionary paths can be constructed with other types of data that have reference links between from one datum to another.[2]

Based on the theoretical GPT discussion and the empirical analysis involving the construction of the evolutionary paths for the GPT-candidate patents, I establish that very general knowledge usually arises from academic and non-profitable research and therefore does not get patented at its earliest stages. Moreover, according to the GPT literature, a GPT at its early stages has a very narrow spectrum of applications. Thus, when the first applications based on the new general knowledge get patented, they are indistinguishable from the regular single purpose technological innovations.

I conclude that, first, it is impossible to identify a GPT in the data when this GPT is at the early phases of the life-cycle.[3] Second, one should not expect the whole GPT to be contained in a single patent.[4] However, one can observe the effects of the GPT in the patent data. The way to trace a GPT in the patent data is to wait until a technology reveals itself as a general purpose technology. Only after it has revealed itself as a GPT, it is possible to examine its evolution and diffusion using the data.

This approach does not allow one to predict upcoming major technological transformations, however it allows the examination of the effects of such a transformation after the transformation took place. A theoretical discovery or a set of discoveries that initiates a GPT usually does not get patented. In the patent data we can observe the earliest applications of these discoveries. These applications will serve as a starting point or points[5] for the construction of an evolutionary path. The evolutionary path would then represent all the patented applications of a given general purpose technology.[6]

[2] One particular type of data that would be interesting to examine using the evolutionary paths is the "scientific publication data".

[3] [Carlaw & Lipsey, 2006] have a discussion about the level of uncertainty associated with a GPT, when it arrives. Given that a GPT reveals all its applications and technological improvements throughout a long period of time, sometimes a century, one would not be able to identify with a 100% certainty a GPT when it just arrived, moreover, one would not be able to predict the way this GPT will evolve and diffuse at its early stages of development.

[4] For example, several types of dynamo machines got patented in the late XIX century (Siemens' patent, Zipeenowsky and Deri's patent, Brush's patent), while the earlier predecessors of these machines (Faraday's disk and Jedlik's dynamo) did not get patented at all.

[5] It is possible for an evolutionary path to start from several nodes and therefore it is possible to describe a technology that originates from a series of patents in the data.

[6] It is necessary to note that some researchers may violate their legal responsibility to quote some relevant previous art, and thus their patents would not to be included in an evolutionary path. On the other hand, some of the researchers may cite an irrelevant

My approach in uncovering GPTs in the patent data is more intuitive than the "search for outliers" approach in [Hall & Trajtenberg, 2004]. My approach allows us to describe and examine a GPT, to establishe its features, and to trace its diffusion. Unlike previous research, I conclude that, given the time frame of my dataset, it is unlikely that we can find 20 developing or active[7] GPTs as [Hall & Trajtenberg, 2004] have suggested in their article. Based on previous analysis and history, at most 3 active GPTs are likely to be observed during that time frame.[8]

As an example, I describe the "Microprocessor Based Computers" general purpose technology that I call "microprocessor" for short. The choice of this technology was dictated by the time range of the data set. The latest edition of the NBER US Patent Citation Data is in the range from 1976 to 2006. I, naturally, wanted to find a GPT that exists and is "active", that is evolves and diffuses within this time range. Luckily there was such a GPT - microprocessor-based computers.

I construct the evolutionary path based on the earliest microprocessor-relevant patents. Then I examine the effects of a "microprocessor" technology on the aggregate technological process expressed using number of patents granted. The observed effect of a technology on the aggregate knowledge process resembles a positive permanent technological shock since the mean of the process is shifted to a permanently higher level.[9] Previous researchers, for example [Kortum & Lerner, 1999] and [Rafiquzzaman & Whewell, 1998], at least partially attribute this surge in the innovative and patent activity to the fertile computer technology effect rather than the effect from some particular patent policy. I also find that the acceleration of the patenting activity in 1980s should be at least partially attributed to the presence of an active and very fertile GPT that I denote "microprocessor".

The effect of the microcomputer GPT on the aggregate technological knowledge can be observed when one subtracts the patents that belong to the microcomputer evolutionary path from the total number of patents in the dataset. The upward trend segment of the process in the 1980s-1990s is not observed when we exclude the patents derived from the microprocessor technology. The process becomes mean-stationary throughout the time range of the data. Therefore, the effect of the arrival of the microprocessor GPT is the growth in the patenting activity that can be considered a positive permanent technological shock. This effect is in accordance with the theoretical GPT literature, where it is stated that GPT opens the door for the new stream of innovations that are applied in various sectors of the economy.[10]

The paper is structured the following way: Section 1.2 will present the current trends in the GPT literature concentrating on the identification of GPTs in the patent data. Section 1.3 is dedicated to the description of the NBER patent citation dataset and the data processing methodology used in this study. The authors of the dataset supplied it with the accompanying paper ([Hall et al., 2001b]) where they supply a detailed analysis of the data features and provide multiple descriptive statistics. The presence of this accompanying paper allows me not to go in detail discussing the dataset itself and spend more time on presenting the construction of the evolutionary path from the data. The first natural candidates for construction of evolutionary paths are the GPT-candidate patents selected in [Hall & Trajtenberg, 2004]. Section 1.4 provides the analysis of the single patent GPT candidates along with the discussion about why it is unlikely for those patents to represent multiple co-existing GPTs. In Section 1.5 I present an alternative approach to identifying GPTs in the patent data. Instead of looking at a patent and speculating

patent in order "to be safe" or guided by some other motivation, especially given that some of the citations are added by the patent bureau examiners.

[7]In [Oystrakh, 2014] readers can find a discussion of the phases of a typical modern GPT based on the GPT theory. The phase of a GPT life-cycle that is associated with the active development of the applications for a given GPT the "active phase of the life-cycle".

[8]According to [Lipsey et al., 2005] for the given time period there are "Computers", "Internet", and "Biotechnological" GPTs that are active.[Lipsey et al., 2005] separate "Computer" and "Internet" GPTs while one can claim that this is the same technology that allows us to process information electronically.

[9]Two types of unit root tests: Dickey-Fuller and Phillips-Perron were applied to the time series aggregate technology measure. Both tests indicated absence of the unit root in the data prior to the beginning of the active phase of the microprocessor technology and the presence of the unit root afterwards.

[10]See [Lipsey et al., 2005] for example

whether this particular patent can or cannot be a GPT based on certain characteristics of the patent, I take a well-known general purpose technology that is commonly and unarguably believed to be of general purpose (the microcomputer or microprocessor computer technology) and then trace it within the patent data. In section 1.6, I examine the microprocessor GPT identified within the patent data and observe its affect on aggregate patenting activity over time. I find that presence of this GPT stimulated research activity and gave birth to multiple applications in various sectors thus increasing patenting activity. In Section 1.7, I return to the discussion of the old GPT identifying methodology presented in [Hall & Trajtenberg, 2004] and numerous other studies. So why are certain patents outliers by some of their characteristics? I find out that a close relation of a given technology to an active GPT makes this technology different from an "average technology" of the time. By studying a sample of 1000 random patented technologies and introducing a measure of the connection for these technologies to a GPT, I find that the connection to the GPT plays an important role in the subsequent life-cycle of a given technology. Section 3.5 concludes the study.

1.2 Literature Review

The term GPT was introduced for the first time in [Bresnahan & Trajtenberg, 1995] where these technologies were described as "pervasive technologies with inherent potential for technical improvements and innovational complementarities that give rise to increasing returns to scale". Since [Bresnahan & Trajtenberg, 1995] various authors have studied the role GPTs have played in the economy.

The GPT literature can be roughly divided into three groups according to the types of research questions. The first type of publications describe the nature of GPTs and thus contribute to the GPT theory. The second group of researchers conduct empirical analyses in order to identify GPTs in the data. To accomplish this goal they exploit the GPT features described by the first group of publications. Finally, the third group works on constructing economic models that incorporate GPT into the economic analysis. A detailed discussion of publications from each of the groups can be found in [Oystrakh, 2014]. In this section I will list and review some past contributions to the identification of GPTs in the patent data since this branch of the GPT literature is the most relevant to the current study. However, it necessary to note that many GPT-identifying empirical studies contain a theoretical part about GPTs, therefore the distinction between the two types of articles is sometimes vague.

First it is necessary to give credit to the creators of the US patent citation dataset for the detailed analysis of the data in [Hall et al., 2001b]. In their article some important statistics as well as techniques that are useful in detecting GPTs are established. [Hall et al., 2001b] include a very detailed description of the dataset along with the characteristics of the variables and statistics that are in the data as well as some general observations about the dataset.

In [Hall et al., 2001b] among other issues related to the patent citation dataset, the authors discuss the arrangement of patents into technological categories and subcategories. They propose measures of "generality" and "originality" for the patent.[11] These are the measures indicating whether this patent is cited or cites a greater variety of technological fields relative to the other patents. Measures of generality and originality become useful in the GPT identifying analysis presented in [Hall & Trajtenberg, 2004]. [Hall et al., 2001b] also propose ways of dealing with the citation inflation[12] in the patent data.

The second paper produced by the authors of the dataset, [Hall & Trajtenberg, 2004], describes research methodologies useful in detecting GPTs. The article presents a list of 20 patents that best fit the predetermined criteria of a general purpose technology. In order to identify a GPT the authors assume

[11] See Appendix 1.9.1where these variables are re-defined and slightly improved.

[12] An increase of the average number of citations per patent in the recent decades is sometimes called "citation inflation" int he literature.

that a GPT-patent would be more general (cited by patents from a wider spectrum of technological fields) than an average patent, the followers of the GPT-candidate patents would also be more general. Moreover, the technological field of the GPT-candidate, after the introduction of that patent, is expected to grow in size (contain more patents) than an average technological field. Thus, a GPT-candidate patent would be a "black sheep" by certain criteria.

The article "Was Electricity a General Purpose Technology? Evidence from Historical Patent Citations" by [Moser & Nicholas, 2004] is also based on patent citation data. The authors use two sets of the data, the first set includes patents from the 1920s. The second set used in [Moser & Nicholas, 2004] contains modern patents i. e. the patents granted after the year 1976. However, the authors use only those modern patents that directly cite the patents from their 1920s patent set. [Moser & Nicholas, 2004], like [Hall & Trajtenberg, 2004] also apply the measures of generality and originality[13] as well as other patent describing variables to compare the patents from the "Electrical" technological fields to the patents in the other fields. The goal of their analysis is to establish whether the technological breakthroughs of the early XX century are all due to the "Electrical" GPT.

Nanotechnology is often considered to be a potential GPT in the literature. For example, [Lipsey et al., 2005] and [Nikulainen & Kulvik, 2009] list it as a potential GPT of the 21st century, while [Menz & Ott, 2011] take the fact that nanotechnology is a GPT for granted in their study of its effect on regional development.

[Youtie et al., 2008] attempt to identify whether nanotechnology is an emerging GPT using the patent citation data and the analytical tools presented in [Hall et al., 2001b]. The approach taken by the authors is simple. They assume that Information-Communication Technology (ICT) patents are definitely patents representing a GPT, while drug and medical patents are definitely not representing a GPT. Based on this assumption they compare the nanotechnological patents to both computer and drug patents of the same time cohort. The criterion for the comparative analysis is the generality index presented in [Hall et al., 2001b] and [Hall & Trajtenberg, 2004]. The authors find that the nanotechnological patents are closer to the computer patents according to their generality measure and therefore it looks like nanotechnology is indeed an emerging GPT.

[Nikulainen & Kulvik, 2009] examine three technologies: ICT as the established GPT, and also nano and biotechnologies as potential GPTs. Among other sources, the authors use Finnish patent data to examine the links between these technologies and the industries. The authors look at the companies assigning the patents in the technological fields of interest, if, for example, a company listed in the forest industry filed an ICT patent, the authors would observe the link between the forest industry and the ICT technology that indicates the diffusion of the technology into that industry. Applying this methodology to the patent data from various sources, the authors establish that ICT and nanotechnology diffuse to much larger extent and find a much broader set of applications than biotechnology did.

Another article that uses the US patent Citation Dataset is [Gress, 2010]. The article includes a discussion of backward and forward citations in the dataset and provides a good description of the analytical tools that will help in drawing conclusions about generality, originality of a technology. The author also provides an illustrative discussion about how patents are related to each other using graph theory. Unfortunately, the graph theory analysis presented in the article does not allow the examination of some particular technology, rather it shows a segment extracted from the data with no beginning and no ending. These chunks of data are used for illustrative purposes, to show how the data looks in general, what are the main features and general observations about the structure of the US patent citation dataset.

It is also necessary to mention the discussion paper "Knowledge Flow and Sequential Innovation: Implications for Technology Diffusion, R&D and Market Value" by [Belenzon, 2006] in which the author constructs tree-looking structures, similar[14] to the "evolutionary paths" proposed in my research, using

[13]The authors in [Moser & Nicholas, 2004] define their originality measure in a slightly different way from the way used in [Hall et al., 2001b] and [Hall & Trajtenberg, 2004]

[14]Unfortunately I was not able to establish the exact way the tree structures are constructed in [Belenzon, 2006] in order to be

patent citation data. The research question in[Belenzon, 2006], however, differs from the question in this study. [Belenzon, 2006] uses his structures to trace whether the original inventor participated in the subsequent rounds of research.

1.3 Data and Methodology

1.3.1 Data

In this section, I briefly describe the dataset. For more details and descriptive statistics a reader should review [Hall et al., 2001b]. I also present a way of processing data into tree-structures that I called evolutionary paths in order to observe evolution and diffusion of a technology. I use the 2006 vintage of the US patent citation data. The dataset spans from 1976 to 2006. It contains various information about all patents issued in the US during that time period. This information includes patent characterizing variables, information about patent assignees, and patent citation links.

The following two tables from the dataset are used: the general patent information table that includes the list of all the patents in the dataset along with relevant variables and statistics and the citation table that includes pairwise cited-citing related patents. In order to generate the evolutionary paths these tables were combined into one table that has pairwise cited-citing related patents and the variables and statistics relevant to the citing patent. Table 1.1 schematically represents the dataset.

Table 1.1: Initial Dataset Structure

Citing patent ID	Cited patent ID	Grant year of the citing patent	Other variables describing the citing patent
2	1	1976	···
3	1	1977	···
3	2	1977	···
4	2	1978	···
5	1	1979	···
5	3	1979	···
5	4	1979	···

Description: Assume there are five patents in the data. Patent 1 is the oldest one and it is cited by patents 2, 3, and 5. Patent 2, is the second oldest and is cited by patents 3 and 4. Patents 3 and 4 are cited only by patent 5, while patent 5 is the latest and it is not cited by any patent.

The variables that are used in my research can be divided into three groups: the original information derived from the patent citation data, the variables and statistics created by the authors of the dataset and the new statistics I created. I provide a description of the variables in Appendix 1.9.1.

1.3.2 Methodology

In their identification of a GPT, [Hall & Trajtenberg, 2004] select some criteria that a GPT-candidate patent as well as its direct followers have to meet in order to qualify for a GPT. In my research, I go beyond just considering the initial GPT-candidate patent and its direct followers. I construct the evolutionary path for a given technology[15]. Thus, when I am looking at a particular technology, I first identify the patent

more precise on how much they resemble evolutionary paths. The attempt to contact the author on that matter did not provide any additional insights.

[15] I need to thank Guillermo Fuentes from University of Waterloo Arts Computing Office for helping in coding a program that extracts the patents relevant to a given technology from the data

itself, I call it "the basement patent", then I identify the patents that quote this patent directly "the level 1 patents", then the patents that quote level 1 get extracted and so on until the end of the dataset.

Due to the size of the data and in order to save the computing time as well as not to observe the same patents multiple times, I do not keep the duplicate observations of the same patent in the resulting dataset. The following important assumption allows me to eliminate the duplicate observations in a rational and justifiable manner:[16]:

Assumption 1 *All the prior art*[17] *is necessary in order for a given technology to exist.*

This assumption allows one to structure the resulting data in the following way: If the same patent appears on both levels n and $n-1$, only the observation on the level n will be kept since there was some research activity on the level $n-1$ necessary for this patent to exist, therefore this patent must be placed after the level $n-1$. Besides the variables that existed in the original data two new variables are added to the output data: the variable "Level" indicates the position level of the current observation relatively to the basement patent, the variable "Group" is just the basement patent's seven digit ID and it is the same for all the observations belonging to the evolutionary path of the same technology. Table 1.2 schematically describes the resulting evolutionary path datasets.

Table 1.2: Resulting Evolutionary Path Dataset Structure

Citing patent ID	Cited patent ID	Grant year of the citing patent	Other variables	Level	Group
2	1	1976	$\cdots$	1	1
3	2	1977	$\cdots$	2	1
4	2	1978	$\cdots$	2	1
5	4	1979	$\cdots$	3	1

Table 1.2 is derived from the example in Table 1.1. The variable "Group" denotes that all the patents in Table 1.2 are the followers of patent number 1. The level variable represents stages of research activity. For example, there is a patent with the identification number "1" that, besides representing some practical application of knowledge of its own, also provides grounds for further research. According to the above assumption, this patent opens the door for the whole group of technologies represented by the patents with the identification numbers "2", "3", "4", and "5".

Unfortunately there is no perfect choice of data and ways of dealing with it when conducting empirical research. All methods and choices have some caveats. Foreseeing possible directions of criticism of the choices made, I list certain problems and disadvantages associated with the choice of data and methodology in [Oystrakh, 2014]. The problems can be split into three groups: genuine patent data problems, technical problems, and methodology problems. Some of the problems can be dealt with, others will have to be ignored.

[16]However one should remember that as with all assumptions, there are caveats, in our case inventors sometimes cite "previous art" only because of the legal motivation and not because the previous art played an important role in their research. Other times citations are added by an examiner and there is a possibility of an inventor being unaware of certain prior art included to the patent. However, a study by [Duguet & MacGarvie, 2007] finds that "patent citations are indeed related to firms' statements about their acquisition and dispersion of new technology, but that the strength and statistical significance of this relationship varies across geographical regions and across mechanisms for technology diffusion."

[17]In most of the systems of patent law, prior art is all the information that was available before a certain date and is considered to be relevant to the given patent's claims and originality, and therefore must be cited in the patent documentation. [United State Patent Act, 1836,]

1.4 Analysis of Single Patent GPT Candidates

[Hall & Trajtenberg, 2004] provided a methodology for identifying a GPT in the patent citation data as well as listing 20 patents that best fit all the criteria for the GPT.[18] When I began constructing evolutionary paths of technologies, the obvious decision was to look at these patents first.

When the evolutionary paths were constructed, it appeared that only 9 out of the 20 patents selected by Hall and Trajtenberg (2004) were independent i. e. did not belong to another evolutionary path originating from one of the patents in the group of 20 GPt-candidate patents. The other 11 patents belonged to at least one of the evolutionary paths derived from the nine independent patents. Moreover, when the remaining 9 GPT-candidate patents' paths are analyzed, it appears that more than 95% of observations in these paths are also the observations of some other path within the group of 9. For example 99.8 % of the followers of the patent 4528643 are also the followers of the patent 3956615. Therefore these paths are not independent at all.

Table 1.3 lists the GPT-candidate patents from [Hall & Trajtenberg, 2004]. We observe that some of the patents are not independent and belong to an evolutionary path(s) of another patent (other patents) within the same group of the 20 GPT-candidate patents. The patents that are called independent in Table 1.3 do not belong to any of the evolutionary paths within the same group of 20 patents. However, the subsequent observations in their evolutionary paths are mainly coincident within the observations of the other evolutionary paths in the group of 20.

Table 1.3: [Hall & Trajtenberg, 2004] Selected Technologies†

PID	HJT cat.*	HJT Sub-cat.**	Grant Year	Subsample of‡	on the Level	Subsample of‡	on the Level	Independent⋆
3624019	1	15	1971					√
3636956	3	32	1972					√
3842194	2	24	1974					√
3956615	2	21	1976					√
4575621	5	59	1986	3956615	5			
4528643	2	22	1985					√
4558413	2	24	1985					√
4672658	2	21	1987					√
4783695	4	46	1988					√
4821220	2	22	1989	3956615	5			
4885717	2	22	1989	4558413	1	3956615	5	
4916441	2	21	1990	3956615	8			
4953080	2	22	1990	4558413	1	3956615	8	
5133075	2	22	1992	4558413	3	3956615	9	
5155847	2	22	1992	4558413	3	3956615	6	
5093914	2	22	1992	4558413	2	3956615	6	
5347632	2	22	1994	4528643	4	3956615	8	
5119475	2	23	1992					√
5132992	2	21	1992	4528643	4	3956615	9	
5307456	2	23	1994	3956615	10			

† "Highly Cited Patents with High Generality, Class Growth, and Citing Patent Generality" From Table 8 in [Hall & Trajtenberg, 2004]
* Technological category, see Appendix 1 in [Hall et al., 2001b]
** Technological Subcategory, see Appendix 1 in [Hall et al., 2001b]
‡ The patent belongs to the evolutionary path(s) of another patent within the table
⋆ The patent does not belong to any evolutionary path of another patent within the table

In Table 1.4 I demonstrate the degree of coincidence within the evolutionary paths of the independent patents. I show the percentage of coincident observations within the patents that are themselves the

[18] See Hall and Trajtenberg, 2004 , Table 8; also re-listed in Table 1.3.

basement patents for the 11 non-independent patents.[19] Thus, most of the 20 GPT-candidate patents share a common history in terms of subsequent research.

Table 1.4: Coincident observations in historical paths

PID	3956615		4528643		4558413	
	Number of coinciding	%	Number of coinciding	%	Number of coinciding	%
3956615	638375	100	224633	99.8	273452	98.4
4528643	224633	35.2	225023	100	183946	66.2
4558413	273452	42.8	183946	81.7	277806	100

I collected a sample of 1000 evolutionary paths originating in the year 1976, so that the information from the whole dataset could be represented by the sample. The sample was used for various purposes among which was the calculation of pairwise coincident observations of every path in the sample with every other path (999000 observations in total). It appears that the average percentage of coincident observations between two paths originating in the year 1976 is 10.4%, which is very far from 99.8% or 98.4% coincidence between the GPT-candidate paths.

There are four possible explanations of the interdependence of these patents:

1. All the GPT candidate patents are closely related to the same GPT. None of them stands for a GPT of its own. Each of the patents in Table 1.3 is an important SPT derived from the GPT.

2. Only one of the patents in the group stands for a GPT (the earliest patent), the rest are important derived SPTs

3. Some of the patents in the group stand for independent GPTs and some are important SPTs

4. All the patents represent independent GPTs. And the followers, the secondary inventions are just some technologies that use combinations of these GPTs.

It seems that the fourth explanation is the least realistic. The degree of interdependence of these patents makes them too close to be independent GPTs. Moreover, observing 20 active and evolving GPTs at the same time would be an unprecedented case for the history of human kind. [Lipsey et al., 2005] using historical analysis have identified 24 GPTs from the ancient times up until now: 8 ancient GPTs that were active between 10000BC and 1450AD, 6 New Age GPTs that were active between 1450 - 1850, 9 modern GPTs, and 1 potential 21st century GPT (Nanotechnology). Even this list of 24 GPTs observed throught history of human kind is criticised as too extensive in [David & Wright, 1999]. The presence of 20 or more GPTs in the time range of 30 years seems too unrealistic from the GPT theory perspective. In [Oystrakh, 2014] I provide a historical analysis of GPTs and discuss how often we should expect to observe a new GPT. Many other authors agree that new GPTs emerge once in a while and take some time to develop. Observing multiple emerging GPTs is highly unlikely.

If the third explanation was true then we could conclude that the three patents listed in Table 1.4 could be GPTs[20]. However, the patents 4528643 and 4558413 have most of their subsequent history coinciding with the history of the patent 3956615. Thus, if the patents 4528643 and 4558413 are independent GPTs,

[19] Note that patents with earlier grant years will have a smaller percentage of coincident observations with later patents for obvious reasons.

[20] the remaining 6 of the nine independent patents while are closely connected to the three patents in in Table 1.4, do not contain any other evolutionary path derived from the group of 20.

then the rest of the patents in Table 1.3 could also be independent GPTs. Therefore, case 3 is a special case of case 4 that itself is very unlikely.

Now we need to compare the explanations 1 and 2. If there is a GPT patent in Tables 1.3 and 1.4 , this would be the patent 3956615 since most of the observations in the evolutionary paths of the other patents also belong to the path of this particular patent. Alternatively, if the first explanation was true, this patent could be one of the important SPTs derived from an active GPT but not a GPT of its own. In that case, the rest of the patents are closely related to this particular patent because this patent was granted earlier and had a higher chance to be cited.

Patent 3956615 requires closer attention. The technology this patent stands for is "transaction execution system with secure data storage and communications". Without diminishing the importance of this technology, one can hardly claim that this technology is of general purpose given that from its name it is designed for transaction processing. The patent was filed in 1974. Since then it would be acknowledged as an active or at least a potential GPT by many researchers and transaction execution systems would make it to the list along with computers, Internet, biotechnology, and nanotechnology as a subject of modern GPT research which clearly never happened.

A more formal argument in favour of patent 3956615 not standing for a GPT on its own is that most of the observations in the evolutionary path of that patent coincide with the observations of the microcomputer GPT identified in the next section. Thus, we exclude patent 3956615 from the list of potential GPTs for the same reasons we excluded the rest of the 20 patents. The subsequent history of that particular technology is a subset of another technology.

Finally, the list of 20 patents in Table 1.3 is not exclusive. By replicating [Hall & Trajtenberg, 2004] GPT-identifying technique, one can come out with more patents that could qualify for a GPT within the same data. Thus, the argument that we observe that many active GPTs simultaneously becomes even less valid.

As it was discussed earlier, the [Hall & Trajtenberg, 2004] GPT-identifying method requires looking for patents that are outliers by certain attributes. One of such attributes in the subsequent growth of the patent class. This is one of the key attributes and if it is switched off when we search for the outliers, most of the 20 GPT-candidates would not make it to the short list. Instead of them various patents from "mechanical" and "miscellaneous" categories rather than "computer" category would turn out as outliers. Those patents would hardly qualify for GPTs that emerged in 1970s-1990s.

Preserving the "subsequent class growth" selection criterion makes the task of predicting an upcoming GPT by looking at a fresh patent irrelevant. Since you would have to observe the subsequent class growth for some time. One of the reasons for looking at the outlying patents - is the early identification of a GPT at its embryonic-stage, before the GPT establishes itself as such. At this stage the GPT has some limited amount of applications. By specifying the subsequent class growth criterion, we deliberately eliminate the patents that could have been GPTs in their early stages, and thus have limited quantity of applications. We, basically, pre-assign that the patents will be from the fastest-growing ICT technological field. Therefore, the very early GPT-identification motivation of the analysis can be abandoned anyways, no matter whether one uses the [Hall & Trajtenberg, 2004] identification techniques or "first have a GPT, then describe it" scheme that I propose in the next section.[21]

The methodology of identifying GPTs in the patent data based on outlying patents has several weaknesses. The only advantage of this methodology over the methodology proposed in this research is that it could identify an upcoming GPT at the early stages of its life-cycle. Unfortunately, as we had seen, in practice this does not work. A researcher still needs to wait for some time to pass and observe subsequent growth of the class of an outlying patent. Moreover, given there are millions of patents in the data, there are hundreds or thousands of patents that can be considered "the outliers". While we do not expect that

[21] The early prediction task would not be feasible with the [Hall & Trajtenberg, 2004] methodology even when we exclude the "subsequent class growth" criterion. Another criterion in the methodology is the "citations' generality" which implies that a researcher has to wait for a patent to acquire some citations.

many active GPTs at the same time period. Thus, even if there was a GPT that originated from a single patent, it would be difficult to identify it among hundreds of other outlying patents in the data. However, it possible to speculate that most of the outlying patents and evolutionary paths are somehow related to the active GPT. This topic will be discussed in Section 1.7.[22]

1.5 Identifying a Known GPT in Patent Citation Data

1.5.1 Choosing the GPT

I introduce an alternative way of describing a GPT using the patent data. The objective is to find some GPT that, preferably, is in its active phase from the 1970s.[23] The so called ICT revolution that began in the 1970s has definitely reshaped the whole economy. It is questionable whether the ICT revolution consists of more than one GPT. For example[Lipsey et al., 2005] list computers and the Internet to be independent GPTs. Nevertheless it is widely accepted in the GPT literature that the information technology based on microprocessor-computers is of general purpose. Therefore, the GPT that is analyzed using patent citation data is the one that had an enormous influence on economic activities is the GPT that I denote the "microprocessor" and imply the computer-technology based on a microprocessors.

Microprocessors perform all the basic functions of a computer [Summer, 1985] and, in general, modern computers (PCs, Servers, or automobile computers for example) run on microprocessors. Jumping a little bit ahead, it is necessary to mention that the two most famous of the four microprocessor patents, discussed below, are also called "Computer" patents. It is questionable whether the "Microprocessor" should represent a GPT by itself or it is one of the major-amendments of the "Computer" GPT, or it is the invention that opened the door for PC-GPT. The question is more philosophical, than practical since we are talking about the same technology. Nonetheless, I am more inclined towards the microprocessor being the improvement that had moved the "Computer" GPT from its embryo-phase to the active phase. [Summer, 1985] describes the microprocessor: "With so many circuits located on a chip, it was only a matter of time before someone got the idea that one chip could contain all the circuits necessary to perform basic functions of a computer... The microprocessor was a general-purpose computer that could be programmed to do any number of tasks, from running a watch to guiding a missile." Therefore, a microprocessor is a computer, that is more compact and general, but basically it has all the functions formerly performed by various "big computers".

Nevertheless, it is necessary to formally show that the microprocessor-computer technology is indeed of general purpose. For this task I use our theoretical knowledge of GPTs described in Appendix 1.9.3 and the definition of GPT by [Lipsey et al., 2005] that argue that "a GPT is a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects."[24]

Whether microprocessors are a separate GPT or one of the major core improvements of the "Computer" GPT, it should be obvious that there is a GPT in the active phase that is associated with micropro-

[22]This research is not the only one that concludes that outlying patents do not necessarily represent very important or fundamental technologies by themselves. And therefore, there is not enough evidence to claim that certain outliers are GPTs. [Strumsky et al., 2012] use technology codes to investigate novelties of particular patented inventions. By looking at the patent #4723129, which is an outlier by citations received and therefore by it generality the authors conclude that the patent is not novel, all the technologies combined within the patent existed before the invention. The patent just assembles them together. This conclusion demonstrates, first, that patents are applications of some technological knowledge, thus we should not expect a GPT originate from some patented invention. Second, outlying patents are not necessarily the most innovative. [Bresnahan, 2010], in his discussion of the HJT methodology of identifying outlying patents concludes that the methodology "is good for identification of items with GPT features in the list of items, rather than for identification of GPTs themselves".

[23]Since the range of the dataset starts in 1970s.

[24]A reader can find a more detailed description of the microprocessor GPT, how it resembles other historically observed GPTs, and how well it fits the GPT definition and stylized facts by reading [Oystrakh, 2014].

cessors because a microprocessor (or a microcomputer) corresponds to the following GPT features:

- It provides one of the fundamental GPT services[25]: Processing, Storage, and Exchange of Information.
- There were first academic discoveries that led to the development of the technology. It is also true, that first prototype microprocessors were developed for the military sector and therefore were a product of publicly funded research. This feature is shared by most of the modern GPTs. They are discovered by formal scientists and only later some marketable applications were introduced.[26]
- While being improved, microprocessors and computers still preserve their main features and the associated names. Thus, it is a single generic technology, recognizable as such over its whole lifetime.
- It is difficult to find an industry where microprocessors are not used. All modern computers, appliances, means of transportation, and communication technologies rely on microprocessors. Thus, the microprocessor that initially had much scope for improvement eventually became widely used, having many uses and many spillover effects. It also becomes very pervasive and is used in many sectors of the economy; a typical behaviour of a GPT.
- Microprocessor-computers created new industries and sectors of the economy, for example a video game industry.
- Besides simplifying all existing research, [27] microprocessor-computers created or stimulated significant amount of further research that includes both software and hardware engineering. Thus, this technology enhanced the productivity of R&D in downstream sectors.
- A microprocessor on its own has very limited use. It should be applied along with complementary hardware and software, therefore effective use of this technology requires complementary investment in the using sectors.
- Microprocessors double their capacity approximately every two years in accordance with Moore's Law [Schaller, 1997]. Thus, this technology initially had much scope for improvement and it is still being improved on a regular basis.

The microprocessor technology is the most obvious choice for a GPT to analyze given the time frame. However, it is possible to use two alternative technologies for demonstrative purposes: biotechnology and nanotechnology. Nevertheless, microprocessor technology has several advantages over the two aforementioned candidates. First of all, early microprocessor patents are at the very beginning of the dataset, while the earliest patents for the other two technologies are in the 1980s. Therefore more observations are available for the Microprocessor technology. Second, there is almost no doubt about the microprocessor technology being a GPT, while the general status of the other two technologies is questionable. For example, [Youtie et al., 2008] treats nanotechnology as a future potential GPT while [Nikulainen & Kulvik, 2009] find biotechnology to be a less general relative to both computers and nanotechnology.

[25] See [Oystrakh, 2014] for the discussion of the fundamental GPT services

[26] There are still debates about the fact that the Central Air Data Computer(CADC) was the very first microprocessor. Nonetheless, being a military technology it was not patented and was kept in secrecy until 1998. Also, while having many microprocessor features it misses the important feature of having more than one chip. Nonetheless, the existence of CADC only supports the importance of not for profit research in the development of GPTs.

[27] For example, it is much easier to run a regression today, than it was before, when people had to invert big matrices manually.

1.5.2 Taking the GPT to the data

Now, having the GPT it is possible to identify it within the data. The first microprocessors existed in the early 60s. There are two patents issued by Intel that describe the Intel 4004 microprocessor which is believed to be the very first commercial microprocessor. There are also two patents from Texas Instruments Incorporated, one describing the first calculator operated by a microprocessor and the other has the unofficial name "The Microcomputer Patent". All these patents were applied and granted at approximately the same time, the "Microcomputer" patent appearing a little bit later than the others (See Table 1.5). These patents represent the starting point of the GPT in the patent data.

Using these patents as the original nodes[28] I construct an evolutionary path that describes the evolution and diffusion of the GPT. Each position level in the path indicates the next round of research activity. By arranging the resulting evolutionary paths by position levels and also by grant dates one can observe how the quantity of patents related to the GPT grows and the penetration of the GPT-derived applications into various technological fields occurs.

Table 1.5: Initial "Microprocessor" GPT Patents

PID	Grant year	Name	HJT class	Citations Received*	Assignee
3753011	1973	Power Supply Settable Bi-Stable Circuit	Electrical Devices	11**	Intel
3757306	1973	Computing Systems CPU	Computer Hardware & Software	78**	Texas Instruments Inc.
3821715	1974	Memory System for a Multichip Digital Computer	Information Storage	83**	Intel
4074351	1978	Variable function programmed calculator	Computer Hardware & Software	33***, 46	Texas Instruments Inc.

* None of the patents in the table has made a citation to some previous art, therefore the statistics "citations made" and "originality" are equal to zero.
** The parameters were taken from 1999 data edition since these patents are out of scope of the 1976 - 2006 data; the generality is also taken from the 1999 data edition where it was computed by the authors of the data [Hall et al., 2001b].
*** The first statistic was taken from the 1999 data edition for consistence; the second statistic is from the 2006 data edition where generality was computed by the amended formula.

1.6 Results: GPT in the Patent Data

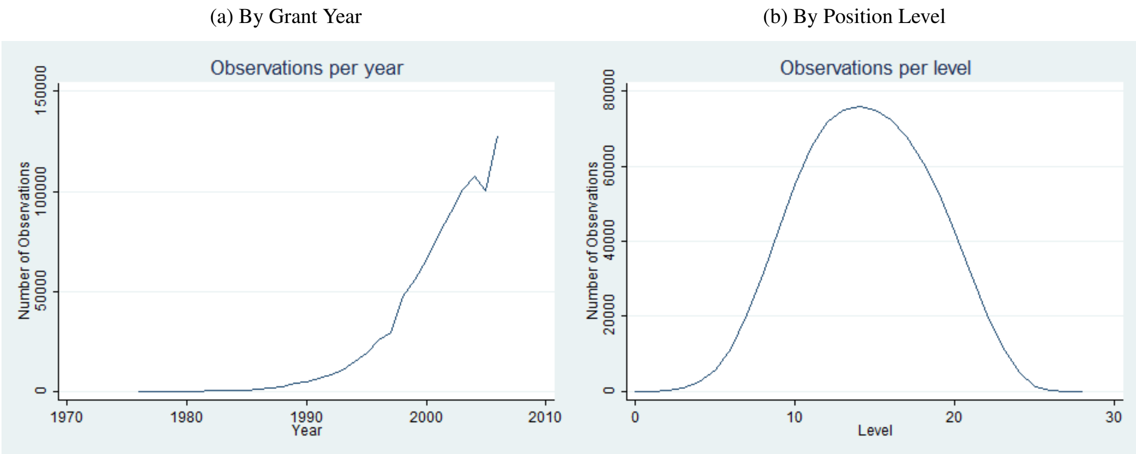
In this section, I describe the resulting evolutionary paths of the Microprocessor GPT and demonstrate the effect the technology has made on the overall level of technological knowledge. Recall that we are looking at the "microprocessor" patent collection that originated from the four patents listed in Table 1.5. The evolutionary path obtained from the four original patents includes 901,968 observations (patents) and 28 position levels (rounds of research activity) by year 2006. The quantity of observations increases both over time and over rounds of innovations (Figure 1.1). When we arrange patents by grant years (left)[29] , we observe an exponential growth of observations over time. When we arrange patents by position levels (right), we observe a bell-shaped curve. Thus, first there is an exponential increase in the number of patents, that eventually reaches its peak and then we observe a decrease. The decrease mainly comes

[28] Note that none of the four patents makes any citations to any prior art!

[29] We could also arrange patents by application years, and we would observe almost the same picture as in the case with grant years. It appears that application and grant years are highly correlated, and therefore we would not achieve much by presenting patents by application years except a shift by the average processing lag.

from the data truncation. As we would expect many of the technologies today have some IT component in them. The evolutionary paths support this notion by exhibiting exponential growth of technologies related to the microprocessor computer.

Figure 1.1: Observations in Microprocessor-Computer Evolutionary Path



All the HJT technological fields and subfields are represented in the evolutionary path of the microprocessor GPT. Figure 1.2 (left) demonstrates the distribution of the observations by HJT categories. We observe that the "home" category that is called "computing" stands for almost 50% of all the patents in the past.[30] However, the technology has diffused a lot to other sectors. The diffusion of the technology can be observed better from the right panel. The right panel shows the percent of observations (patents) that belong to the "home" subcategory ("Computer Hardware & Software") arranged by the position levels. We observe at the beginning of the technological life-cycle most of the observations belonging to the the initial subcategory, however, the technology as it is being developed diffuses to the other fields having roughly 20% of its observations remaining in the "home" subcategory.

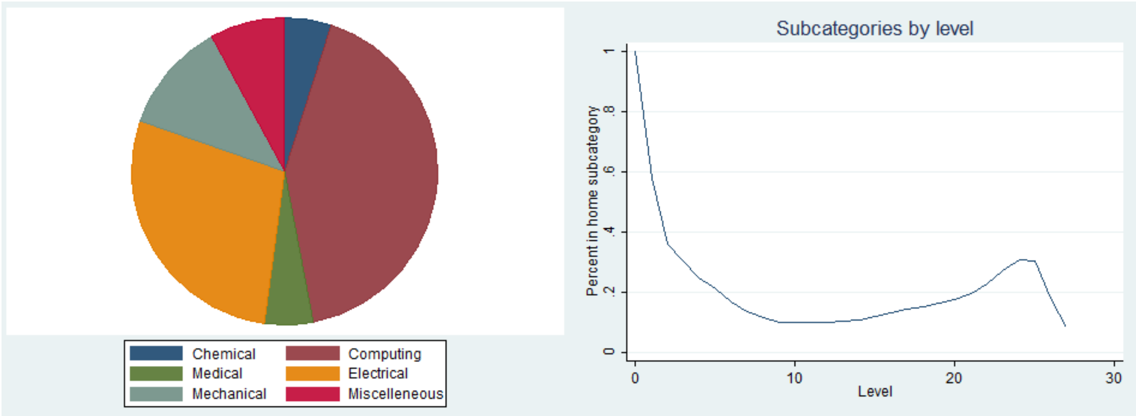
The "microprocessor" patent collection represents a proxy for the microprocessor-computer GPT's applications. By studying the collection and comparing it to the rest of the patents we can observe its effect on technological knowledge. Moreover, given that Assumption 1 holds, we will be able not just to establish the effects of the GPT, but also take a look at what the world would look like if this GPT did not exist. Figure 1.3 demonstrates the total quantity of US patents issued between 1976 and 2006 and how many patents there would be if there were no GPT.

[Kortum & Lerner, 1999] identified that there is a relatively stable quantity of patents from one period to another up until the early 1980s. After that there is sharp increase in patenting activity. They use cross-sectional analysis to investigate whether the growth in the quantity of patents in the US should be attributed to the changes in US-patent policy or to the widespread introduction of information technology. They conclude that information technology, which facilitates research as well as opens all new directions for the innovative activity, is the main cause of the boost in the patenting activity.

[30] Intel decided to break its invention of microprocessor into two patents (see Table 1.5). Moreover, these patents belong to different patent classes first taking the "computer technology" part of the invention and second dealing with the electrical part. It is also worth noting that the first patent received a much greater citing activity, than the second. Most likely, the first patent carried a greater part of the innovativeness while the second was just an adaptation of an earlier existing technology for the current needs. Thus, I consider the computer category to be home category for the Microprocessor technology even though one of the four basement patents belongs to the electrical category.

Figure 1.2: Diffusion of the Microprocessor-Computer GPT

(a) Distribution over HJT Categories (b) Percent in "Home" Subcategory

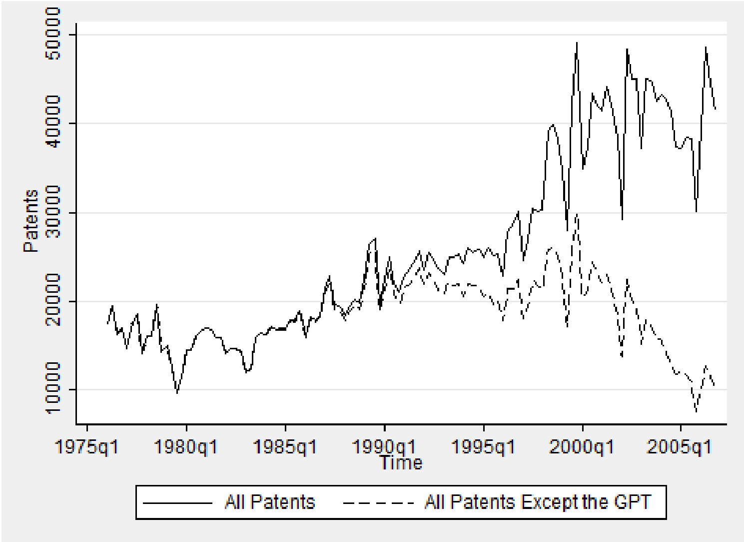


[Rafiquzzaman & Whewell, 1998] derive a conclusion similar to that from [Kortum & Lerner, 1999] but using Canadian data.

In accordance with the aforementioned studies, the current study suggests that it is very likely that we observe the effect of the long run positive technological shock, as seen on Figure 1.3. It is believed that GPTs rejuvenate the economy, and are the only possible engine of economic growth not subject to diminishing returns.[31] This belief appeared out of historical observations, however, now we can support it with the data.

[31] See the discussion on the role of macroinventions in the history of human civilization in [Mokyr, 1990].

Figure 1.3: All Patents Granted in US between 1976 and 2006



Patents contain claims for technological improvements. The US-patent data can be transformed into a proxy of the aggregate technological process. We observe the "stable" and "growing" periods of technological innovation activity. Growing periods, at least the last one[32], can be attributed to the active phase of the "microprocessor" general purpose technology.[33]

1.7 What about the outlying patents?

In this section, I answer the question about the relationship between the outlying patents and the GPT. We have established that the patents identified by [Hall & Trajtenberg, 2004] methodology do not represent the GPT but have some GPT features. We have also established great degree of interdependence of the evolutionary paths constructed from the patents identified by [Hall & Trajtenberg, 2004] methodology. Moreover, these evolutionary paths are also highly related to the "microprocessor" GPT evolutionary path. Therefore, it is necessary to establish whether the relationship of a technology to a currently active GPT makes certain technologies "outliers" and distinguishes them from other technologies.

The premise of the analysis is that any technology that gets related at some point of its lifecycle to a GPT gains from the relationship. The gain is realized the following way, the relative importance of a technology that is closely connected to a GPT increases and it is more likely to become an outlier in comparison to other technologies. Thus, patents identified using the [Hall & Trajtenberg, 2004] methodology are not GPTs by themselves but have GPT features due to their relation to the GPT. The argument can be summarized the following way, if some technology at some point is cited by a patent in a GPT's path, its subsequent citation history changes.

It is necessary to clarify how a technology becomes related to a GPT. A GPTs consist of core and peripheral improvements. Distinction between core and periphery innovations can be found in other literary sources, see [Carlaw & Lipsey, 2011] or [van Zon et al., 2003] for example. Any pre-existing or

[32] Even though it is possible to attribute the preceding 1910-1930 periods to the "Electricity" based on anecdotal evidence.

[33] The importance of the computers in the economic growth in 1980s 1990s is established in other studies, see [Oliner & Sichel, 2000] for example.

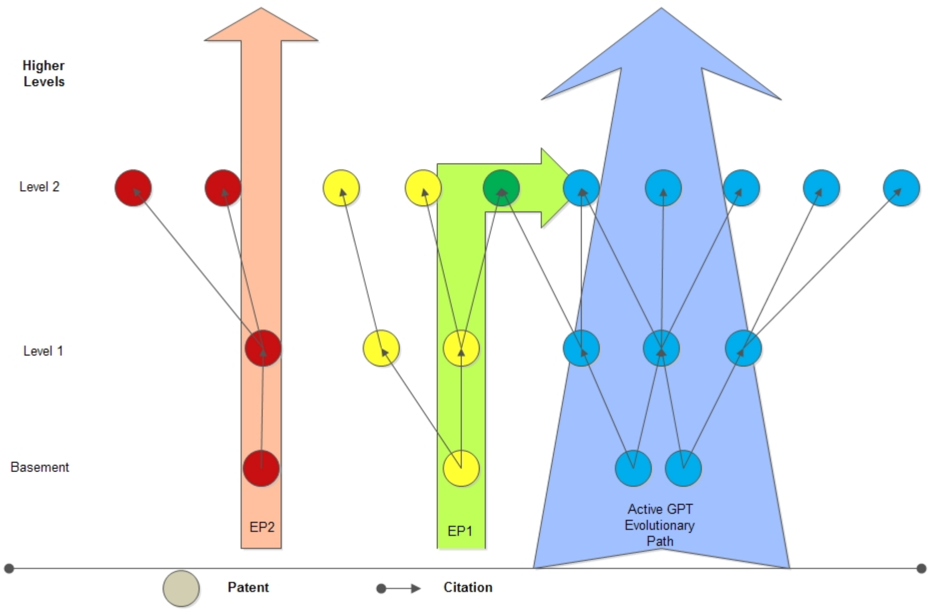
co-existing with a GPT technology can become related to the GPT by merging with one of its periphery branches. For example, automobiles existed before computers, however, modern automobiles are now highly computerized. Therefore, there are certain patents that belong to mechanical category since they mainly describe some mechanical processes in automobiles, however, these patents integrate certain computer technologies and therefore cite certain computer prior art.

Figure 1.4 depicts (over-simplistically) an evolutionary path for some GPT and two additional evolutionary paths. The first evolutionary (EP1) merges the GPT at the second level, the second path (EP2) does not merge the GPT at any point of its history. Note it is not necessarily that the second level of EP1 must coincide with the second level of the GPT, it could be the case that the second level of EP2 merges with the fourth level of the GPT, for example. Figure 1.4 has level 2 merging level 2 only for illustrative purposes, in order to save space on the diagram. Also, 1.4 represents two extremes: one of the technologies gets connected to the GPT very early, the other never gets connected. Technologies that connect to the GPT earlier would have an advantage in terms of being closely related and having a greater share of "common history" with a GTP. Also, it is possible for a technology to get connected to the GPT at several nodes independently.

The following examination should show that the evolutionary path EP1 is more likely to have more observations, and therefore be more technologically fertile than the evolutionary path EP2.[34] Another measure of technological fertility is how many rounds of research given technology had generated given that all the basement patents are granted in the same year 1976. Rounds of research are represented by the "position level" variable. We will examine the dependence of the number of position levels in an evolutionary path on the relation of this path to a GPT. It should also explain how a relation of a technology to a GPT affects parameters of the initial level 0 patent. This way, we can show that outlying patents are outlying due to their relation to an active GPT. Finally, we want to show an implication of the relation with a GPT on the diffusion of a technology over a technological fields. The premise is that technologies that become related to a GPT are more general themselves and diffuse to more technological fields.

[34] Most researchers agree that every patent has at least some technological novelty even though the technological value of patents may vary from one patent to another. See [Beaudry & Schiffauerova, 2011], [Strumsky et al., 2012], or [Rothwell et al., 2013] for example.

Figure 1.4: Schematic Representation of the Initial Levels of an Active GPT and 2 Evolutionary Paths



In order to analyze the effects of the GPT on other technologies, a sample of 1000 patents granted in 1976 that received at least one citation was collected. Evolutionary paths for each of the patents were constructed. As a result we obtained a sample of 1000 evolutionary paths.

The first objective of the analysis is to take a look at the quantity of observations in the constructed evolutionary paths and see if this quantity is affected by the relation to the microprocessor GPT. Similarly the number of rounds of research (position levels) is analyzed. Table 1.6 summarizes the quantity of observations (patents) and number of levels in the evolutionary paths. Figure 1.5 plots the sample densities. As we can see, the number of observations is not evenly distributed but rather leaned towards a few. Most of the randomly selected technologies have few observations. This fact also explains the difference between the mean and the median for the number of observations. The number of levels is more evenly distributed across the sample.

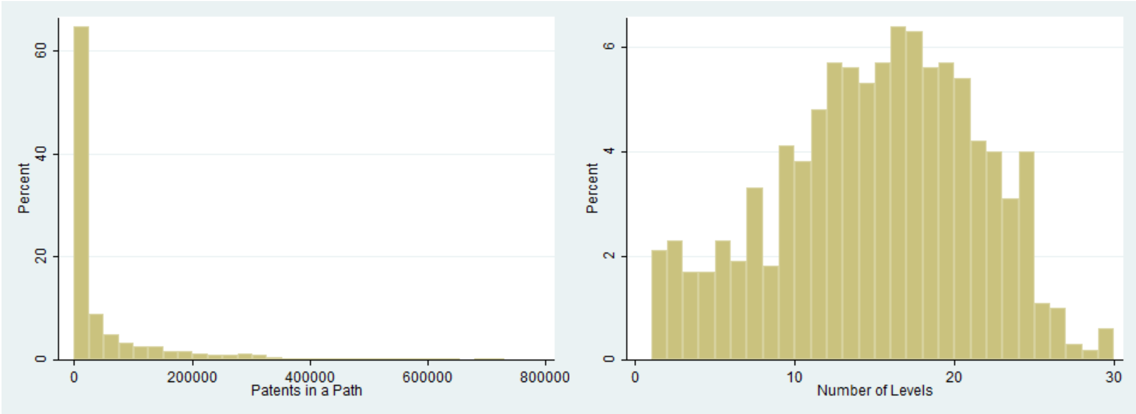
Table 1.6: Observations† in The Sample of 1000 Paths

Characteristic	Mean	Median	Minimum	Maximum
Observations	56771	6462	1	730485
Position Levels	14.625	6.270223	1	30

†Observations, in this context, are patents in a path and not paths in the sample

Figure 1.5: Sample Density of Observations and Number of Levels in Evolutionary Paths

(a) Observations (b) Levels



We may expect that the quantity of the observations in an evolutionary path would depend on the technological field of the initial level 0 patent. Table 1.7 provides the descriptive statistics about the distribution of the observations and levels by HJT technological categories. We observe that the IT patents (HJT Category 2) have much greater quantities of observations and reach higher levels than the rest of the technological fields. The path derived from the IT patents on average have twice as many observations as the next significant field "Electrical" (HJT Category 4). It is also notable that the IT paths have their median closer to the mean implying less skewness of the distribution relatively to the rest of the groups of paths.

The next significant group are the paths derived from the "Electrical" patents. Even though the electrical GPT had passed its active phase long ago, unlike steam and water power technologies, electricity has preserved its wide universal use in most of the sectors of the economy. This GPT remains significantly important as all computers run on electricity. Recall that 1 out of 4 initial "Microprocessor" patents was an electrical patent. Moreover, this category also includes Electronic patents that are closely related to the Computer and Communication technological field.

The next significant field is "Drugs and Medical" (HJT category 3). Even though this field includes the "Biotechnological" sub-field that, according to [Lipsey et al., 2005], would be the "home" for the other modern GPT. We do not observe any biotechnological patents in the year 1976 (the year from which the level 0 patents were drawn) and we observe just 7 biotechnological patents between 1977 and 1981. Thus we would not observe a full life-cycle of a biotechnological path.

The other three categories, are so called "traditional" technological fields [Hall et al., 2001b]. As it is expected, these fields have much lower citing activity than the "modern" fields.

In order to estimate the effect of the GPT on the technological diffusion we use two diffusion variables introduced in Appendix 1.9.1. The percentage of observations in the home category and the simple Technological Diffusion Index (TDI) that are derived in equations 1.10 and 1.12 respectively.[35] We expect that the relation of a technology to an active GPT benefits the technology by making it more general itself. Therefore, the relation to the GPT should stimulate the diffusion across categories and sub-categories.

Table 1.8 presents the results of regressions of the four variables described above. For each of the dependent variables of interest, four log-log regression models were estimated. The log-log approach was

[35]Instead of the percentage in home category the percentage in home sub-category could have been used without affecting the conclusions.

Table 1.7: Observations† and the Number of Levels in The Sample of 1000 Paths By HJT Categories of the Origin Patent

Initial Category*	Observations		Levels		Quantity of Paths
	Mean	Median	Mean	Median	
Chemical	26329	3902	13.3	14	250
Computer & Communications	198201	186957	18.9	22	62
Drugs & Medical	72204	12495	14.8	16	50
Electrical & Electronic	108405	49035	16.9	18	169
Mechanical	37865	4977	14.4	15	245
Others	29882	3580	13.4	14	224

† Observations, in this context, are patents in a path and not paths in the sample
* The category of the initial Level 0 path originating patent

chosen for the following reasons: the histograms of some variables are skewed and logarithmic transformation makes the data more "normal looking". Also, it is easier to interpret log-log regression coefficients. Certain variables, for example "percent of observations coinciding with the Microprocessor GPT" or "percent of observations in home category" have meaningful zero values, in order to transform these variables into a log form without losing the values the following approach was taken: $x'_t = \ln(x_t + 1)$ where x_t is the original variable and x'_t is the transformed variable.

The following are the four log-log regression models were estimated and presented in Table 1.8

- In the first model we regress the dependent variables on the main variable of our interest the fraction of observations in an evolutionary path that coincide with the observations in the "Microprocessor" GPT. This variable indicates how closely related a given evolutionary path is to the GPT.

$$\ln(depvar_i) = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + u_i \tag{1.1}$$

- Relation to the GPT is not the only variable that may affect the significance and diffusion of a given evolutionary path. In the second model we add two other variables that according to the literature are important indicators of the patent quality. What if certain characteristics of the basement patent affect the resulting evolutionary paths? According to [Lanjouw & Schankerman, 1999], [van Zeebroeck & van Pottelsberghe de la Potterie, 2011b] and many others, the number of claims and citations received are important patent quality indicators. These variables are included into the regression model 2.

$$\begin{gathered} \ln(depvar_i) = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + \\ +\beta_2 * \ln(\text{\# of claims})_i + \beta_3 * \ln(\text{\# of citations received})_i + u_i \end{gathered} \tag{1.2}$$

- We add two more explanatory variables to the third model. These are the patent quality indicator variables that, according to [Hall & Trajtenberg, 2004] serve as GPT indicators. In [Hall & Trajtenberg, 2004], [Moser & Nicholas, 2004] and other studies generality indices are used to determine the importance of a patent and even to claim that a patent is a GPT candidate. Inclusion of the generality variables is necessary for consistency of the analysis eith the other studies since these could be the main determinants of the patents' significance and diffusion of the patented technology. The two variables we introduce to the third model are the generality index for the basement patent (the improved version of it calculated by Equation 1.8 of Appendix 1.9.1 and the average generality index of the direct followers of the basement patents.[36]

$$\begin{aligned} \ln\left(depvar_i\right) = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + \\ +\beta_2 * \ln(\text{\# of claims})_i + \beta_3 * \ln(\text{\# of citations received})_i + \beta_4 * \ln(\text{generality})_i + \\ +\beta_5 * \ln(\text{direct citations' generality})_i + u_i \end{aligned} \tag{1.3}$$

- Finally, according to [Hall et al., 2001b] there is strong "field effect" on the citation activity of the patent. Therefore, the fourth model includes HJT technological field dummies. For that purpose 5 dummy variables were created. The variables take value "1" if the initial patent if from one of the six HJT categories and "0" otherwise. The dummy for the HJT category 6 "Miscellaneous" was not used, the intercept value was left in the linear regression model instead.

$$\begin{aligned} \ln\left(depvar_i\right) = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + \\ +\beta_2 * \ln(\text{\# of claims})_i + \beta_3 * \ln(\text{\# of citations received})_i + + \\ \beta_4 * \ln(\text{generality})_i + \beta_5 * \ln(\text{direct citations' generality})_i + \\ +\gamma_1 * D_1 + \gamma_2 * D_2 + \gamma_3 * D_3 + \gamma_4 * D_4 + \gamma_5 * D_5 + u_i \end{aligned} \tag{1.4}$$

The results of the regression of all four models are presented in Table 1.8 of Appendix 1.9.2. The top panel contains the regressions of the number of observations in a path and the number of position levels on the independent variables. According to all four models, a strong influence of the GPT on the evolution of technology can be traced. The GPT coefficients are always statistically significant and of significant size in comparison to other coefficients. Connection to the GPT is the main factor explaining why some evolutionary paths are greater than others both, in terms of number of patents and quantity of rounds of research.

From looking at the other explanatory variables we observe a positive effect of the number of citations and claims of the basement patent on the resulting size of the path. Naturally, more citations received by the basement patent increase the probability of a bigger evolutionary path both in terms of observations and position levels. The number of patent claims of the basement patent, however, does not produce statistically significant result. Surprisingly, the generality index, which is one of the main criteria in attributing a patent to a GPT in previous studies is not found to be statistically significant, while the average generality of the direct followers plays a more important role. From these observations we receive a reaffirmation that the initial patent itself is not as important as the subsequent history of the technology. The generality of the basement patent plays a less important role in determining the subsequent evolutionary path than the generality of the direct followers. Also, the average number of claims of the direct followers if included into the regression model (not reported) produces statistically significant results while the number of claims of the basement patent itself does not.

The field effects, not surprisingly, are of different significance. Naturally, a patent from the "Computer & Communications" technological field is more likely to get connected to the "Microprocessor" GPT. It also appears that paths derived from chemical patents differ from the other paths. Effects of other fields on the resulting paths are uncertain.

The lower panel of Table 1.8 shows the effect of the connection to the GPT and some basement patent attributes on the diffusion measures of evolutionary paths. As with the significance measures, we note that the most important role still belongs to the relation to the GPT variable. It appears that the number of claims in the basement patent has a statistically insignificant effect on the future diffusion of the technology, while the generality of the patent, unlike the top panel, produces significant results.[37] It is also

[36] Recall that direct followers are not only the level 1 patents in the evolutionary path. These are all the patents that cite the basement patent, some of them could appear on the higher levels

[37] It would be surprising if it was otherwise since the generality index is a measure of diffusion itself.

necessary to note that field effects do not play an important role in the diffusion of technology. Surprisingly, most of the fields, including the "Computers & Communication" actually reduce the diffusion.

Thus we have established from the first regression results that a connection to a GPT stimulates the research process for other GPT merging technologies. The evolutionary paths that are closer to the GPT have more patents (applications), more rounds of research activity and diffuse to a greater quantity of technological sectors.

The remaining question is whether a closer connection to the GPT affects certain attributes of the basement patent, makes it look more like outliers and therefore causes them to be identified as GPTs that the methodology of [Hall & Trajtenberg, 2004] helps uncover. Recall, that certain patent attributes (forward citation, generality, centrality) evolve over time[38] and therefore it is possible that a connection to a GPT in year $t + k$ would affect attributes of a patent issued in year t. In order to establish the effects of the connection to a GPT on the basement patent attributes, we run additional regressions on the sample of 1000 paths.

The following regressions will show that a connection with an active GPT does indeed affect certain patent attributes. [Hall & Trajtenberg, 2004] in their GPT identification scheme used patents' generality indices and the average generality of the direct citations as GPT-indicators. Therefore we use these two variables as dependent variables in a regression on the number of coincident observations in a GPT in the derived evolutionary path. We also use the degree centrality index introduced in Appendix 1.9.1 Equation 1.9 as an additional measure of patent generality. Since the 2006 vintage dataset does not include any backward citations for the patents granted in 1976, I extract the citations from the 1999 vintage NBER patent citation data. This version of data includes the backward citations back to 1963.

Table 1.9 provides the regression results of three aforementioned dependent variables that make certain patents look like GPT-candidates on the fraction of coincident observations with a GPT in the derived evolutionary path (model1).

$$\ln{(depvar_i)} = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + u_i \tag{1.5}$$

Besides that I add HJT technological field dummies (model 2) as extra explanatory variables to capture the possible field effects discussed in [Hall et al., 2001b].

$$\begin{aligned} \ln{(depvar_i)} = \alpha + \beta_1 * \ln(\text{\% of observations in GPT})_i + \\ + \gamma_1 * D_1 + \gamma_2 * D_2 + \gamma_3 * D_3 + \gamma_4 * D_4 + \gamma_5 * D_5 + u_i \end{aligned} \tag{1.6}$$

A strong effect on the patents' generality measures from the connection to the GPT is observed in Table 1.9. On the other hand, the technology field effects are insignificant and sometimes counter-factual. For example we would expect a patent belonging to the "Computer" technological field to have higher citation generality and degree centrality. Thus, the fact that the subsequent technological history of a patent is closely connected GPT plays an important role in determining the forward looking attributes of the patent such as the generality index. We can conclude that patents identified as GPT-candidates by the [Hall & Trajtenberg, 2004] methodology do have a connection to the GPT even though they do not represent the GPT by themselves.

1.8 Conclusion

I have the following conclusions and suggestions for future research:

1. General purpose technologies are so big and significant that they cannot be encompassed in one patent. Therefore, the previous methodology of looking for outlying patents allows us to retrieve

[38] See the discussion in [Gress, 2010]

patents that are likely to be closely related to a GPT, but do not represent a GPT of their own. One can find hundreds of outlying patents which does not lead to a conclusion that there are hundreds of co-existing GPTs. Instead of looking for an outlying patent and checking whether it stands for a GPT, one could first find a certain general purpose technology and then trace it within the patent data.

2. An active general purpose technology reshapes and rejuvenates the economy, stimulates aggregate research activity and finds multiple applications in the various sectors of the economy. This study supports the historical GPT literature in providing empirical evidence of a GPT stimulating the research and patenting activity on the aggregate level, leading to a research and patenting boom. We also observe how after a slow start and a narrow set of applications, a GPT evolves into a technology that is widely used in multiple technological fields.

3. We also observe how a GPT affects, transforms, and stimulates other technologies. These technologies in return benefit from being closely related to a GPT. We found that the relation to an active GPT is the key factor thatcan make a patent outlier in terms of certain characteristics. The relation to the GPT would also affect significantly the future evolution and diffusion of a given technology.

4. Finally, this study introduces a useful tool that allows us to examine the evolution and diffusion of a given technology over time and over rounds of research activity. This tool is the evolutionary path.

This study is a small step on the long way of better understanding the nature of GPTs and the nature of technologies in general and their role in economics. There are plenty of issues that have to be addressed by future researchers. One possible extension of the current study is to apply evolutionary paths to examine other historically identified GPTs. Especially if the time range of the patent citation data expands. Another question is to prove that the patent burst of the 1980s is at least partially attributed to the arrival of a very fertile GPT and not to the changes in the US patent policies as some studies suggest.[39]

[39] The debate about the reason(s) for the surge in patenting activity in the United States is still open. There are multiple publications in favour of the patent surge being caused by a change in the US patent policy in 1982 ([Hall & Ziedonis, 2001] or [Turner, 2004] for example) as well as publications that doubt this hypothesis and believe that a fertile technology was mainly or partially responsible for the patent surge ([Kim & Marschke, 2004] [Kortum & Lerner, 1999], or [Rafiquzzaman & Whewell, 1998] for example).

1.9 Chapter 1 Appendices

1.9.1 Variables and Statistics

The variables that are used in my research can be divided into three groups: The original information derived from the patent citation data, the variables and statistics suggested by the authors of the dataset ([Hall et al., 2001b]), and the statistics that I created. In the following section, I present all of the variables used in the research but give more attention to the variables I created or modified. Variables that are introduced in the current study can further be sub-divided in their own turn into patent describing variables and evolutionary path describing variables.

1. Original Variables
   - A seven digit patent identification number (PID)
   - Patent application year
   - Patent grant year
   - Number of citations received by the patent
   - Number of citations made by the patent. Unlike the previous editions, in the 2006 version of the data the "citations made" statistic was not provided. I derived this statistic for each patent from the cited-citing relationship data. Therefore, citations made to the patents issued before 1976 are not counted.
   - Number of citations made by patents granted prior to the year 2000. This is a supplementary measure of backwards citations taken from an earlier vintage NBER Patent Citations Data that has observations up to the year 1999. Even though this dataset starts from 1976, it has backward citations going back to the year 1963, therefore, when we analyze patents granted in the 1970s, this measure of the backward citation activity is preferred to the previous one.
   - Number of claims. Patent claims are the technological novelties contained in the patent that are subject to legal protection. Thus, the claims represent the technological improvements (secondary innovations) a particular patent presents.[40]

2. Variables introduced by HJT
   - 1 digit technological category that splits patents between six technological fields: Chemical, Computing, Medical, Electrical, Mechanical, and Others.
   - 2 digit technological subcategory[41] The HJT classification system introduced in [Hall et al., 2001b] is found to be the most convenient and comprehensive, therefore the technological fields in this paper will correspond to this system unless mentioned otherwise.
   - Generality and Originality Indices Generality measures are the measures of the wideness of the spectrum of technological fields represented by the followers of the given patent. In other words, if the patents that cite the patent of our interest come from the same technological field, the patent is not general at all, while if they come from various fields the patent is general.[42]

[40]Obviously not all the claims are always secondary innovations, however, usually claims that may include some prior art get rejected by patent bureaus. At the same time not all the secondary innovations get claimed since patents' authors sometimes do not include everything they could just to make sure they do not claim something previous patents had already claimed. For more detailed discussion of the patent claims see [Tong & Frame, 1994] and [Hall et al., 2001b].

[41]See [Hall et al., 2001b], Appendix 1, pp. 41 - 42 for the list of subcategories with their description and their relations to the US patent classification system.

[42]Note that the generality is used to describe the knowledge diffusion generality rather than application industry diffusion generality.

In [Hall et al., 2001b] authors present the following kind of generality and originality indices that are variations of the Herfindahl Concentration Index computed the following way:

$$Generality_i = 1 - \sum_{j}^{n_j} s_{ij}^2 \tag{1.7}$$

Where s_{ij} denotes percentage of citations received by patent i that belong to patent class j.[Hall et al., 2001b]. For the originality index the formula is the same except that s_{ij} denotes percentage of citations made by patent i that belong to patent class j. The indices presented above are higher when more citations received (made) from (to) the patents in different technological fields (different HJT subcategories in out case). However, as it was stated in [Hall & Trajtenberg, 2004] the indices are biased and give more weight to the patents with a greater total amount of citations received (made). Following Hall and Trajtenberg we are making the following adjustment to the indices: $\tilde{G}_i = \frac{N_i}{N_i - 1} G_i$, where N_i is the number of citations received (made).

- Citation truncation weight as of 2006. Allows to adjust the "citations received" measure for data truncation.

3. Variables Introduced in this Paper

   - Patent Characterizing Variables

     – Improved Generality and Originality Indices.
     In the conclusion of [Hall & Trajtenberg, 2004], the authors state that: "All of the generality measures suffer from the fact that they treat technologies that are closely related but not in the same class in the same way that they treat very distant technologies." Therefore, traditional generality indices would give the same values to a patent with a follower in a close field and to a patent with a follower in a completely distinct technological field. In reality, the second should be more general than the first. Below I present the original way of computing generality and originality indices. Then I propose a simple improvement that will treat contiguous technologies differently from the distant ones and thus fix the problem mentioned by the authors of [Hall & Trajtenberg, 2004]. The Herfindahl indices introduced above take into account the sub-categorical variety of the citing (cited) patents relatively to the patent i. A simple amendment will still catch the moves of the citing (cited) patents into different subcategories but will give more weight when the categories of the citing (cited) patents also differ.[43]

$$Generality_i = 1 - \sum_{j}^{n_j} \sum_{k}^{m_k} s_{ij}^2 s_{ik}^2 \tag{1.8}$$

     Where s_{ij} denotes percentage of citations received by patent i that belong to the subcategory j out of n_j subcategories, while s_{ik} denotes percentage of citations received by patent i that belong to the category k out of m_k categories. Obviously, the adjustment for unbiasedness discussed above applies to the improved generality and originality indices.

     – Degree Centrality is a simple statistic borrowed from the graphing theory, it is calculated the following way:

$$C_D(v) = \frac{deg(v)}{n-1} \tag{1.9}$$

[43] I introduce indices for the HJT patent classification system since it is mainly used in this study. However, the measures can be easily adopted for other patent classification system that have classes, sub-classes and even sub-sub-classes.

Where $deg(v)$ is a number of nodes connecting to the vertex, in our case it is the sum of citations made and received by the patent i; n is the quantity of vertices (patents) in the data. This variable indicates whether the relative importance of agiven patent in terms of both backwards and forward citations.

  - Adjusted number of citations received variable that is calculated by multiplying the citation received variable by the citation truncation weight variable introduced in [Hall et al., 2001b].

- Evolutionary Path Characterizing Variables
  - Observations. Number of patents in an evolutionary path.
  - Position Level (Level). Location of a particular patent in an evolutionary path relatively to the initial basement level zero patent. Following assumption 1, each patent appears on the highest level possible relative to the basement.
  - Group. Common indicator for all patents of the same evolutionary path. Becomes useful when comparing different evolutionary paths or looking for coincident observations. For the evolutionary paths with single basement patent the value of the group variable equals the PID of that patent. For the multi-patent path can take any meaningful value.
  - Percent of Observations in Home Category/Subcategory. After obtaining an evolutionary path one question we can ask is how far or by how much the technology diffuses. That is how far it spreads into other technological fields out of the initial "home" field. One of the ways to measure this diffusion is to compute the percentage of the patents in the evolutionary path that belongs to the home category or subcategory[44]. The variable indicating the percentage of the observations in the "Home" category (subcategory) is constructed the following way:

$$\forall t = (1976:2006) \quad X_t = \frac{\sum OBS_{t,i}}{\sum OBS_t} \tag{1.10}$$

Where $OBS_{t,i}$ is the observation in year t that belongs to HJT category/subcategory i.

  - General Technology Diffusion Index (TDI), the way of measuring technological diffusion using percentages in home categories/subcategories presented above has a significant disadvantage since it measures the diffusion from the initial category to "somewhere else". On the other hand, we would like to find out by how much a technology disperses over all the categories. For that purpose we introduce the Technological Diffusion Index. TDI is a time series variable from the family of Herfindahl Indices that is defined the following way:

$$TDI_{ti} = 1 - \sum_{j}^{n_j} \sum_{k}^{m_k} s_{tij}^2 s_{tik}^2 \tag{1.11}$$

Where TDI_{ti} denotes technological diffusion index for the collection of patents i in the time period t.[45] Also s_{ij} denotes percentage of patents in the evolutionary path i granted at the period t that belong to the subcategory j out of n_j subcategories, while s_{ik} denotes percentage of patents in the evolutionary path i granted at the period $t$that belong to the category k out of m_k categories.[46] TDI is helpful in our analysis in two ways: first it

[44]The category/subcategory of the original "basement" patent or the group of patents, from which the path was derived.

[45]Instead of a time period t may represent any patent grouping variable, for example it is possible to calculate TDI for the patents arranged by the position levels.

[46]The TDI equation is written specifically for the HJT technological classification, however, it can be easily adjusted for another classification scheme.

allows us to examine how a particular technology diffuses over technological categories and subcategories as time passes or by a position level. Second, it allows us to compare the diffusion in one evolutionary path to the diffusion in another both by time and by position levels.

– Simple Single Value TDI. This simplified formula produces a single value for the whole evolutionary path and is computed the following way:

$$TDI_i = 1 - \sum_{j}^{n_j} \sum_{k}^{m_k} s_{ij}^2 s_{ik}^2 \tag{1.12}$$

1.9.2 Tables

Table 1.8: Regression Results Part 1

Path Significance

Dependent Variable	Observations				Levels			
Model	1	2	3	4	1	2	3	4
Ind. Variables:								
% of Coincidence[1]	14.066**	11.727**	10.959**	10.6**	2.772**	2.356**	2.197**	2.245**
Number of claims		0.106	0.105	0.074		0.017	0.016	0.015
Citations Received		1.167**	1.179**	1.223**		0.208**	0.21**	0.212**
Generality			-0.273	-0.234			-0.05	-0.048
Cites' Generality			1.982**	2.033**			0.407**	0.41**
Chemical				0.67				0.118*
Computer & Communications				1.811**				0.113**
Drugs & Medical				0.27				-0.02
Electrical & Electronic				0.579**				-0.021
Mechanical				0.131				0.022
Constant	0.548*	-0.482*	-0.688**	-0.966**	1.005**	0.826**	0.782**	0.716

Path Diffusion

Dependent Variable	% in home Category[2]				TDI			
Model	1	2	3	4	1	2	3	4
Ind. Variables:								
% of Coincidence[1]	-2.308**	-2.126**	-1.858**	-2.543**	0.489**	0.427**	0.355**	0.396**
Number of claims		-0.005	-0.013	-0.03		-0.005	-0.004	-0.003
Citations Received		-0.093**	-0.061*	-0.025		0.032**	0.029**	0.027**
Generality			-0.389**	-0.36**			0.03**	0.029**
Cites' Generality			-0.444**	-0.388**			0.157**	0.155**
Chemical				-0.148*				0.009
Computer & Communications				1.145**				-0.051**
Drugs & Medical				0.189†				-0.008
Electrical & Electronic				0.596**				-0.045**
Mechanical				0.061				-0.01
Constant	-0.044	0.035	0.155	0.307**	0.351**	0.336**	0.312**	0.304**

[1] Fraction of patents in the evolutionary path that that also belong to the "Microprocessor" GPT path
[2] Percentage of patents in the evolutionary path that have the same HJT category as the basement patent
† Significant at 10%
* Significant at 5%
** Significant at 1%
Red Statistically Insignificant

Table 1.9: Regression Results Part 2

Patent Generality and GPT-resemblance

Dependent Variable	Generality		Direct Citations' Generality		Degree Centrality[1]	
Model	1	2	1	2	1	2
Ind. Variables:						
% of Coincidence[2]	0.375**	0.396**	0.426**	0.436**	4.84e-06**	5.32e-06**
Chemical		0.005		0.003		-2.42e-07
Computer		-0.101*		-0.048†		-8.58e-07
Drugs & Medical		0.004		-0.005		2.14e-06**
Electrical & Electronic		0.001		0.003		-8e-07*
Mechanical		-0.01		-0.013		-4.29e-07
Constant	0.178**	0.174**	0.125**	0.125**	1.59e-06**	1.58e-06**

[1] Degree Centrality is computed according to Equation 1.9 in Appendix 1.9.1. Due to the large size of the network (1999 vintage data has 2,923,922 observations), the degree centrality values appear to be very small.The degree centrality of an average patent in the 1999 vintage data is 4.33e-06.

[2] A fraction of patents in the evolutionary path that that also belong to the "Microprocessor" GPT path

† Significant at 10%

* Significant at 5%

** Significant at 1%

Red Statistically Insignificant

1.9.3 What is a GPT?

There is no well-established universal GPT theory up to this day. Therefore, I combine information from various sources about GPTs in order to come up with a well-defined picture of the phenomenon. As a result, I provide a list of historically observed GPT-related stylized facts.

When we are looking at a GPT, we are looking for some invention or innovation that started with some narrow set of applications. After that, while being instantly improved, it gave rise to numerous "side effects". Those "side effects" are some particular applications in some particular sectors and sub-sectors of the economy (SPTs) that make the technology a GPT. We have historical examples of technologies with GPT potentials remaining SPTs. Nuclear power is the case that is sometimes used in the GPT literature to show that a technology that could have been a GPT for various reasons did not become one (or has not developed into a GPT yet).

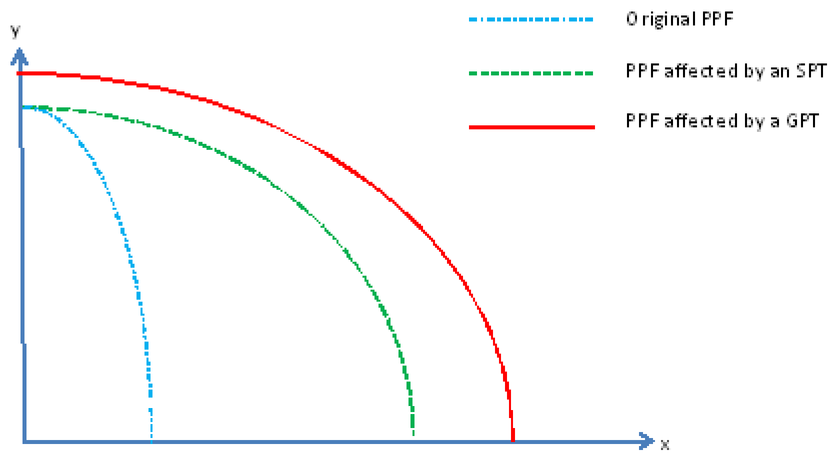
When one tries to identify a GPT, one is looking for some technology that first started with a narrow set of applications, which makes it look similar to an SPT at the early stages. Moreover, its potential may not be recognized even within its original industry.[47] This feature of a GPT makes it difficult to identify this technology at the immature stage of its life-cycle.

A technology that acquires numerous applications and becomes a GPT, completely changes the way of doing things. It may establish new industries, or destroy old ones. It may introduce new capital and consumer commodities and make certain old commodities obsolete. The whole economy after the development of such an invention or innovation changes its shape. Also, as it is stressed in [Lipsey et al., 2005], GPTs unlike regular innovations, not only decrease the amount of inputs necessary to produce the same amount of output but also open the door for the types of output that would not be possible to produce otherwise.

Figure 1.6(b) illustrates the difference between the effects of a applied innovation or a single purpose technology (SPT) and a fundamental innovation or a general purpose technology (GPT) on production possibilities in a simple two-sector economy.

[47] In [Bresnahan, 2007] the reader can find how computers developed into a GPT after they were applied to complete the tasks they were not initially designed for.

Figure 1.6: Effects of SPT and GPT on two-sector economy



Having all this information lets attempt to select an appropriate definition for the phenomenon. The term "General Purpose Technology" or a GPT was first introduced in [Bresnahan & Trajtenberg, 1995]. There are a few definitions for a GPT in the literature. [Lipsey et al., 2005]define GPT the following way:

A GPT is a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects.

This definition while being very close to the phenomenon still "creates a web too wide so it would miss some of the fundamental innovations and at the same time too narrow so it may catch some non GPTs". For example, many transportation technologies considered to be GPTs by the authors of [Lipsey et al., 2005] themselves have one particular use: move people and cargo from point A to point B. However, the appearance of these technologies had significant effects on every sector of the economy facilitating labour mobility and product delivery, making profitable something that otherwise would be unprofitable.

The term GPT, most often used in the literature, can be misleading due of the word "Technology" in it. According to the [Webster's Dictionary, 1997]: *Technology is the Application of scientific knowledge to serve man in industry, commerce, medicine, and other fields.* However, by technology we can imply relative broad applications of scientific knowledge, for example: *Information Technology* or *Nanotechnology* as well as more particular applications that serve one very specific purpose. One can call a particular device or group of devices *technology*.

A broader definitions of "technology" with multiple meanings are extracted from the following dictionaries: According to [The BBC English Dictionary, 1992]

Technology is:

1. *The activity or study of using scientific knowledge for practical purposes*

2. *Scientific equipment and methods used in particular area of activity*

While according to [The Oxford English Dictionary, 1989]:
Technology is:

1. *A discourse or treatise on an art or arts; the scientific study of the practical or industrial arts*

2. *Practical arts collectively*

3. *A particular practical or industrial art*

It is important to separate a technology in its "application of scientific knowledge to serve man in industry, commerce, medicine, and other fields" meaning from science on the one hand and from a capital tool (a particular practical or industrial art) on the other. GPTs are sometimes also called "core innovations" and "macro-inventions" and are the applications of new knowledge for multiple purposes that eventually reshape the economy.

Let steam power be an example to describe how the term "technology" can be misused and confused in the literature. [Lipsey et al., 2005] have a detailed discussion about classification of technological knowledge. They specifically mention in one part of their book that a steam engine and a car are pieces of capital rather than technologies. Nevertheless, in their list of GPTs, later in the book, they claim the "steam engine" to be one of the GPTs. It is important to understand that it is the "method of transforming thermal energy into mechanical energy via steam" rather then the steam engine was one of the major GPTs of the 19th century. Thus, a more appropriate name for that GPT would be the "steam power" instead of the "steam engine". Moreover, there are many kinds of steam engines that all work on steam power.

Nevertheless, it is often a convention to give the same name to a technology and the capital commodity the technology is mainly associated with. For example the technology called "the wheel" caused the creation of various capital commodities called "wheels" from tiny ones used in mechanical watches to huge ones used for haul trucks.

It is also often a source of confusion in the technological literature whether "electricity" is a technology and, therefore, a general purpose technology. Alternatively one may claim that electricity is a scientific field, while a dynamo machine is a GPT. The answer to that debate is that there is a section of physics that is called "electricity" as well as there is a technology of transforming mechanical and electrical energy one into another for various practical purposes that is also called "electricity". Even though it would be more consistent to call the GPT "electrical technology" and leave the term "electricity" to science. However, historically we call the electrical GPT "electricity".

The importance of the above discussion arises when one tries to look for a GPT in the data. It is important to remember that we are actually looking for the fundamental innovations that give a rise to a whole new group of applied innovations in various sectors of the economy. And not for some particular technology as a tool or capital commodity that was designed for a specific purpose.

For that purpose I am providing the following definition of a GPT: *A general purpose technology is an application of knowledge that had raised in some particular sector of economy and initially had much scope for improvement but eventually came to be widely used for multiple purposes, had many scientific and economic spillover effects.*

1.9.4 Stylized Facts About GPTs

After studying various sources in GPT literature as well as historical literature on the technologies I provide a list of historical stylized facts about GPTs. Note that there have not been two GPTs that would be alike in every aspect. Also note that most of the facts are based on historical observations rather than data analysis and therefore these facts should not be considered as universal. Nevertheless, the facts present common features observed in general purpose technologies and therefore are useful in describing the phenomenon.[48]

Stylized Facts of GPTs

- GPTs' have been observed through out the history of human civilization.
- GPTs' arrival may be exogenous or endogenous. [49]
- GPTs reshape the economy and the society.
- A potential GPT, in order to succeed requires additional research and investment into the microinventions derived from it.
- A potential GPT, in order to succeed requires a "fertilized soil" i. e. technology and innovation favouring social institutions.
- A potential GPT, in order to succeed needs to be technologically feasible. That is, the industries should be physically ready for switching to the new technology.
- A potential GPT, in order to succeed needs to be economically feasible. In simple words, it should outperform the existing technologies.
- Several GPTs may coexist at the same time and be widely applied.
- GPTs may complement or substitute each other. If a newer superior GPT provides the same service to the society as an old one it is likely that the newer GPT will eventually replace the old one like steam power and later electricity replaced water power. If the services are different, the technologies may complement each other.
- An active phase of research and secondary innovation for a modern GPT[50] is around 30 years.
- Some GPTs may play an important role in the production process after their active phase ends for literally thousands of years (for example, the wheel); while others may quickly loose their economic and technological significance, (for example steam power).
- Up to the twentieth century, GPTs would be discovered first inside the industry that would become their first and primary user, and only afterwards described by the formal science. Modern GPTs, on the other hand, usually get discovered in the labs and then find their applications.
- A typical GPT through its lifetime would be adopted in most of the existing industries and also would create new industries. It is also possible for a GPT to eliminate old industries that are not fit for its adoption and whose products would lose their demands due to the new substitutes created by the GPT.

[48] The facts were extracted from various literature sources, the most important ones are the following: [Rogers, 1983], [Lipsey et al., 2005], [Mokyr, 1990], [David, 1990], [Freeman & Soete, 1997], [Lipsey et al., 1998].

[49] There is no consensus in the literature about whether GPTs arrive because there is economic and/or technological demand for them or they are discovered by accident or out of scientific curiosity and then find their applications. Probably, both ways are possible.

[50] There is no strict convention in the literature about how to distinct modern GPTs from the older ones. At the same time there is a universal agreement that modern GPTs differ from their older counterparts in several ways. In my research I am mainly concern with the contemporaneous GPTs starting with electricity. Therefore, by modern GPTs I would imply GPTs that developed in the late 19th - 20th centuries.

- GPTs increase factors' productivity and stimulate economic growth in the long run. [51]
- The effect of GPTs on the economy in the short run is ambiguous. [52]
- GPTs diffuse (get adopted by various sectors of the economy) following an S-shaped logistic curve.[53]
- Early adopters of a new GPTs benefit more from it.
- The discovery of the microinventions for a particular GPT is subject to diminishing returns, while the arrival of new GPTs is not.
- There are no two identical GPTs that would have all features alike.

The task of characterizing GPTs is not easy due to uniqueness of each of the historically observed GPTs. However, in this section we provided a working definition of a GPT and a list of common features most of the historically observed GPTs share. In the next section we are going to dig deeper into the GPT theory and will attempt to classify all the historically observed GPTs.

[51] [Mokyr, 1990] clarifies that GPTs themselves have little effect on productivity and economy, the microinventions that follow the GPT and the increase in the research activity of these microinventions affect the economy.

[52] Some authors believe that a GPT in its early stage has little effect and microinventions arrive gradually reshaping the economy, therefore, no effect of a GPT is observed in the data. Others believe that there is a "productivity puzzle" that implies a negative initial effect on the productivity and output because the economy goes through a re-shaping process.

[53] This stylized fact is supposed to be relevant to all kinds of technologies, not only the general purpose ones. [Rogers, 1983].

Chapter 2

Effect of Technological Innovations on Hours Worked

2.1 Introduction

While there is more or less an established agreement about the effect of a permanent positive technological change on some economic measures, for example, on long run economic growth, the effect of such a change on the labour market in the short run is controversial. Theoretically, an improvement in technology allows production of more output with the same quantity of inputs. From that perspective, a technological improvement has two effects on labour. First, when labour productivity improves following a positive technology shock, it causes an increase in the demand for labour (the income effect). Second, since the technological change increases the relative cost of labour to capital, the demand for labour falls (the substitution effect).

Which of the two aforementioned effects is predominant in the aggregate macroeconomic data is the subject of disagreement. From the economic modeling perspective, the two major competing macroeconomic models, the Real Business Cycles (RBC) model and the New Keynesian Model (NKM) provide different results at least for the short run reaction of the labour market to a positive technological improvement.

The RBC model features perfect competition and the ability of producers to adjust prices and wages immediately in response to exogenous changes. The ability to adjust wages allows employers to compensate the increase in the relative price of labour. Therefore, in this model the substitution effect is dominated by the positive income effect. As a result, a positive technology shock leads to a rise in employment. On the other hand, if the employer was not able to adjust the wage immediately, a feature in the NKM, the relative cost of labour remains high, causing the substitution effect to dominate the income effect. In this case a positive technology shock actually leads to a fall in employment. Thus, while the RBC model predicts that working hours should increase after a positive technological shock, the NKM featuring staggered prices and wages implies a fall in hours worked following a technology shock.[1]

The objective of this chapter is to empirically identify the short term effects of a positive technology shock on aggregate hours worked. A vector autoregression model (VAR) with theoretically identified sign restrictions is used for that purpose. A feature that makes this study different from the previous attempts is the patent-based measure of aggregate technological process incorporated into the VAR model. The combination of the methodology and technological metric allows to identify a positive short run effect of a technology shock on aggregate hours worked in the recent US macroeconomic data.

[1] See [Gali, 1999] and [Rebelo, 2005] for a discussion.

Several researchers have attempted to resolve the issue empirically rather than theoretically. The typical empirical framework used by researchers in this literature is the VAR model. VARs are less theoretical but it is believed that they fit the data better. As a result, they can better identify the effects of technological shocks on the economy. Various specifications of VAR have been used to establish the effect of a technological shock on employment. The results from the VAR literature are also mixed. For example, [Gali, 1999] and [Francis & Ramey, 2002] identify the negative effect in accordance with the NKM. On the other hand, [Peersman & Straub, 2004] and [Debola & Neri, 2007] find that working hours respond positively to a technological improvement.

In this study, a DSGE model that features two types of workers (skilled and unskilled) and two types of technological improvements (skill-specific and general) is introduced. The model may imply both an increase and a decrease in aggregate hours depending on the type of technology improvement and assumptions made regarding the proportion of skilled workers in the economy. However, the effect of both types of shocks on some of the other variables presented in the model is identical. Such a model provides the necessary restrictions and identification to be imposed on the data in order to uniquely identify a skill-specific and a skill-neutral technology shocks and observe their effects on aggregate hours worked.

In general, the problem with the ambiguity of the responses of the employment variables to the technological innovations arises because it is difficult not only to identify a technology shock but also to select an appropriate measure of technology. Identifying correctly a technological shock and being able to select a measure of technology is the key to understanding the response of the economy to this type of shock.

It is important to know how technological innovations affect hours worked. First of all this knowledge would be helpful to policy makers for making appropriate decisions regarding the labour market.. Second, this knowledge provides insights on the importance of technological innovations of economic cycles. Third, it would allow economists to select models that explain more correctly the interaction of variables and their effects on each other.

2.1.1 The Novelties of This Study: How Technology Is Measured

Most of the controversies in the VAR literature on technology shocks occur when one is choosing or identifying the measure of technology. The technological measures that are most commonly used in the VAR literature are the total factor productivity (TFP) and labour productivity. Both of these variables are often identified from the Solow residuals and are usually computed as a natural logarithm of output per hour, as in [Gali, 1999] and [Debola & Neri, 2007]. However, many researchers have argued that TFP as measured by the Solow residuals is not a good measure of technology shocks.

[Rebelo, 2005] indicates that TFP measured using Solow Residuals can be predicted by military spending and monetary policy. Both of the predictors have little relevance to technology. According to [Alexopoulos & Cohen, 2010], both TFP and output per hour are affected by various factors other than technological change, while the elimination of the non-technological effects from the TFP measure data is costly and not always successful. Finally, [Lipsey & Carlaw, 2004] have a detailed study of various measures of TFP and conclude that "changes in TFP do not represent technological changes but rather measure the externalities associated with the technological change". By the externalities, the authors mean the unpaid effects of technological change, for example when an invention in industry A improves productivity in industry B.[2]

One of the novelties I introduce in this paper is the measure of technology used in the VAR model. I use the quantity of patent claims as the measure for technological innovations. Patent claims represent novelties of a particular patent that are subject to legal protection. In general, a patent is comprised of "previous art" - the pre-existing technological knowledge and the novelties that are claimed. Therefore,

[2]The alternative TFP-based measure of technology will be used for the robustness check. See section 3.4.

in theory, each of the claims represents some technological improvement that previously did not exist. In practice, there is an incentive to claim more in order to receive more payments, however, a claim that, when examined, appears to encompass some previous art may be rejected causing no revenue and thus providing the incentive to not over-claim.

There are some drawbacks associated with using patent variables as indicators of technological knowledge. [Griliches, 1990] discusses that changes in patenting activity may not only be supply-driven (caused by technological improvements) but also demand-driven (profit motivated and caused by economic conditions). Patent-related measures of technology find both critiques and supporters in the literature. In section 2.3 of the chapter I will take a more detailed look at the literature that discusses patent measures of technology in general and patent claims in particular in order to justify my selection.

2.1.2 The Novelties of This Study: How Technology is Identified in VARs

Even if we can actually choose a variable to measure technology shocks, identifying the effects of this variable on the other variables is another issue that has to be addressed. The difficulty of identifying technology shocks comes from the fact that these shocks are identical to each other. Moreover, the economy experience many other shocks besides the technology shocks and it is difficult to distinguish the effects of technology shocks from the effects of the other shocks [Bocola et al., 2011].

The literature employing VAR has used various techniques to identify a technology shock. Some of the authors identify technological shocks in the data by using short run restrictions on their VAR. Short run restrictions require a researcher to assume that a certain shock will have no immediate effect on a certain variable. For example [Alexopoulos & Cohen, 2010] assume that technological improvements affect output and TFP with a one year lag.

Other authors implement long run restrictions. The long run restrictions require a researcher to assume that a certain shock will have only transitory (non-permanent) effect on a certain variable and therefore, in the long run this variable will return to its original value. For example, [Gali, 1999] distinguishes between technological shocks that have permanent long run effects on both working hours and productivity, and non-technological shocks that have permanent long run effects on working hours only.

Finally, there is a so called "agnostic" method based on sign restrictions. This method implies making the least assumptions possible and letting the data speak for itself. When imposing a sign restriction, a researcher makes an assumption only about a direction a certain shock would cause a certain variable to move. The sign restriction method was introduced and successfully implemented in the monetary policy literature to identify a monetary shock in [Canova & De Nicolo, 2002] and [Uhlig, 2005]. It was then adopted to identify technological shock in [Peersman & Straub, 2004] and [Debola & Neri, 2007]. In terms of the structural VAR identification techniques, I will follow the so called "agnostic" method based on sign restrictions.[3]

Sign restrictions have several appeals. It requires less controversial and arguable assumptions and it produces robust results. Different theoretical assumptions on the VAR system may cause different results. In terms of robustness, long run restrictions may not always be robust to the specification of the variables in levels or in differences [Rebelo, 2005]. At the same time, the sign restrictions implemented for identification of a technological shock in [Peersman & Straub, 2004] produce robust results no matter whether the variables are specified in levels or in the first differences.

The agnostic sign restriction method requires specifying whether there is a positive or a negative response to a shock from certain variables rather than specifying whether or not there is an effect to a shock. Thus, it implies some set of assumptions of the following type: a shock i causes variable j not to increase/decrease for a length of time k. Obviously, since the objective is to retrieve the direction of the response of the labour market variable to the technological shock, I will avoid imposing any sign

[3] In order to check the robustness of the results received using the sign restriction method, I will also implement the short and long run restrictions to the data. The results will be presented in Section 3.4.

restrictions on the labour market variable. More details on the VAR model with sign restrictions and its application to technology shock identification are presented in Section 2.5.

2.1.3 Structure

The paper will be structured as follows: Section 2.2 will present some important past contributions to the debate about the effect of technological innovation on hours worked. Section 2.3 will include a discussion about the choice of variables selected for this research. Most of the attention in section 2.3 will be devoted to the choice of the patent variables over the alternative measures of technology since the choice of the correct technological measure is of importance when one tries to observe the effect of technology on other variables.

Section 2.4 is devoted to the choice of sign restriction and therefore to the technology shock identification.In order to do that, a macroeconomic model with skill heterogeneity will be introduced. The model will feature two types of technology shocks causing different effects on aggregate hours worked and will provide a framework for the restrictions that uniquely identify these shocks.

In Section 2.5, I will present the results of the model while in Section 3.4, I provide the robustness checks of the results derived in Section 2.5. I perform several robustness checks. I change the technology measures, the specifications of the VAR (levels versus first differences) and I use a more conventional short run and long run identification strategy instead of the sign restriction. Section 3.5 concludes the chapter.

2.2 Literature Review

The two competing theoretical explanations of the effect of technology shock on hours worked are found in the RBC and NKM models. One of the conclusions of the RBC macroeconomic model is that a positive technology shock raises hours worked (See RBC models in [King et al., 1988] and [King & Rebelo, 2000] for an example). An alternative class of macroeconomic models, the New-Keynesian models that feature price and/or wage stickiness predict a different outcome from a technological shock: a positive technology shock has a negative short-run effect on hours worked according to these models (See the model in [Gali, 1999] or the model [Gali et al., 2011], that features the effect of an investment-neutral productivity shock).

Empirically, many authors starting with [Gali, 1999] claim that the results produced by the RBC models are inconsistent with the data. Gali uses a structural VAR with long run restrictions[4] to show that the estimated conditional correlations of hours and productivity are negative and the impulse responses show a persistent decline of hours in response to a positive technology shock. These results, while being in contradiction to the RBC model, are in accordance with the Neo-Keynesian staggered prices/wages models where price rigidities imply that aggregate demand cannot change immediately, which makes firms contract employment after an exogenous increase in productivity [Peersman & Straub, 2004].

Empirical findings in [Gali, 1999] found both supporters and critics. There are researchers who also find negative effects of technology shocks on working hours using long run VAR restrictions (see [Francis & Ramey, 2002] for example). There are also researchers who question the validity of the empirical results found in [Gali, 1999]. [Rebelo, 2005] discusses the literature that finds Gali's results to be a product of misspecification and not to be robust to the choice of the variables in the VAR system being in levels or in first difference. Also, the long run restriction VAR identification scheme that was used in [Gali, 1999] was criticized in [Chari et al., 2007]. [Chari et al., 2007] generate artificial data by running an RBC model. Then they apply the long run VAR restrictions introduced in [Gali, 1999] on this

[4] Structural VAR methodology with long run restrictions implies that a researcher can predetermine a shock to one of the endogenous variables in the VAR to have no permanent effect on another endogenous variable(s).

data. They use Gali's identification scheme and show that a positive technology shock negatively affects working hours in the short run, while the true data generating process from the underlying RBC model implied the opposite.

The results of [Chari et al., 2007] are crucial in the discussion since they show that an imposition of long run restrictions on the data does not necessarily lead to the uncovering true effects of shocks on various variables. The effect uncovered with these restrictions in [Chari et al., 2007] contradicts the known in advance data generating process. Therefore, when one implements long run restrictions on the actual data, the results may also contradict the unknown or unobserved true data generating process. At the same time it is necessary to mention that the methodology of the [Chari et al., 2007] paper is not free from criticism itself. The authors use a structural VAR with a finite number of lags to analyze the data derived from the RBC model with an infinite number of lags.

Empirical results, different to those obtained in [Gali, 1999] are derived when sign restrictions instead of long run restrictions are applied to the VAR system. For example, [Peersman & Straub, 2004] and [Debola & Neri, 2007] find evidence of a positive response of working hours and employment to the positive technology shock. What is necessary to note, is that [Peersman & Straub, 2004] works with European data, while the vast majority of the literature considers US data. For example, [Debola & Neri, 2007] working with the US data finds that hours worked are likely to increase. However, according to [Debola & Neri, 2007], technology shocks leave unexplained most of the variation in hours worked. Thus, the empirical effects of the technology shocks on working hours are still open to discussion.

[Peersman & Straub, 2004] and [Debola & Neri, 2007] select sign restrictions in a similar way. [Peersman & Straub, 2004] introduce an RBC model and a New Keynesian model and investigate on which variables the effect of a technology shock is common in both models. These common effects (positive or negative) determine the sign restrictions on the variables in their VAR. The approach in [Debola & Neri, 2007] is to introduce a DSGE model that is New Keynesian in nature, but could reproduce the RBC results when you switch off certain model parameters (assign certain constants to be equal to zero). The rest of the approach is to identify the effects of a technology shock on certain variable that would be common no matter whether the switching parameters are "on" or "off" and therefore to determine the sign restrictions.

One of the ways the literature is trying to deal with the nondeterminacy of the effect of technology shock on employment and working hours is to distinguish between two types of the technology shocks: investment specific shocks, and investment neutral shocks. [Rebelo, 2005] calls an investment specific technological change to be a natural alternative to the "regular" RBC-model type technological change. The classical RBC-type technological shock, the investment-neutral shock, improves both labour and existing capital productivity. Therefore, this shock increases quantity of hours worked. On the other hand, the investment-specific shock makes new capital less expensive while leaving labour and old capital's productivity unchanged, therefore it causes a substitution towards capital augmenting technologies and therefore makes some labour services obsolete.

Although the investment-specific shock was called an alternative to the "regular" technology shock, most of the modern macroeconomic New-Keynesian DSGE models use both shocks (see [Fernandez-Villaverde, 2009]). Some of the empirical studies have also found that investment-specific and investment-neutral technological shocks produce different effects on working hours. [Michelacci & Lopez-Salido, 2007] provide a simple, Solow-type model according to which they describe the theoretical effects of the investment-neutral and specific shocks on the working hours. Then, the authors construct a VAR system with a long-run restriction identification scheme that supports their theoretical conclusions, which are that the results of the two types of technological shocks are different.

[Gali et al., 2011] use a more sophisticated DSGE model than the model in [Michelacci & Lopez-Salido, 2007]. The authors estimate the model on US data. Along with the other findings, the authors also derive the effects of the investment specific and neutral shocks to be different, however, the effect of the investment-neutral shock on working hours is found to be weak.

The question is how to measure investment specific and investment neutral shocks? The literature deals with this question the following way: The investment specific technological shock can be measured using the relative price of investment goods in terms of consumption goods ([Rebelo, 2005]). For instance, [Michelacci & Lopez-Salido, 2007] use the price of a quality-adjusted unit of new equipment in terms of the price of a consumption unit in their analysis.

For the "regular" or investment-neutral technological shocks, the literature often uses productivity measures based on [Solow, 1957] growth accounting approach. For example, [Michelacci & Lopez-Salido, 2007] and [Debola & Neri, 2007] use output per working hour, while [Gali, 1999] uses the log of GDP minus the log of non-farm labour hours. These measures are criticized because they contain non-technological effects [Rebelo, 2005]. Also, [Lipsey & Carlaw, 2004] claim that these measures capture only the externality effect of the technological shock. Moreover, while labour and total factor productivity increases are associated with technological progress, that is positive technological shocks, the decreases in factor productivity are more difficult to explain by technological regress ([Rebelo, 2005]) since historically we do not observe technological regress very often.

It is also necessary to mention studies that use patent application based technological indicators to study the effect of technological shocks on labour productivity. One of the pioneers in establishing the importance of the relation between patent counts to productivity on industry level is [Lach, 1995]. [Shea, 1998] uses annual patent application and R&D spendings as technological indicators. He constructs tri-variate VARs with quantity of inputs, total factor productivity and one of technological measures (patents or R&D). [Shea, 1998] uses a short run identification scheme,[5] he puts technological measures last into the VAR to reflect the notion that patent applications and R&D spendings may be demand driven[6]. The author also uses time-dummies to reflect patent policy changes. [Shea, 1998] finds that technology shocks tend to increase inputs, especially labour in the short run, but decrease inputs in the long run.

[Christiansen, 2008] follows [Shea, 1998] in using annual patent applications as a measure of technological innovation. Also, like [Shea, 1998], [Christiansen, 2008] imposes short run Cholesky restrictions on the bi-variate VAR that includes patent applications as the technological measure and output per hour as the productivity measure. However, unlike [Shea, 1998], [Christiansen, 2008] places the technology measure first thus, restricting patents to react to changes in productivity with a lag. [Christiansen, 2008] shows that productivity may decrease in the short run due to a technology shock. Thus, the patent data technology measure with short run restriction VAR produces results that are not robust and depends on the order of variables.

Finally, most of the aforementioned studies treat employment uniformly using aggregate measures of working hours or aggregate employment rates. However, in theory a technological improvement affects skilled and unskilled labour differently. A technological improvement raises the demand for the skilled labour and reduces the amount of unskilled workers needed in the production at the same time. [Allen, 1996] finds that technology variables account for the 30% increase in the wage gap between high school and college graduate in the 1980s. When a technological change is of general nature, rather than an industry or task specific, the wage inequality between skilled and unskilled workers is expected to be even wider [Aghion et al., 2002]. Thus, there will always be "winners" and "losers" from the introduction of a technological improvement within the labour market. The goal of an aggregate study is to observe

[5] The short run identification scheme requires a researcher, after estimating reduced form VAR, perform Cholesky decomposition of the estimate variance-covariance matrix $\hat{\Sigma}$. The resulting matrix P is lower triangular. This matrix is used to uncover structural shocks given the reduced form shocks. The reduced form VAR is transformed into the moving average form. In order to identify structural shocks u_t, reduced form shocks e_t are multiplied by P^{-1}. Note that the order of the variables in VAR matters. The very first variable's shock will have its corresponding row in the Cholesky matrix P with the first coefficient and the rest of zeros that would imply that the first variable in the system is immediately affected only by its own shock and nothing else. The last variable in the VAR system, on the other hand, will have its corresponding row in the Cholesky matrix P full of coefficients implying that this variable is immediately affected by all the shocks from all the endogenous variables.

[6] See [Griliches, 1990] for a discussion.

whether there are more of the former or latter and therefore to establish the overall effect of the innovation on the labour market.

The current research adds to the literature by using a unique combination of the patent-based measure with the sign restriction VAR methodology and distinguishes between general and skill specific technology shocks. Each of these components requires some attention and justification of its use. Therefore in the next two sections, I present the measure and the methodology respectively along with the reasoning for my choices.

2.3 The Data: Choice of the Variables

One of the widespread approaches to measure technological innovation is to construct a so-called Solow residual. This method implies that all the empirical fluctuations in output that cannot be associated with the fluctuations in labour and capital are counted as changes driven by the improvement in factors' productivity. Thus technological innovations, in this case, would be recorded indirectly from observing the changes in economic variables. This technique has a certain amount of criticism.

[Rebelo, 2005] indicates that total factor productivity (TFP) measured using Solow Residuals can be predicted by military spending and monetary policy. Both of the predictors have little relevance to technology. According to [Alexopoulos & Cohen, 2010], both TFP and output per hour variables are affected by various factors other than the technological change. The elimination of the non-technological effects (as it was done in [Basu et al., 2006]) is costly and not always successful.[7] Finally, [Lipsey & Carlaw, 2004] have a detailed study of various measures of TFP and conclude that "changes in TFP do not represent technological changes but rather measure the externalities associated with the technological change".

An alternative approach to using the Solow Residual measures of technological improvements would be to use one of the direct measures of technology. Instead of waiting for technology to affect certain economic variables and then record these effects, numerous researchers look at the variables that directly react to technological innovations. Two main kinds of data that directly reacts to technological innovations are patent and R&D data.[8]

Both R&D and patent based measures are approximations for some true uncovered measure of the quantity of technological improvements. Both of these measures have advantages and disadvantages. The advantage of R&D measures is the dollar value assigned to the innovation. On the other hand, patent data has important advantages over R&D measures. As summarized in [Lanjouw et al., 1998], patent data contains more information and covers a wider scope of inventions and innovations. While trustworthy R&D data is available mostly for large corporations , it is not available for small firms and individual inventors.[9]

In my research, I am using a patent-derived measure rather than R&D measure for the following reasons. Patents represent successful outputs of the research and development process while R&D expenditures measure the inputs. Moreover, patent data is more aggregate and covers a wider scope of inventions [Lanjouw et al., 1998]. [Lin & Chen, 2005] note that scientists and researchers often use the number of patents as an R&D performance indicator since R&D inputs do not always lead to successful technological improvements. Therefore, a patent data measure is a more appropriate measure of technology for an aggregate study.

One big disadvantage of the patent data in comparison to the R&D data is the unavailability of a certain economic value attached to every patent. There are different ways to weigh and measure the

[7] However, we use the utilization adjusted TFP as an alternative measure of technology in the Robustness section.

[8] It is possible to find other variables that react directly to technological innovations. For example, [Alexopoulos & Cohen, 2010] looks at the new book titles, while [Archibugi & Pianta, 1996] use innovation surveys.

[9] Patent and R&D measures are not necessarily always substitutes for one another. They can also be used as complements. [Lanjouw & Schankerman, 2004] use ratio of patents to R&D as a measure of research productivity.

economic and/or technological significance of a given patent. In the next section, I discuss the "number of patent claims" indicator of technological improvement that will be used in my research as a way of improving the significance of the patent data. Using patent claims instead of patent counts is one of the common ways to reduce patent counts' skewness and improve the accuracy [Archibugi & Pianta, 1996].

Since the true measure of technological innovation is unknown, there is no perfect approximation for it. Number of patent claims is also an imperfect measure of innovation. In order to justify the "number of patent claims" measure, it is first necessary to justify the overall use of the patent data rather than other types of data or a Solow residual in general. After that, it is possible to proceed with justifying the use of the "claims" parameter instead of some other patent derived parameter.

2.3.1 Advantages and Disadvantages of Patent Data

Below is a brief comparison of patent data measure of technology to alternative possible measures of technology:

1. Patent data is a superior measure of technology than the indirect Solow residual type measures because it does not capture the effects of technology-unrelated events. Even though it is still difficult and sometimes impossible to establish whether patents are mainly supply or demand drive i.e. whether quantity of patents reflects the economic situation (demand) or technological opportunities (supply), [Griliches, 1990] argues that the "news" component in the patent statistics reflects technological opportunities. Therefore bursts and declines in the patent application activity reflect changes in the technological opportunities rather than current economic situation.[10]

2. Patent data outperforms R&D expenditure (the main alternative direct technology measure) in several ways. It accounts for a greater scope of technological improvements. [Griliches, 1990] notes that the importance of R&D rises with the size of the firm, while not all the patents reflect the R&D activity. Also R&D data represents input into the innovative process rather than a successful output. For that reason patents are sometimes used to evaluate R&D performance (see [Lin & Chen, 2005]). A disadvantage of the patent data is that patents differ in economic and technological value.[11] In order to improve the measured value of the inventions and reduce skewness of raw patent count data, the number of patent claims measure is used.

3. Patent data may be preferred to other more "exotic" direct measures of technological activity such as scientific citations databases, number of publications, or innovation surveys due to the limited availability of the latter [Lanjouw et al., 1998].

4. Even though patents are of different economic and technological significance every patent contains at least a "minimum quantum of invention" [Griliches, 1990]. Moreover, since patent protection is costly, those invention that are of some economic significance, the inventions that promise certain economic return are more likely to be patented [Archibugi & Pianta, 1996]. It has been noted that the vast majority of significant inventions since the first industrial revolution were patented at some point [Rothwell et al., 2013].

2.3.2 Justification for the Use of the Number of Claims

In the following segment, we present a brief comparison of the number of patent claims variable to the raw patent count and alternative patent-derived measures of technology.

[10][Kortum & Lerner, 1999] and [Rafiquzzaman & Whewell, 1998] test whether the jumps in the patent activity in the USA and Canada respectively are due to "fertile patent policy" or "fertile technology" both studies come to the conclusion that the patent activity jumps are due to "fertile technology".

[11]Another disadvantage is that not all the inventions are patented. However, not all inventions are results of organized R&D thus the same critique applies to the R&D data.

1. It is commonly known that not all patents are of the same value. The value of a particular patent is difficult to measure. Patents differ in economic value as well as technological value. Also the private value of a patent may be different from the social value of the invention. The use of patent claims as a measure of technological innovation was first proposed by [Tong & Frame, 1994]. The authors found that the number of patent claims is a superior measure of technology than raw patent counts. Each claim, according to [Tong & Frame, 1994], represents a distinct inventive contribution, while patents are bundles of inventions. It was found that patent claims count performed well as an indicator of national technology capacity. Higher correlation of patent claims rather than patent counts with other technological variables was found.

2. Claims in a patent define the scope of legal protection. Any patented technology can be roughly split into two parts: "prior art" and the novel part. The novel part is the subject of claims. Claims themselves are also split into principal and subordinate. Principal claims define essential novel features, subordinate describe detailed features of the innovation [Lanjouw & Schankerman, 2004]. Unfortunately, there is no data available that would count principle and subordinate claims differently.[12] Nevertheless, secondary or subordinate claims are not completely useless. They may be less valuable technologically but still determine the economic value of a patent. According to [Beaudry & Schiffauerova, 2011], higher number of claims indicates a broader scope of application and thus of greater profitability of a patent. Also, according to [Rothwell et al., 2013], claims define patents invention and what is legally enforceable, thus, patents with multiple inventions would have multiple claims.

3. An important issue is the difference in the number of claims across technological fields. A solution to this problem was suggested in [van Zeebroeck & van Pottelsberghe de la Potterie, 2011a]. The authors suggest dividing the number of claims in each patent document by the median number of claims in the technological field and the year of application in order to capture the time and field effects.

4. [van Zeebroeck & van Pottelsberghe de la Potterie, 2011b] compare 50 different patent value indicators. The indicators are roughly grouped into four different groups: patent characteristics, ownership, insider information, and filing strategy. Each of the groups is also split into sub-groups. It is necessary to note that not all the indicators are easily available for each patent. The indicators from the "insider information" group come from surveys and thus would not be available for an aggregate study. Other indicators, for example "firm's technological strength" and "patenting motives" is also subjective and not freely available on the aggregate level. According to the comparative study of [van Zeebroeck & van Pottelsberghe de la Potterie, 2011b], the indicators from the "filling strategy group" that include the number of claims indicator, demonstrated the best results as a group. The number of patent claims was the second best indicator after the number of forward citations from the "patent characteristic" group.

5. There are two important problems associated with forward citations but irrelevant to the number of claims or backward citations. The first problem is that unlike the number of claims for a given patent that is established once and for all, the number of forward citations increases over time as new patents are granted that may refer to a given patent. The second problem is derived from the first one. The number of forward citations is subject to truncation of the data. Patents that are closer to the end of the data set (in terms of time) have less chance to receive a citation than the early patents.[13]

[12] If there were such data, number of principal claims could be a better measure of technological value.

[13] See [Hall et al., 2001b] and [Gress, 2010] for a discussion of the citation inflation and truncation problems associated with forward citations.

6. Another study by [Lanjouw & Schankerman, 2004] considered various patent measures that would indicate the technological quality of a given patent and found that, in six out of seven technological fields, number of claims is the most important patent quality indicator.[14]

Based on a detailed literature review summarized above, we establish that the number of patent claims measure is expected to suit its purpose well and adequately approximate the unobserved technological process. Nevertheless, as a robustness exercise, an alternative direct indirect measure of technology will also be used in the VAR model. Besides that, the third chapter of the dissertation will present a methodology that allows to overcome the choice of the "best" measure of technology and incorporate all the information available.

2.3.3 The other variables

The choice of "regular" economic variables examined in this study is less controversial and therefore does not require as much attention as the choice of the technology measure. Aggregate weekly hours of production and non-supervisory employees was taken from the U.S. Bureau of Labor Statistics as the main response variable of interest. OECD data on real GDP and prices in the U.S. is used.

In order to distinguish between employment and wages of skilled and unskilled workers, U.S. Census Bureau's historical time series data on the education level of the labour force was used.

Figure 2.1 plots skilled workers' wage premiums and the share of skilled workers in the labour force in the United States. One may wonder who do we consider a "skilled" worker. Historically, the definition of a skilled worker has changed over time so that higher and higher standards are expected of the skilled employee. At the beginning of the 20th century, a person with a high school diploma could be considered skilled. Today we may consider a Bachelor's degree a prerequisite to be a skilled worker.[15] Thus, I adopt two schemes to identify a skilled worker. According to scheme A (top panel on Figure 2.1) workers with a Bachelor's degree and up are considered skilled.[16] In the middle panel, a more conventional way that counts workers with a college diploma as skilled workers is presented. Both panels demonstrate a similar pattern of an increase in the wage premium of skilled workers over time as well as an increase in the share of skilled workers in the production process. The lower panel of Figure 2.1 represents scheme C according to which a high school diploma determines whether a worker is skilled or unskilled. The wage premium according to this scheme is even more drastic. However, due to the extremely low proportion of scheme C unskilled workers in the economy, this scheme will not be further used or discussed. However, one can speculate that the conclusions of this chapter would be even stronger if scheme C was implemented. Overall, the data features depicted in Figure 2.1 will be important in identifying technology shocks and explaining their effects on aggregate hours worked.

2.4 Theoretical Shock Identification

In my analysis, I identify two types of technology shocks: a general productivity shock and a skilled labour specific productivity technology shock. [Katz & Margo, 2013] note that the at the time of the industrial revolution technological improvements were mostly de-skilling the labour force. Very skilled artisans were replaced by unskilled labourers doing simple mechanical tasks in the production chain. Modern technological innovations, on the other hand are skilling, and tend to increase the share of skilled workers in the workforce as well as leading to wage premiums for skilled workers.[17] Further, several

[14] Only in the "Drugs and Medical" field the number of citation received technological measure, that was presented in [Hall et al., 2001a], outperformed the number of claims measure.

[15] [Goldin & Katz, 2007] note that high school graduate of 1910's were more of an elite class than college graduates today.

[16] It would be more correct to call skilled workers "high-skilled" and unskilled workers "low-skilled" since both high school and college graduates have some of their own sets of skills.

[17] See [Aghion et al., 2002] for the discussion.

studies indicate a skill-biased technological improvements in the last decades that lead to an increase of the wage premium and in the participation rate of skilled workers.[18] We also observe these trends in Figure 2.1. Thus, the objective is to identify and compare a skill-biased technology shock with a general technology shock, and then examine the effects of these shocks in the data.

It is possible that the literature that looks for the effects of a technology shock on hours worked often finds contradictory results since it does not distinguish these different kinds of technology shocks. In order to identify skill-specific and skill-neutral (aggregate) technology shocks, I use a DSGE model with skill heterogeneity that is similar to the model presented in [Prasad, 1996] with certain additional features such as a preference shock and money in utility. The changes are introduced in order to accommodate the model for the tasks of deriving theoretical sign restrictions on economic variables that identify technology shocks. The objective of the model is to uniquely identify two types of technology shocks: a general productivity improving shock and a skilled worker productivity improving shock. The model provides sets of sign restrictions that let us distinguish technology shocks of one kind from the other, as well as technology shocks from monetary and preference shocks.

2.4.1 The Model

The household sector consists of a continuum of identical, infinitely lived two-person households, skilled and unskilled. Households in the model derive utility from consumption and two types of leisure: skilled labour leisure and unskilled labour leisure as in [Prasad, 1996]. However, unlike [Prasad, 1996], in this model households also derive utility from holding real balances and are subject to a stochastic preference shock.[19]

$$U(c_t, l_{st}, l_{ut}) = \ln(c_t - z_t) + D\ln\frac{m_t}{P_t} + \theta(\eta \ln l_{st} + (1-\eta)\ln l_{ut}) \tag{2.1}$$

Where the preference shock[20] follows an AR1 process: $z_t = \rho_z z_{t-1} + \epsilon_z t$; l_{st} and l_{ut} are skilled and unskilled workers' leisures respectively.

To reflect the notion that fluctuations in the aggregate hours worked are caused by fluctuations in employment rather than in the number of hours of those who are employed, the model restricts the labour day to be $0 < h < 1$ hours long. Therefore agents either work h hours or 0 hours each periods. In order to convexify the space, we let agents optimize over lotteries that assign certain probability of working h hours next period.[21] The assumption of the fixed labour day provides full insurance for the households over their consumption as it was shown in [Rogerson, 1988]. Therefore it significantly simplifies the aggregation procedure since all the households consume the same quantity of the consumer good c_t in the period t. Assigning the fixed labour day, lets us convert 2.1 into:

$$U(c_t, N_{st}, N_{ut}) = \ln(c_t - z_t) + D\ln\frac{m_t}{P_t} - \eta a N_{st} - (1-\eta)aN_{ut} \tag{2.2}$$

Where $a = -\theta\frac{\ln 1-h}{h}$; N_{st} is the number of hours worked by skilled workers, and N_{ut} is the number of hours worked by unskilled workers.

[18] See the studies by [Goldin & Katz, 2007], and [Marquis et al., 2011] for example.

[19] The demand shock is taken from the [Wen, 2006] model, while the money in the utility technique is described in details in [McCandless, 2008].

[20] Besides skill-biased and general technology shocks the model features a preference shock and a monetary shock. Due to the fact that number of patents can be both demand and supply driven, when we examine a technological improvement using the patent data, we need to exclude the demand driven case. Therefore, I introduce the demand shock to be able to assign unique restrictions that identify a technology shock, but not a demand shock. The monetary shock in the flexible price model is more redundant than other shocks due to money neutrality. However, the price level is one of the key empirical variables used for the shocks' identification. Money is introduced in order to be able to define the aggregate price level in a simple manner.

[21] See the models in [Hansen, 1985], [Rogerson, 1988], or [Prasad, 1996] for more details.

Firms use skilled and unskilled labour along with capital to produce consumption goods:

$$Y_t = A_t K_t^{1-\alpha} N_t^{\alpha} \tag{2.3}$$

Where the labour component is of the CES form:

$$N_t = (B_t\omega(N_{st} - \frac{d}{2}(N_{st} - N_{st-1})^2)^{1-\upsilon} + N_{ut}^{1-\upsilon})^{\frac{1}{1-\upsilon}} \tag{2.4}$$

The above production function has the following features:

1. Two technological processes: $\ln A_{t+1} = \rho_A \ln A_t + \epsilon_{At+1}$ and $\ln B_{t+1} = \rho_B \ln B_t + \epsilon_{Bt+1}$ represent total factor productivity improving technology and skilled labour productivity improving technology respectively.
2. Term $\omega > 1$ makes skill labour more productive under any circumstances (with or without skilled labour productivity technological innovation).
3. Term $\frac{d}{2}(N_{st} - N_{st-1})^2$ incorporates skilled labour adjustment (training) cost. Without this term the ratio of skilled to unskilled workers would remain constant.

The output in the model is either consumed or invested:

$$Y_t = C_t + I_t \tag{2.5}$$

The capital stock evolves according to the following equation:

$$K_{t+1} = (1-\delta)K_t + I_t \tag{2.6}$$

The household maximization problem is the following:
The utility function determined in 2.2 is maximized by infinitely lived households subject to a resource constraint.
$\max_{E_t \sum_{t=0}^{\infty}[C_t, m_t, N_{st}, N_{ut}, K_{t+1}]} U_t = E_t \sum_{t=0}^{\infty} \beta^t [\ln(c_t - z_t) + D \ln \frac{m_t}{P_t} - \eta a N_{st} - (1-\eta) a N_{ut}]$

Subject to:

$C_t + K_{t+1} + \frac{m_t}{P_t} = (1-\delta)K_t + w_{st} h N_{st} + w_{ut} h N_{ut} + r_t K_t + \frac{m_{t-1}}{P_t} + (g_t - 1)\frac{M_{t-1}}{P_t} + \chi_t.$

Where r_t is the interest rate household receives from renting firms capital, w_{st} and w_{ut} are skilled and unskilled wages, and χ_t is the firms' profit, $(g_t - 1)\frac{M_{t-1}}{P_t}$ is a lump sum transfer from the monetary authority.

By solving the household problem we derive the following equilibrium conditions (in aggregate terms):

$$\frac{1}{C_t - z_t} = \beta E_t [\frac{1}{C_{t+1} - z_{t+1}}[1 - \delta + r_t]] \tag{2.7}$$

$$w_{st} = \frac{\eta a}{h}(C_t - z_t) \tag{2.8}$$

$$w_{ut} = \frac{(1-\eta)a}{h}(C_t - z_t) \tag{2.9}$$

$$\frac{1}{C_t - z_t} = D\frac{P_t}{m_t} + \beta E_t [\frac{1}{C_{t+1} - z_{t+1}} \frac{P_t}{P_{t+1}}] \tag{2.10}$$

Equation 2.7 represents the inter-temporal Euler consumption equation. The equation shows how households value goods received in the future in terms of goods received in the current period. Equations 2.8 and 2.9 represent the optimal labour-leisure choice for the skilled and unskilled workers respectively. The equations allow to determine equilibrium consumption, labour, leisure, and wages for the both types of workers. The equilibrium wage rates are evaluated in terms of consumption they can afford. Equation 2.10 represents the optimal allocation of real balances. The real balances are determined by the current and inflation-adjusted future consumption. Note that the households will always hold positive balances by construction given that positive utility is derived from holding the monetary asset.

By combining 2.8 and 2.9 we can find the wage ratio of this economy:

$$w_{st} = \frac{\eta}{1-\eta} w_{ut} \tag{2.11}$$

Firms in the economy maximize profits that equal to the output produce by skilled and unskilled labour and capital less the cost needed to be paid for these inputs. The firms problem looks the following way:
$\max_{[N_{st},N_{ut},K_t]} \sum_{t=0}^{\infty} \chi_t =$

$= \sum_{t=0}^{\infty}[A_t K_t^{1-\alpha}[(B_t\omega(N_{st} - \frac{d}{2}(N_{st} - N_{st-1})^2)^{1-\upsilon} + N_{ut}^{1-\upsilon})^{\frac{1}{1-\upsilon}}]^{\alpha} - r_t K_t - w_{st}hN_{st} - w_{ut}hN_{ut}]$
leads to the following equilibrium conditions:

$$r_t = A_t(1-\alpha)K_t^{-\alpha}N_t^{\alpha} \tag{2.12}$$

$$w_{st} = A_t\alpha K_t^{1-\alpha}N_t^{\alpha-1-\upsilon}\omega B_t(N_{st} - \frac{d}{2}(N_{st} - N_{st-1})^2)^{-\upsilon}(1 - d(N_{st} - N_{st-1})) \tag{2.13}$$

$$w_{ut} = A_t\alpha K_t^{1-\alpha}N_t^{\alpha-1-\upsilon}N_{ut}^{-\upsilon} \tag{2.14}$$

Where N_t is given in 2.4.

Equation 2.12 equates the rate of return on capital to its marginal product. Equations 2.13 and 2.14 equate the skilled and unskilled wages to the respective marginal products of labour.

After solving for the steady-state of the model, we linearize the model by applying a Taylor approximation around the steady-state. We obtain a system of linear equations where the endogenous variables are a function of past variables and current shocks. These equations can be converted into forward looking equations. This allows us to trigger the shocks one by one in period 1 and construct the impulse-response functions (IRFs). The IRFs will allow us to observe how the variables react following a given shock.

Before constructing the IRFs, it is necessary to calibrate the model. Within the class of models that we employ, calibration is often used to obtain the response of the model to policy shocks, for example in [King & Rebelo, 2000]. For the benchmark case, the parameters $\eta, d, \omega, \upsilon$ were calibrated using similar values as in [Prasad, 1996]. However, there is more uncertainty on the values of the parameters $\upsilon, \eta,$ and ω. By altering the values of these parameters, I present the alternative calibration scheme that produces a different response of aggregate hours to a skill-specific technology shock while not violating the sign restrictions (keeping the responses of all the other variables to the skill-specific technology shock the same as under the benchmark calibration) as shown in Figures 2.3, 2.4, and 2.5. This allows my model to feature any kind of responses of the aggregate hours to a technology shock, while imposing sign restrictions on the other variables. The calibrated parameters under the baseline and alternative calibration schemes are summarized in Table 2.1.

The shock to skilled labour was calibrated identically to the total factor productivity shock. The demand shock was calibrated according to [Wen, 2006]. To obtain the IRF to a one unit positive shock to total factor productivity and skilled labour productivity, we employ the Blanchard-Khan methodology.

2.4.2 Shock identification

The VAR system v_t introduced in Equation 2.15 consists of the following endogenous variables:

Impulse Variable Measure of technological knowledge

Response Variable Measure of aggregate hours worked

Support Variables Measures of GDP, prices, difference in employment between skilled and unskilled workers

Where the reaction of the response variable to the impulse is the subject of examination. The reaction of the support variables to the impulse is predetermined and helps to identify the shock of interest. The model provides us with the sign restrictions that uniquely identify the aggregate technology shock, the skill-specific technology shock, and the preference shock.

Figure 2.2 depicts the impulse responses of the variables to the aggregate, skill-specific, and preference shocks derived from the model.

Table 2.1: Model Calibration

Parameter	Symbol	Baseline	Alternative	Sources
Proportion of skilled to unskilled workers	η	0.59	0.5	[Prasad, 1996] points to empirical studies that identify 0.59 but speculates on possibility of using 50/50 depending on how "skilled" is defined
Labour share of output	α	2/3	2/3	The range of the values of that parameter in RBC Literature goes from 0.7 to 0.58
Capital depreciation rate	δ	0.025	0.025	[Prasad, 1996]
Elasticity of substitution between skilled and unskilled labour	υ	0.4	0.6	[Prasad, 1996] quotes the range in the US manufacturing being equal to [0, 0.65], and uses 0.4
Skilled labour productivity premium	ω	1.67	1.4	Determined by the values of η and υ in the steady state condition: $N_s = (\frac{\omega*(1-\eta)}{\eta})^{\frac{1}{\upsilon}}$
Discount factor	β	0.985	0.985	Common in RBC literature, implies long run returns of about 6% per annum
Skilled labour adjustment cost parameter	d	100	100	[Prasad, 1996] picks the value equal to 100 to yield a ratio of standard deviation of skilled to unskilled employment of between one third and one fourth. These values are marginally higher than the variability of college-educated workers relative to workers without a college degree.
Fixed number of hours worked by an agent	h	0.53	0.53	[Hansen, 1985], [Prasad, 1996]
Wage elasticity	θ	3.42	3.42	Determined using the steady state condition for total employment $\bar(N) = \frac{\alpha*\frac{\bar{y}}{\bar{c}}}{\theta+\alpha*\frac{\bar{y}}{\bar{c}}}$
Demand shock parameters	ρ_z and σ_z	0.98 and 0.04	0.98 and 0.04	[Wen, 2006]
Technology shock parameters	ρ_a, ρ_b, σ_a and σ_b	0.979 and 0.0072	0.979 and 0.0072	[King & Rebelo, 2000]

Output From Figure 2.2 Panel (a) we find that output increases following all the three types of shocks. Therefore, just by assuming that output increases, we would not be able to distinguish any of the shocks of interest from the rest of the shocks. At the same time, it would be very strange not to impose the positive reaction of output to a technology shock. It is difficult to think of a technological improvement that would reduce output in the economy. Therefore, the positive reaction of the output to both types of the technology shock will be one of the identifying restrictions we will adopt.

Prices Figure 2.2 Panel (b) presents the response of the aggregate prices to the three shocks. We observe that both types of technology shock lead to a decrease of the aggregate price level, while the preference shockleads to an increase in the price level. This reaction of the price level allows us to distinguish the supply side (technology) shocks from the demand side (preference) shocks. The reaction of prices to the different types of shock explains why we placed money in the utility function. It is well known that money in the RBC class of models is neutral and does not affect real variables, therefore besides the reaction of the aggregate prices to the three shocks money plays no role in the model and is redundant.

Employment Difference Figure 2.2 Panel (c) demonstrates the difference in the employment of skilled and unskilled workers in the model. As aggregate prices allowed us to distinguish the technology shocks from the preference shock, the employment difference allows us to distinguish the skill-specific and skill-neutral technology shocks from each other. We observe an increase in the employment difference following a skill-specific technology shock that makes skilled workers relatively more productive. The aggregate technology shock, on the other hand, reduces the employment difference since all the workers become more productive, while there is a "training cost" to increase the number of skilled workers.

Aggregate Technology Figure 2.2 Panel (d) depicts the reaction of aggregate technological knowledge to the shocks. In the model the aggregate technological knowledge is assumed to be a sum of two technological processes A_t and B_t. Therefore both types of the technology shocks will bring a positive reaction of the aggregate technological knowledge.[22] At the same time the aggregate technological knowledge, as it is modeled, shows no reaction to the preference shock. If this was the case with the data, we would not need to put money in the utility function and impose any restrictions on the aggregate price level. We would be able to distinguish supply side (technology) shocks from the demand side (preference) shocks by imposing a zero reaction of the measure of the technology to the preference shock and positive reaction to the technology shocks. However in reality we do not have a perfect measure of the aggregate technological process. The number of patent claims measure with all its advantages still reacts to the changes in the aggregate demand as it was discussed in [Griliches, 1990]. Thus we will not be able to use our empirical measure of the technological knowledge to distinguish technology shocks from the preference shocks.

Aggregate Hours Worked The responses of the aggregate hours worked are presented on Figure 2.2 Panel(e) and Figures 2.3, 2.4, and 2.5 Panels (d). Figure 2.2 Panel (e) demonstrates that the aggregate hours increase following the aggregate technology shock as well as the preference shock and decrease following the skill-specific technology shock. Therefore, the model allows aggregate hours to react in both directions and thus allows us to uniquely identify all three shocks without imposing any restriction on hours. As a result, we will let the data decide which way hours worked react following different types of shocks. Moreover, the model leaves room for hours worked to react differently to the skill-specific technology shock (Figures 2.3, 2.4, and 2.5 Panels (d)).

The calibration of the elasticity of substitution between skilled and unskilled labour in the production function υ, the skilled productivity premium ω, and the adjustment cost parameter d were taken from [Prasad, 1996]. The baseline calibration results show a decrease of the aggregate hours worked following a skill-specific shock. [Prasad, 1996] discusses the possibility for the calibrated parameters to be altered. The available range for the υ is between $[0, 0.65]$. The choice of υ pins down the value of ω. Altering the value of the adjustment cost d, as it is shown in [Prasad, 1996], affects the cyclical variability of the skilled labour and subsequently aggregate hours worked. Altering the values of υ, ω, and d within the acceptable range leads to an increase in the aggregate hours following the skill specific technology shock. Therefore, we are not restrictive about the way

[22] Studies of General Purpose Technologies (see [Lipsey et al., 2005] or [Helpman & Trajtenberg, 1998b] for example) indicate that important innovations that affect aggregate productivity for the whole economy open the door for multiple further improvements at the same time. Thus one important invention will lead to a further increase in innovative activity and will result in multiple applications in different sectors of the economy. Therefore it is reasonable to assume that a technological improvement will generate further innovative activity and therefore the aggregate technological knowledge will be above its steady state value for a period of time.

the aggregate hours will react to "some technology"[23] in any way.

Overall the model implies an increase in aggregate hours worked following an aggregate technology shock and either a decrease or an increase in hours worked following the skill-specific shock as shown on Figures 2.3, 2.4, and 2.5 Panels (d). Note that alternative calibration does not affect the responses of other examined variables to any of the examined shocks qualitatively as it is shown on Figures 2.3, 2.4, and 2.5 Panels (a), (b), and (c).[24] Therefore, the model provides a robust identification of aggregate and skilled technology shocks without imposing strong qualitative restrictions on the reaction of aggregate hours worked.

Based on the above discussion, I apply sign restrictions to the data. Table 2.2 summarizes the restrictions for three shocks analyzed earlier. Note that each shock can be uniquely identified using these restrictions.

It would also be possible to include the wage difference between skilled and unskilled workers as a support variable in order to distinguish between skill-specific and aggregate technology shocks. The wage premium is not supported by the current model, but supported by multiple studies ([Allen, 1996], [Aghion et al., 2002] ,[Goldin & Katz, 2007], [Marquis et al., 2011], [Katz & Margo, 2013]) and by the data (Figure 2.1).

Table 2.2: Summary of Sign Restrictions

Restriction	Hours worked	Output	Prices	Employment Difference	Patent claims
Aggregate Shock	None	+	-	-	+
Skill-specific Shock	None	+	-	+	+
Preference Shock	None	+	+	-	None

2.5 Empirical Results

According to the theoretical model, a technology shock of either kind increases output, reduces prices, has a positive effect on the aggregate technology process and has an ambiguous effect on total hours worked. The response of the latter depends on the type of shocks and the calibration of certain parameters. The combination of the theoretical restrictions with the data through the VAR methodology will reveal the effects of the three aforementioned shocks on aggregate hours worked.

Empirically we proceed with the sign restriction VAR described in details in [Fry & Pagan, 2010]. Such a methodology was successfully implemented in [Faust, 1998], [Canova & De Nicolo, 2002], and others. Moreover, sign restrictions are now popular in the modern VAR literature and are used to answer other questions, including the effects of a technology shock on economic variables.[25]

Suppose we have a VAR of the following structural form:[26]

$$v_t = \Gamma_p L^p v_t + \epsilon_t \tag{2.15}$$

[23]Recall that the patent data reports both skill specific and skill neutral innovations. When we use an aggregate patent measure of technology we are not able to distinguish one type of technological improvement from another.

[24]Figures 2.3, 2.4, and 2.5 do not present the responses of the aggregate technology to the shocks since those are not affected by the alternative calibration.

[25]See aforementioned [Debola & Neri, 2007] or [Peersman & Straub, 2004] for example.

[26]I am omitting a constant term and a deterministic time trend for notation simplicity.

Where v_t is the vector of n endogenous variables: $v_t = \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}$, $v_{1,t}$ representing economic indicator variable(s) and $v_{2,t}$ standing for a technological indicator. L is a lag operator, p is the number of endogenous lags: $p = 0 \ldots P$. $\epsilon_t = \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix} \sim IID\Big[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\Big]$ and Γ_p are the matrices of coefficients.

The system in 2.15 cannot be estimated due to endogeneity. However it can be transformed into a reduced form the following way:

$$\begin{aligned} [I_n - \Gamma_0]v_t = Bv_t = (\Gamma_1 L + \Gamma_2 L + \ldots + \Gamma_p L)v_t + \epsilon_t \\ v_t = (B^{-1}\Gamma_1 L + B^{-1}\Gamma_2 L + \ldots + B^{-1}\Gamma_p L)v_t + B^{-1}\epsilon_t \\ v_t = (C_1 L + C_2 L + \ldots + C_p L)v_t + e_t \end{aligned} \tag{2.16}$$

Where $Var(e_t) = Var(B^{-1}\epsilon_t) = E(B^{-1}\epsilon_t\epsilon_t'(B^{-1})') = \Sigma$.

The system in 2.16 can be estimated using OLS, GLS, or MLE. The OLS approach was used in this study. After estimating the reduced form VAR, we can orthogonalize the estimated $\hat{\Sigma} = \tilde{P}\tilde{P}'$ using a Choleski decomposition.[27] Then for any orthonormal matrix Q such that $Q^{-1} = Q'$ and $Q'Q = QQ' = I_n$, we can check if the new orthogonal decomposition

$$\hat{\Sigma} = \tilde{P}QQ'\tilde{P}' = \hat{P}\hat{P} \tag{2.17}$$

provides IRFs satisfy the sign restrictions that we imposed in Table 2.2 for a predetermined period. If the draws do not satisfy the sign restrictions, we discard them and repeat the procedure (i. e. repeat placing different orthonormal Qs) until the sufficient number of draws satisfying the restrictions is obtained[28]. From the draws that survived the elimination process, we extract the median, the 16th and the 84th percentile so that we have approximately one standard deviation bound around the median.[29]

There are two common ways of obtaining the orthonormal matrix Q. One could use Givens matrices or QR decomposition. In [Fry & Pagan, 2007] it was shown and proven for the bi-variate VAR case that the methods are equivalent. The latter method is believed to be more efficient for larger VAR systems[30] and therefore is adopted here. The QR decomposition method implies drawing random variables $W \sim N(0, I_n)$, and decomposing $W = QR$ where Q is an orthogonal matrix and R is a triangular matrix. Many draws of W generate many matrices Q that can be plugged in equation 2.17.

Figures 2.6, 2.7, and 2.8 depict the resulting empirical IRFs for the logarithmized and differenced data with the restrictions held for two periods after the shock.[31]

Figure 2.6 shows the IRFs for the aggregate technology shock with one standard deviation bounds. We observe an increase in aggregate hours following the shock. The rest of the variables also react in accordance with the theory. Moreover, we also observe an increase in aggregate hours following the skill-specific technology shock (Figure 2.7) which lets us to conclude that there is a strong pro-cyclical effect in aggregate hours generated by technological innovations. Both shocks lead to an increase in output, hours, and an initial increase in the innovative activities. Prices fall after both shocks. As we expected, aggregate hours worked increase following the aggregate technology shock. The increase in aggregate hours worked following a skill specific technology shock is explained by growing number of skilled workers (as shown in Figure 2.1) that benefit from this shock.[32]

[27] Note that eigenvalue-eigenvector decomposition would also work.

[28] The baseline results of this study are obtained by accumulating 5000 successful draws for each shock. However, the results are confirmed with 10000 successful draws experiments.

[29] This is a common way to analyze sign restriction IRFs in the literature. See [Uhlig, 2005] or [Canova, 2007] for examples.

[30] See [Rubio-Ramirez et al., 2005] or [Fry & Pagan, 2010] for the discussion.

[31] All the IRFs in this study were constructed using Ambrogio Cesa-Bianchi, 2014; "VAR Toolbox for Matlab", `sites.google.com/site/ambropo/`.

[32] Recall that the model presented earlier was flexible enough to imply both an increase and a decrease of the aggregate hours following the skill-specific shock.

Finally, Figure 2.8 depicts the empirical IRFs for the demand shock. We observe an increase in output and hours worked as predicted by the model. We also observe an increase in prices, a fall in the difference between skilled and unskilled employment and a very weak reaction of the aggregate technological process represented by the patent data. All of these responses are consistent with the theory and the model.

The increase in hours worked following the aggregate technology shock is an expected result and is consistent with real business cycle theory. As it is discussed in [King & Rebelo, 2000], in the RBC framework planned investment increases in response to a positive productivity shock. It is optimal for the representative household to save some fraction of the higher current output and it is efficient to lower consumption and raise work effort.

The increase in aggregate hours worked following a skill-specific technology shock can be explained by the growing proportion of skilled workers in the economy, as well as by the elasticity of substitution between skilled and unskilled workers and the skilled labour productivity premium. All of these factors determine the level of employment of skilled and unskilled workers and therefore the impact of skilled technology shock on both groups and on aggregate hours in general. The theoretical model used two alternative sets of parameters to retrieve both positive and negative theoretical responses of aggregate hours worked to a skill-specific technology shock. The ranges of values for these parameters were obtained from [Prasad, 1996] who mentions lack of direct evidence available to restrict the skilled-unskilled labour elasticities of substitution and productivity premium and determines the potential range based on 1979 empirical study conducted by Halvorsen and Ford.[33] In general, a more recent empirical study of skilled and unskilled workers' employment and contribution to production would be beneficial.

Besides examining IRFs, a common tool used in the VAR analysis is forecast error variance decomposition (FEVD). When sign restriction methodology is used, similarly to the IRFs, we obtain and empirical distribution of FEVDs instead of a single observation. Figure 2.12 depicts median, 16th, and 84th percentiles of FEVDs for the skilled technology shock when the data is specified in first differences. We observe the technology shock explaining a relatively substantial part of the aggregate hours variation, approximately 30%. The FEVDs for the aggregate technology shock are similar to those reported for the skilled shock. It is necessary to note, however, that when the level data is used, the amount of information contributed by the technology shocks is much more modest.

2.6 Robustness

In this section we provide more supporting evidence for the positive reaction of the aggregate hours worked to a technology shock. To show the robustness of the aforementioned results we will conduct the following analysis:

1. Use level instead of first difference data to observe whether data specification matters for our results.
2. Use more common identification techniques. Instead of using the sign restriction approach, we will first impose short run restrictions and then long run restrictions on the data.
3. Finally, we will try out an alternative measure of technological knowledge to see the resulting effect of aggregate hours a technological innovation.

There were other types of robustness checks attempted but not reported, besides the three listed. Different VAR specifications with and without trends and constant terms were examined for the data in log-levels and first differences. The resulting IRFs support the conclusion that aggregate hours increase

[33] Unfortunately I was not able to obtain a copy of this study.

following both types of technology shocks. Different number of lags in the underlying VAR were also tested. The default of two endogenous lags were reported since this number was supported by the majority of information criteria and is in accordance with the literature. However, for some data specification certain information criteria suggested to use an alternative quantity of endogenous lags. All the suggested cases were examined with no impact on the conclusions.

We also impose different assumptions regarding how long the sign restriction had to hold. Besides the default case of two periods, four and six periods cases were examined. Increasing the number of periods for the restrictions to hold, lead to a significant increase of the computing time and very low rate of survival especially for the data in first differences. The IRFs that survived the sign restrictions had the same positive effect of a technology shock on aggregate hours worked.

2.6.1 Levels or Differences?

One of the weaknesses of the long run restriction approach that is commonly used to estimate the effects of technology shocks on economic variables is that the data specification matters. [Christiano et al., 2003b] find that if hours worked are specified as a difference, they fall after a technology shock in the long run VAR specification. On the other hand, when hours worked are represented in levels, they obtain the reverse, that is hours worked increase following a technology shock.section, we investigate if the results discussed in section 2.5 are robust to specifying our variables in levels and including a trend in the VAR.

Figures 2.9, 2.10, and 2.11 depict the IRFs resulting from the data specified in levels with a trend included in the VAR. The sign restrictions imposed on the data are the same as the ones used for the first difference specification. As we observe from the diagrams, the level specification does not affect the conclusions we make about the behaviour of the aggregate hours or other variables after the technology or preference shocks. The aggregate hours worked increase after both kinds of technology shock whether hours are presented in levels or in first differences.

2.6.2 Structural VAR with Short Run Restrictions

In this section, we conduct an analysis of the effect of technology shock on aggregate hours worked and the other variables in Equation 2.15 using more conventional short run restrictions instead of sign restrictions. Short run restrictions imply that some of the components in the matrix B from Equation 2.16 must be assigned zero values. To be more precise, the system with 5 variables requires $\frac{5*(5-1)}{2} = 10$ restrictions. Each of the restrictions implies that we assume no immediate effect from a shock to variable j on variable k.

Unlike the sign restrictions, when we conduct the traditional SVAR analysis, we cannot analyze each of the shocks separately. Given the five variables in the system we need to examine five structural shocks simultaneously. We have already examined two types of technology shocks, a preference shock and a monetary shock in the model.

The short run SVAR framework will not allow us to distinguish between skill-specific and skill-neutral technology shocks. Both of these shocks affect the employment difference variable according to the model. Therefore we will analyze "a technology shock" that is a structural innovation to the patent variable that combines both skill-specific and skill-neutral technology shocks.

Unfortunately, we can find limited information about the short run restrictions from the model.[34] Both types of technological innovations in the model are completely exogenous and unaffected by other variables. We assume that none of the shocks should have any immediate effect on technological knowledge

[34][Kilian, 2012] notes that often there are not enough zero short run restrictions to achieve VAR identification. The weakness of the short run SVAR is that the economic theory often does not provide enough restrictions or the restrictions imposed based on one model may contradict numerous alternative models.

except the technology shock itself. Given that technological knowledge is approximated by the patent data and if we take into account the fact that it takes time to file a patent and get it granted, it is safe to assume that there is no short-run immediate effect of any other shock on the technology. One can also think that it would still require some time to invent something even if there is a strong market demand for that invention. We also know that the monetary shock will not influence any real variables in the model and only the price level will be affected.

When we orthogonalize the VAR coefficients' variance covariance matrix using a Cholesky decomposition, and selecting the variables' order, the patent claims variable will be placed first and the aggregate measure of prices will be placed last. Since we are only interested in the identification of the technology shock, the order of the other variables does not matter for us. What matters is that patent claims do not contemporaneously respond to any shocks except their own.

Table 2.3: Short Run restrictions

Variable/Innovation	Technology	Preference	Labour Supply	Training Cost	Monetary
Patent Claims	.	0	0	0	0
Output	.	.	0	0	0
Hours Worked	.	.	.	0	0
Empl-t Diff-ce	.	.	.	.	0
Prices	.	.	.	.	.

Table 2.3 presents the short run zero restrictions on the effects of structural innovations on the endogenous variables. Figure 2.13 represents the structural impulse responses of the endogenous variables to the technology shock. We find that the response of the technology variable to a positive technology shock is volatile. This volatility is observed when we look at the responses of the other variables to the technology shock. Therefore, the short run identification methodology does not provide any conclusive results about the effect of technology on aggregate hours or any other variable. Note that the earlier studies that used short runs restriction SVAR models to examine the impact of technology on inputs into production also lead to mixed conclusions.[35]

We also observe a similar result when different ordering of the variables were implemented (we do not report the results here). The short run identification scheme in table 2.3 was chosen because most of the restrictions are justifiable and also because all the other structural innovations, except the technology, behave as they are expected. The resulting IRFs are compatible with the theoretical IRFs from the model.

We conclude that the combination of data and methodology presented in this subsection does not allow us to identify a technology shock. It could be possible to explain the volatile reaction of aggregate hours to a technological innovation on figure 2.13. However, it is difficult to explain such a reaction for the output and the technological variable itself. The reason for this incompatibility may be hidden in the patent data that is volatile or in the short run identification methodology. Use of non-patent technological measures (not reported) produced less volatile impulse-responses. However, the reaction of output to a positive technology shock derived from a SVAR model with short run restrictions was still not intuitive. No matter the reason for the incompatibility, the fact remains that the sign restriction approach outperforms the short run zero restrictions when the same data is used to identify technology shocks.

2.6.3 Structural VAR with Long Run Restrictions

An alternative approach to imposing the short run restrictions is to impose long run restrictions that imply an assumption of the following kind: a permanent shock to variable j has no long run effect on variable k. This approach was first introduced in [Blanchard & Quah, 1989] and since then has become

[35] See [Shea, 1998] for example.

one of the most popular ways to identify shocks in SVAR models. This is the most common approach in the literature when the effects of a permanent technology shock on economic variables are analyzed.

Table 2.4: Long Run restrictions

Variable/Innovation	Technology	Training Cost	Labour Supply	Preference	Monetary
Patent Claims	.	0	0	0	0
Eml-t Diff-ce	.	.	0	0	0
Hours Worked	.	.	.	0	0
Output	.	.	.	.	0
Prices	.	.	.	.	.

Table 2.4 presents the proposed long run identifying restrictions. The zero long run restriction on the technological measure from the preference shock is the one that requires our special attention. First we assume that no shock besides the technology shock will have a permanent effect on technological knowledge. Thus, even demand side shocks that are believed to affect the patent data are assumed to have only transitory effects. We also assume that the labour supply and consumption preference shocks have a symmetric effect on skilled and unskilled workers. Thus the long run effect of these shocks on the employment difference is zero. Note that a general technology shock will also have only transitory effect on the employment difference. The alternative restriction with zero long run effect of technology shock on the employment difference was attempted. The results are similar to the results reported below.

Figures 2.14 and 2.15 present the resulting IRFs form the long run restrictions that we imposed on the data. Figure 2.14 presents the IRFs when the level data is used. Figure 2.15 shows the IRFs when aggregate hours worked are first-differenced in order to observe whether the variable specification matters as it did in some previous studies.[36] In both cases hours worked increase following a positive technology shock when long run restrictions are imposed.

2.6.4 Alternative Measure of Technological Process

To answer the debates about how well patents represent the aggregate technological process, in this section we introduce an alternative measure of technology. Using this alternative measure of technology, we again run our SVAR using sign restrictions. The alternative measure is the utilization adjusted total factor productivity introduced in [Basu et al., 2006] and re-introduced and described using quarterly values in [Fernald, 2012].

In response to the critique of the indirect Solow residual based, measures of technology, [Basu et al., 2006] construct a series for technology by controlling for non-technological effects in aggregate TFP. The authors take into account varying capital and labour utilization, nonconstant returns and imperfect competition.

$$Y_{,t}i = F^i(A_{i,t}K_{i,t}, E_{i,t}H_{i,t}N_{i,t}, M_{i,t}, Z_{i,t}) \tag{2.18}$$

[Basu et al., 2006] estimate equation 2.18 for industry i, where Y_i stands for gross output, K_i for the capital stock, N_i for the number of workers, M_i for the other inputs into production such as intermediate goods, energy, and raw materials, and the residuals Z_i for the technology. Equation 2.18 also takes into account hours worked per employee H_i, the effort of each employee E_i, and the capital utilization rate A_i. After estimating equation 2.18 for each industry, the authors define the aggregate technological process as the weighted sum of the resulting residuals Z_i.

[36] See [Christiano et al., 2003b] for the discussion.

[Basu et al., 2006] constructed their utilization adjusted measure of technology for yearly frequency. [Fernald, 2012] notes that certain industry-level detailed parameters on the input/output flows and industry investment are unavailable for the quarterly and higher frequencies. Therefore the quarterly measure of utilization adjusted TFP, introduced in [Fernald, 2012], does not control for all non-technological sources of fluctuations. However, according to the authors, it is still superior to the alternative "naive" simple Solow residual measures since it takes into account the heterogeneity of labour and capital inputs and labour and capital utilization at the quarterly frequencies.

$$Y_t = F(A_t K(K_{t,1} \dots K_{t,J}), E_t L(H_{t,1} \dots H_{t,N}), Z_t) \tag{2.19}$$

Unlike [Basu et al., 2006], [Fernald, 2012] starts with the aggregate production right away (equation 2.19). He then distinguishes between J types of capital inputs: $K_{t,1} \dots K_{t,J}$ and N types of labour inputs $H_{t,1} \dots H_{t,N}$, where $H_{t,n}$ stands for the aggregate hours worked by the workers of type n.

[Fernald, 2012] assumes a Cobb-Douglas functional form for the production function represented by Equation 2.20 without time subscripts for simplicity.

$$Y = ((AK())^{\alpha} * (EL())^{1-\alpha})^{\mu} * Z_t) \tag{2.20}$$

In equation 2.20, α is share of capital in production such that $\alpha = \sum_j \gamma_j$ and γ_j is the share of capital of type j. $(1-\alpha)$ is the labour share in the production such that $1-\alpha = \sum_n \beta_n$ and β_n is the share of labour of type n. μ is the markup over the marginal cost such that $(\frac{\delta Y}{K_j})(\frac{K_j}{Y} = \mu(\frac{R_j K_j}{PY}) = \mu\gamma_j$ and $(\frac{\delta Y}{H_n})(\frac{H_n}{Y} = \mu(\frac{W_n H_n}{PY}) = \mu\beta_n$ where P, W, and R are the prices.

By taking natural logarithms and differentiating the production in Equation 2.20, we obtain equation 2.21.

$$\Delta \ln Y = \mu(\alpha \Delta \ln K + (1-\alpha)\Delta \ln L) + \Delta \ln U + \Delta \ln Z \tag{2.21}$$

In equation2.21, $\Delta \ln K = \sum_j \frac{\gamma_j}{\alpha} \Delta \ln K_j$. Similarly, $\Delta \ln L = \sum_n \frac{\beta_n}{(1-\alpha)} \Delta \ln H_n \equiv \Delta \ln Q + \Delta \ln H$, where $\Delta \ln H = \Delta \ln (\sum_n H_n)$ and $\Delta \ln Q = \Delta \ln L - \Delta \ln H$. The utilization effect in equation 2.21 is: $\Delta \ln U = \mu[\alpha \Delta \ln A + (1-\alpha)\Delta \ln E]$.

We can re-arrange equation 2.21 to solve for the conventional TFP:

$$\Delta \ln TFP = \Delta \ln Y - (\alpha \Delta \ln K + (1-\alpha)\Delta \ln L) = \Delta \ln U + \Delta \ln Z \tag{2.22}$$

We can also extract the utilization adjusted TFP that we will use in our VARs later.

$$\Delta \ln Z = \Delta \ln TFP - \Delta \ln U \tag{2.23}$$

The utilization adjusted TFP data has already been used for uncovering the effects of technology on labour hours in the VAR context. The authors of the original annual dataset, [Basu et al., 2006], apply the data to estimate the effects of the technology on total hours worked. They, first estimate the effect using the Solow residual equation itself. Next, they use the utilization adjusted TFP in a series of unrestricted VARs[37] and SVARs with long run restrictions. Using all three approaches, the authors find that technology shocks have a negative effect on aggregate hours.

[Christiano et al., 2003a] also use the utilization adjusted TFP as their measure of technology and long-run restrictions to examine the effects of technology on aggregate hours. They find, as it is common when long-run restrictions are employed, that the response of hours worked depend on whether or not the VAR is specified in level or first difference. [Christiano et al., 2003a] insist that aggregate hours must be specified in log-levels and conclude that hours worked increase following a technology shock.

[37] The authors claim that the technology process uncovered at the annual frequency is completely exogenous and therefore pure technology shocks can be revealed without any further restrictions.

When we use utilization adjusted TFP as our measure of technology in the VAR with sign restriction, we examine both level and difference specification for the aggregate hours. Figures 2.16 and 2.17 show the empirical IRFs drawn for the skill-specific technology shock identified in Table 2.2. In figure 2.16 aggregate hours worked as well as prices, output, and employment difference are in log-levels. The utilization adjusted TFP variable is in first difference as it is specified in equation 2.23. Figure 2.17, on the other hand, has the aggregate hours in first difference of the log-levels while prices, output, and employment difference are in log-levels and the utilization adjusted TFP variable is as it is specified in equation 2.23. These two approaches were taken to address the debate about the correct way of including the labour variable in the VAR. Using sign restrictions and this alternative measure of technology we observe that aggregate hours worked increase following a skill specific technology shock. This result is robust to whether or not we specify hours worked in levels or in first differences.

A similar conclusion can be stated about the general technology shock (figures2.18 and 2.19). We observe a positive impact of technology on aggregate hours whether the VAR is specified in levels or first differences.

2.7 Conclusion

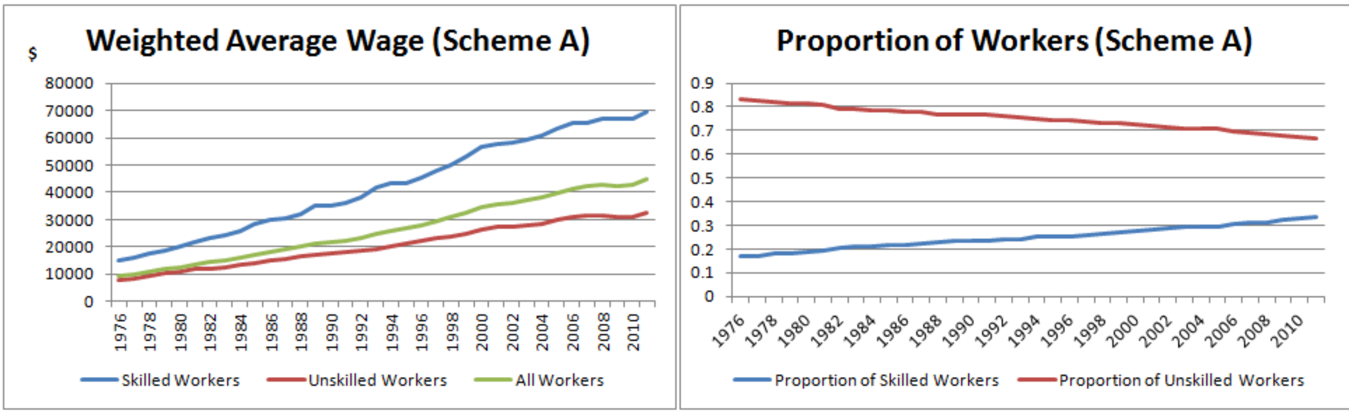
In this paper, we examined the effect of a technological innovation on aggregate hours. We used a direct patent data based measure of technological process and implemented the sign restriction VAR identification technique on the data. The identifying restrictions were determined by a simple RBC class macroeconomic model that features heterogeneous labour and two types of technology shocks. We found that both technology shocks lead to an improvement in aggregate hours worked. The conclusion is based on the observation that we have a larger and growing proportion of skilled workers who mainly benefit from a technology shock that improves their productivity.

Implementation of more conventional SVAR methodologies led to inconclusive results. The short run restrictions did not allow us to conclude anything about the effect of technological innovation on aggregate hours since the effects of technological innovations on all variables, when the short run restrictions are used, are volatile. The long run restrictions were found to be sensitive to the variable specification. This effect is in accordance with numerous studies that attempt to use long run identification methodology to examine the effect of a technology shock on aggregate hours. We believe that the sign restriction methodology we employ in this paper produces more robust results compared to the other identification methods used in the literature. The resulting effect of a technology shock on aggregate hours coincides with the conclusion of other studies that used sign restrictions but different identification mechanisms to examine how technological innovation affect economic variables empirically.

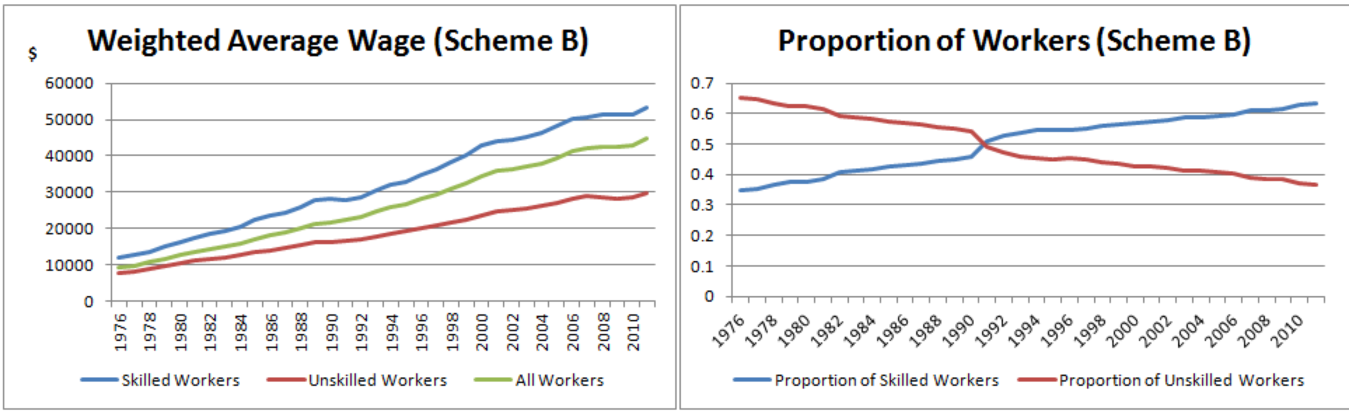
2.8 Chapter 2 Appendix

Figure 2.1

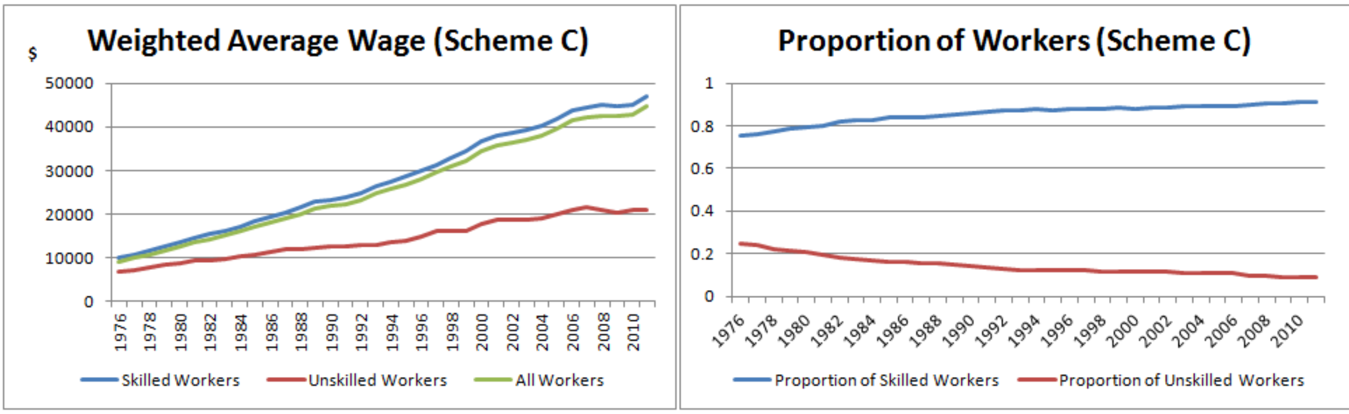
(a) Scheme A†



(b) Scheme B‡



(c) Scheme C⋆

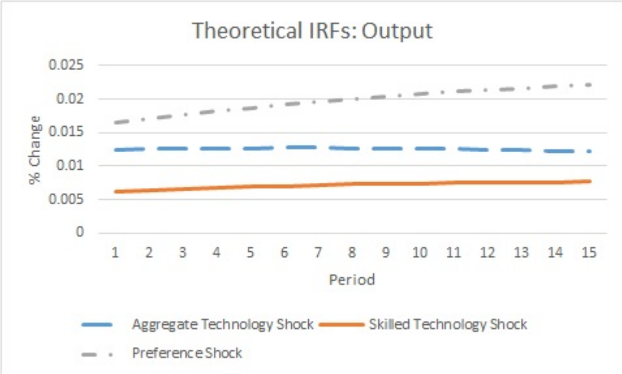


† Employees with a bachelor's degree or higher are considered to be skilled.
‡ Employees with a college diploma or higher are considered to be skilled.
⋆ Employees with a high school diploma or higher are considered to be skilled.

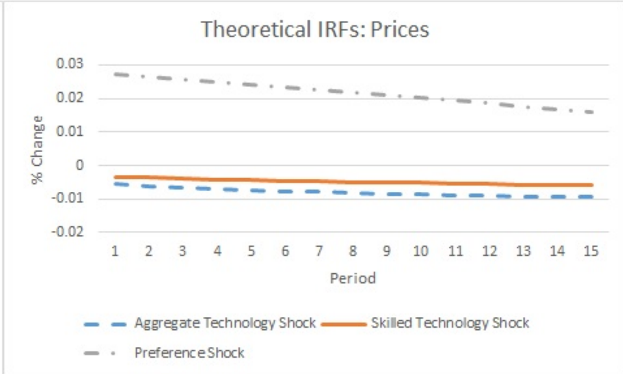
Source: United States Census Bureau, CPS Historical Time Series Tables, Table A3.

Figure 2.2: Theoretical Impulse Responses

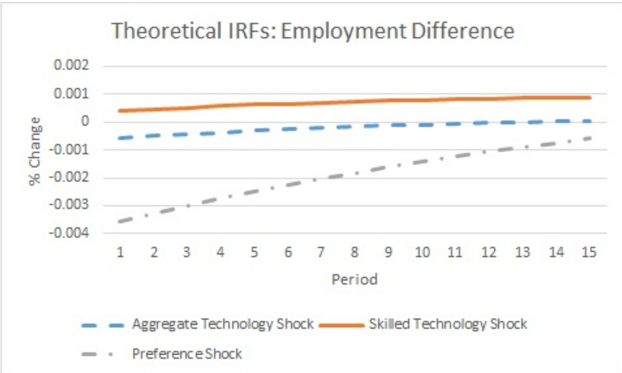
(a) Responses of Output



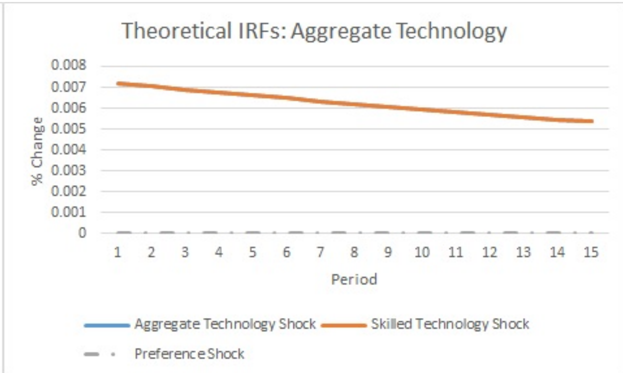
(b) Responses of Prices



(c) Responses of the Employment Difference



(d) Responses of Aggregate Technology



(e) Responses of Aggregate Hours Worked

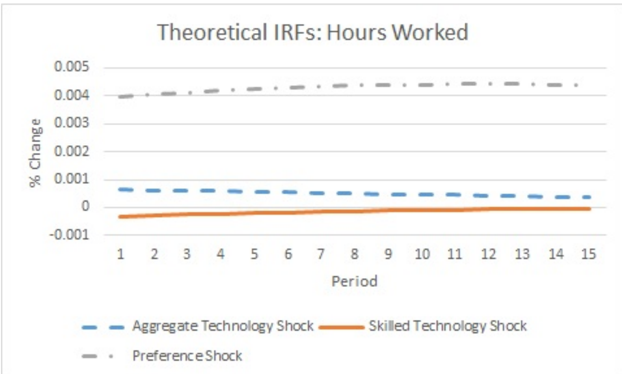
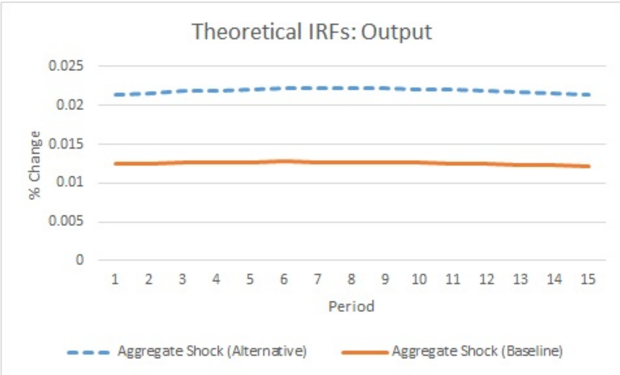
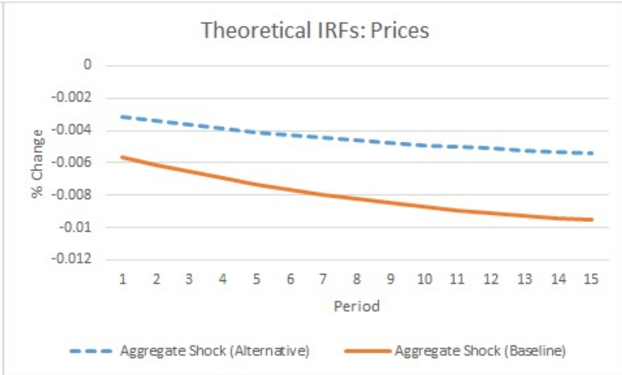


Figure 2.3: Theoretical Impulse Responses to the Aggregate Technology Shock: Alternative Specifications

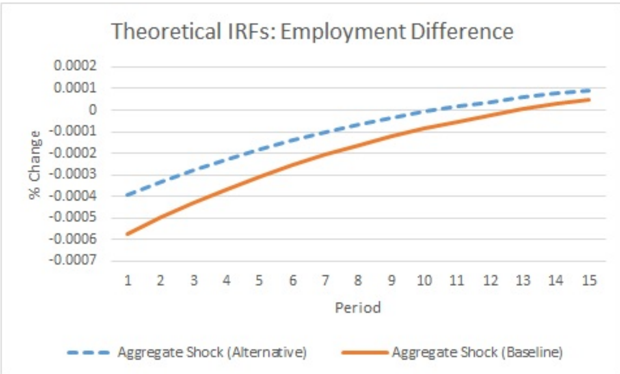
(a) Responses of the Output



(b) Responses of the Prices



(c) Responses of the Employment Difference



(d) Responses of the Aggregate Hours Worked

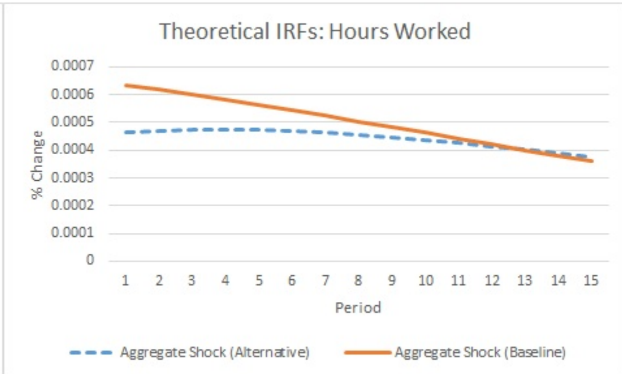
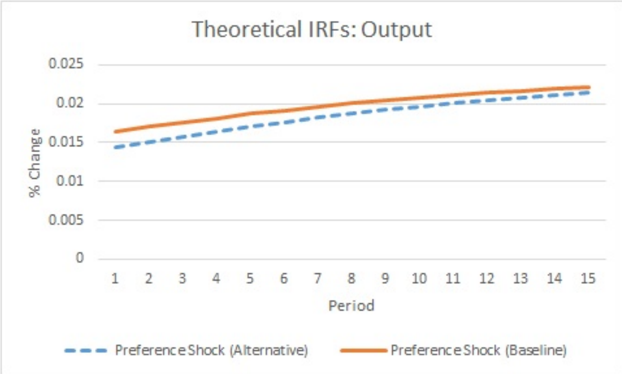


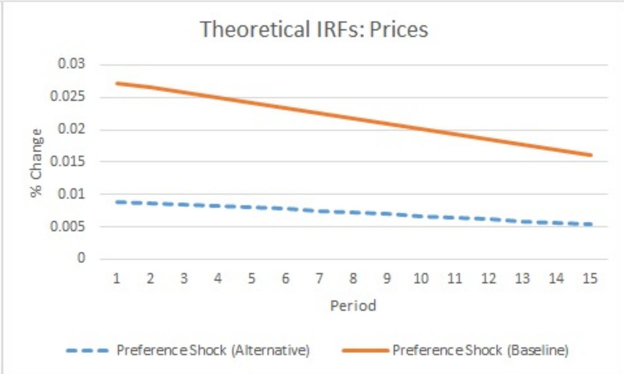
Figure 2.4: Theoretical Impulse Responses to the Skilled Technology Shock: Alternative Specifications

Figure 2.5: Theoretical Impulse Responses to the Preference Shock: Alternative Specifications

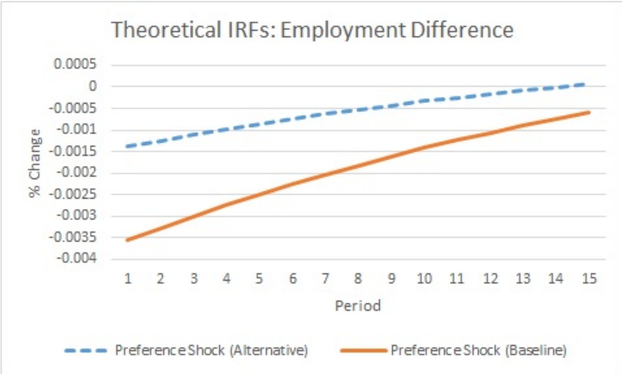
(a) Responses of the Output



(b) Responses of the Prices



(c) Responses of the Employment Difference



(d) Responses of the Aggregate Hours Worked

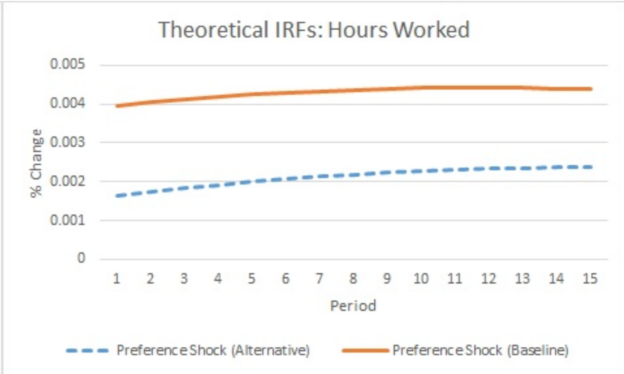


Figure 2.6: Empirical IRFs: Aggregate Technology Shock (First Differences)

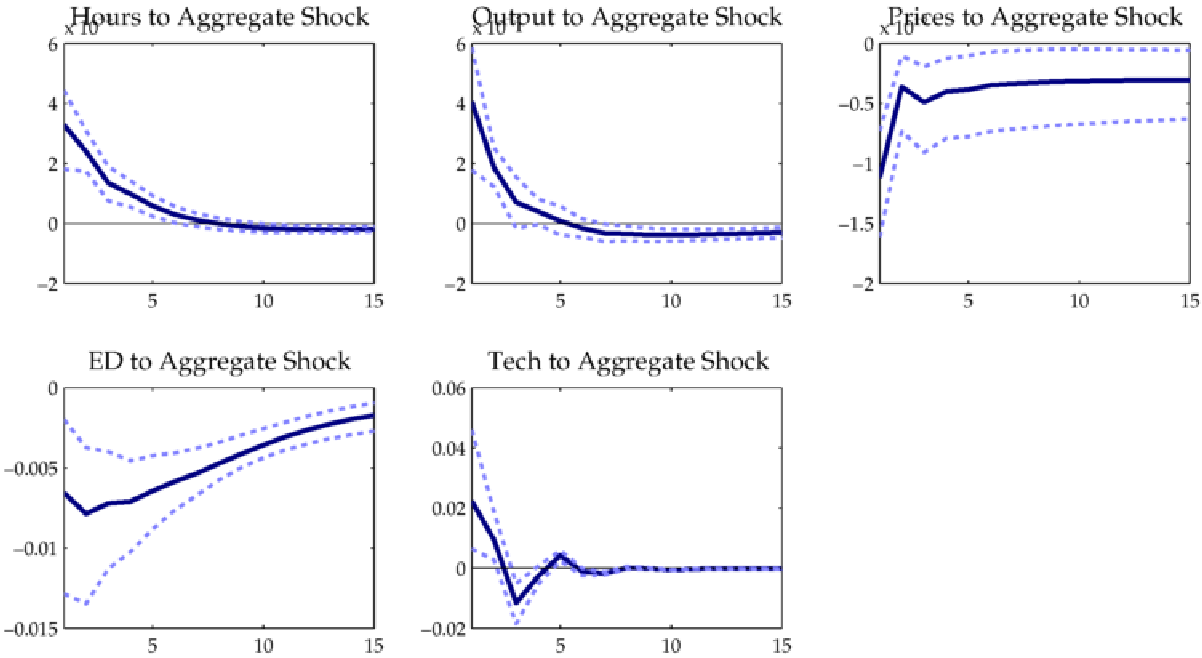


Figure 2.7: Empirical IRFs: Skilled-Specific Technology Shock (First Differences)

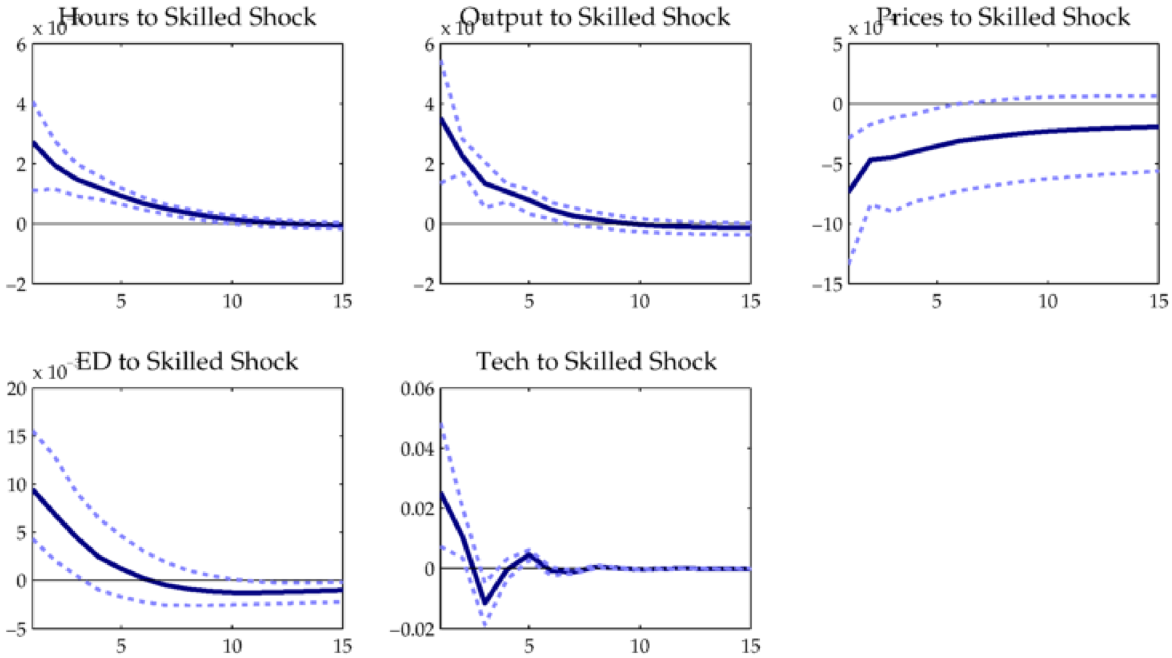


Figure 2.8: Empirical IRFs: Preference Shock (First Differences)

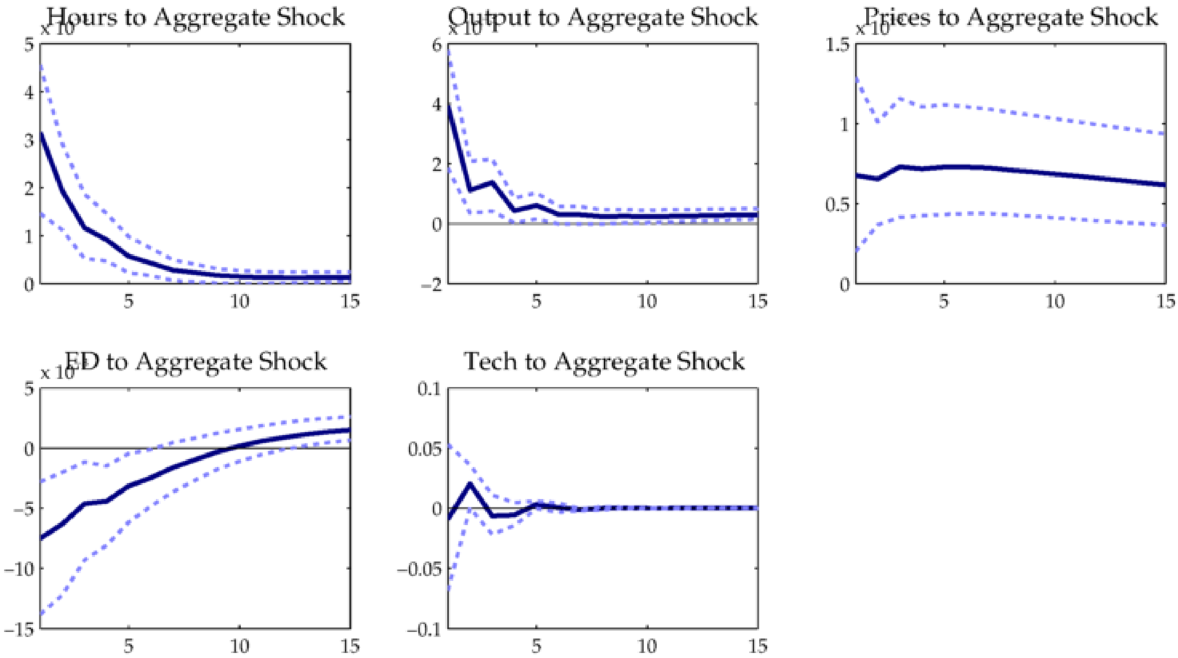


Figure 2.9: Empirical IRFs: Aggregate Technology Shock (Levels)

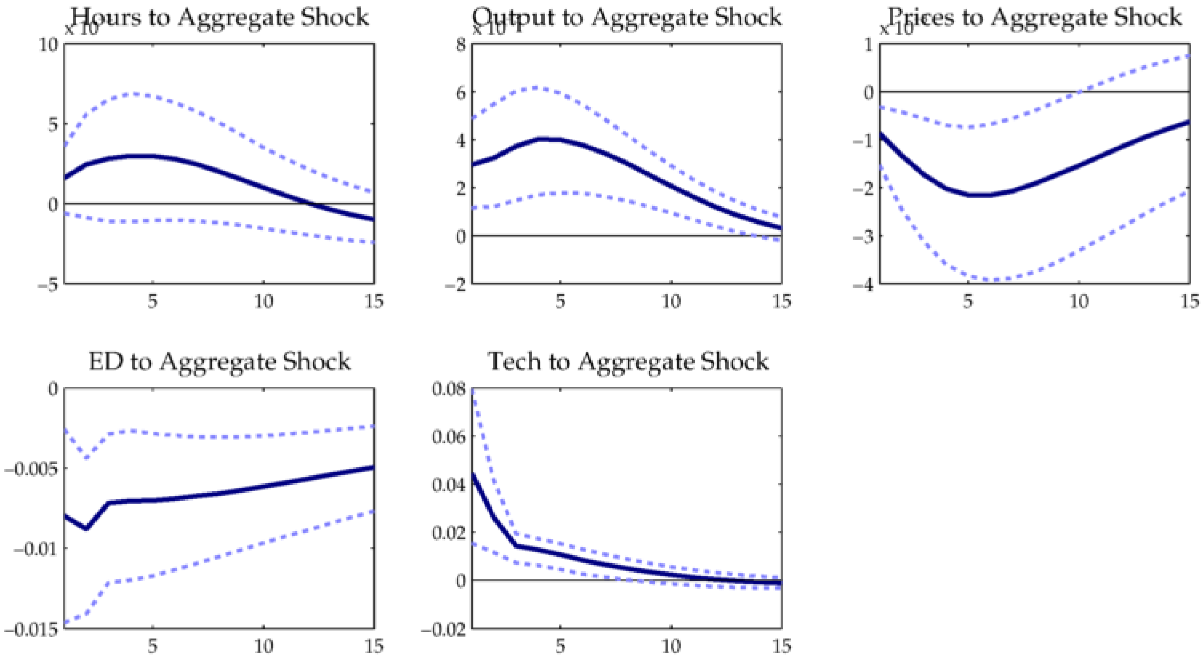


Figure 2.10: Empirical IRFs: Skilled-Specific Technology Shock (Levels)

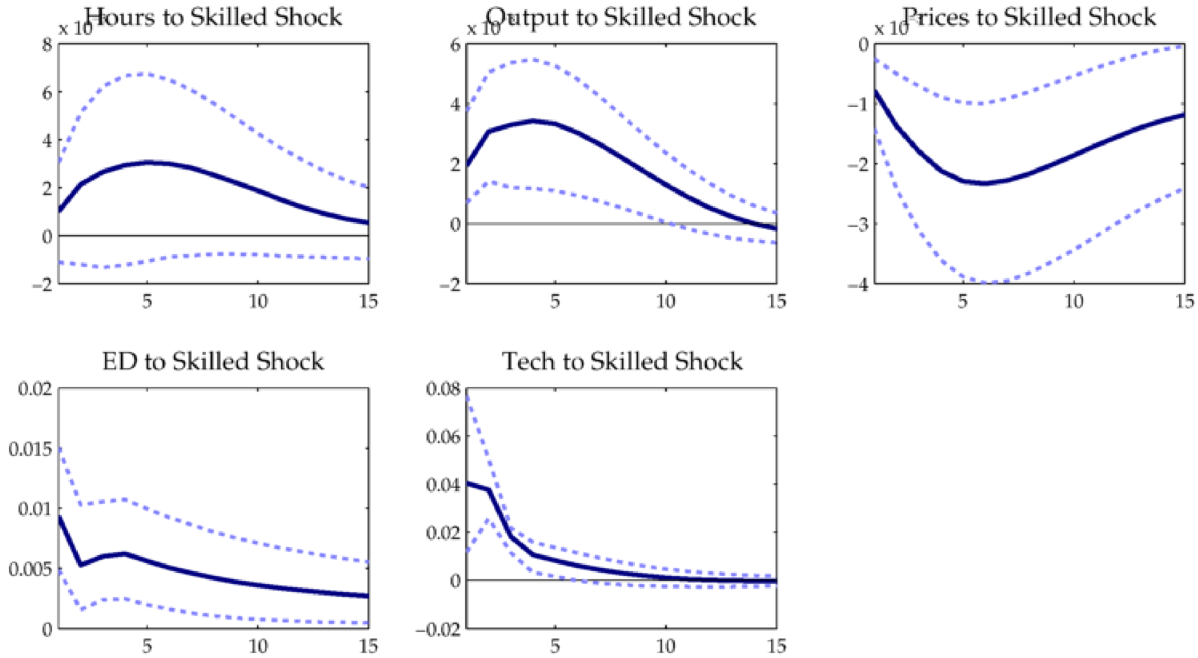


Figure 2.11: Empirical IRFs: Preference Shock (Levels)

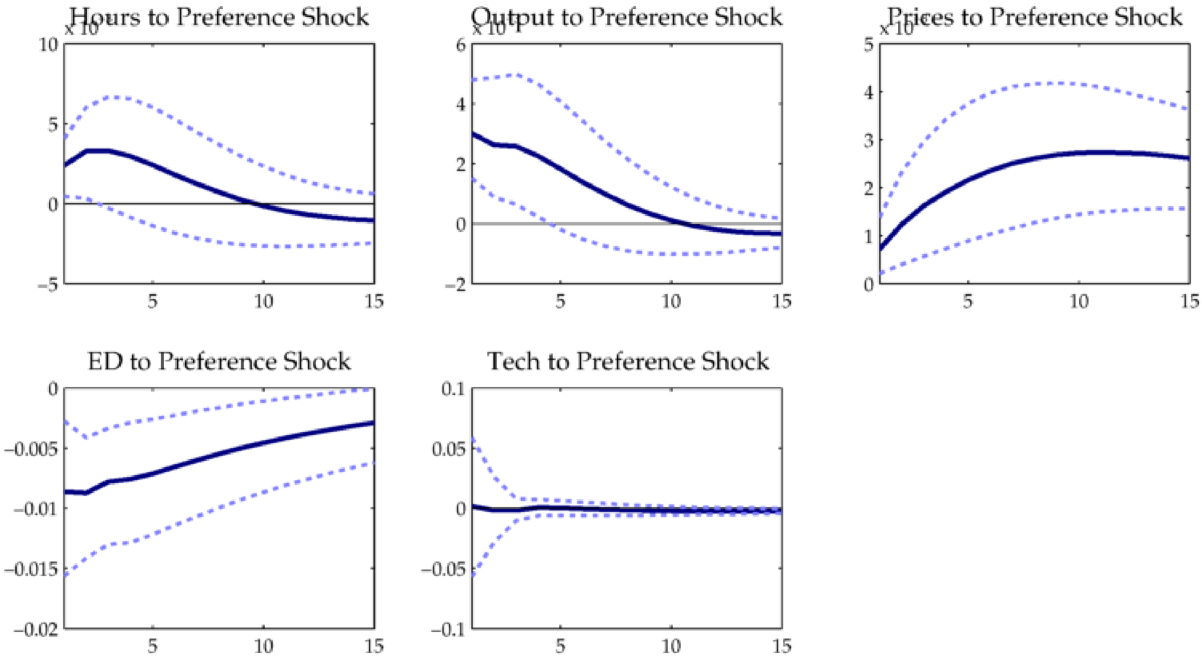


Figure 2.12: FEVD: Skilled Technology Shock (First Differences)

Figure 2.13: Short Run SVAR Impulse Responses to a Technology Shock

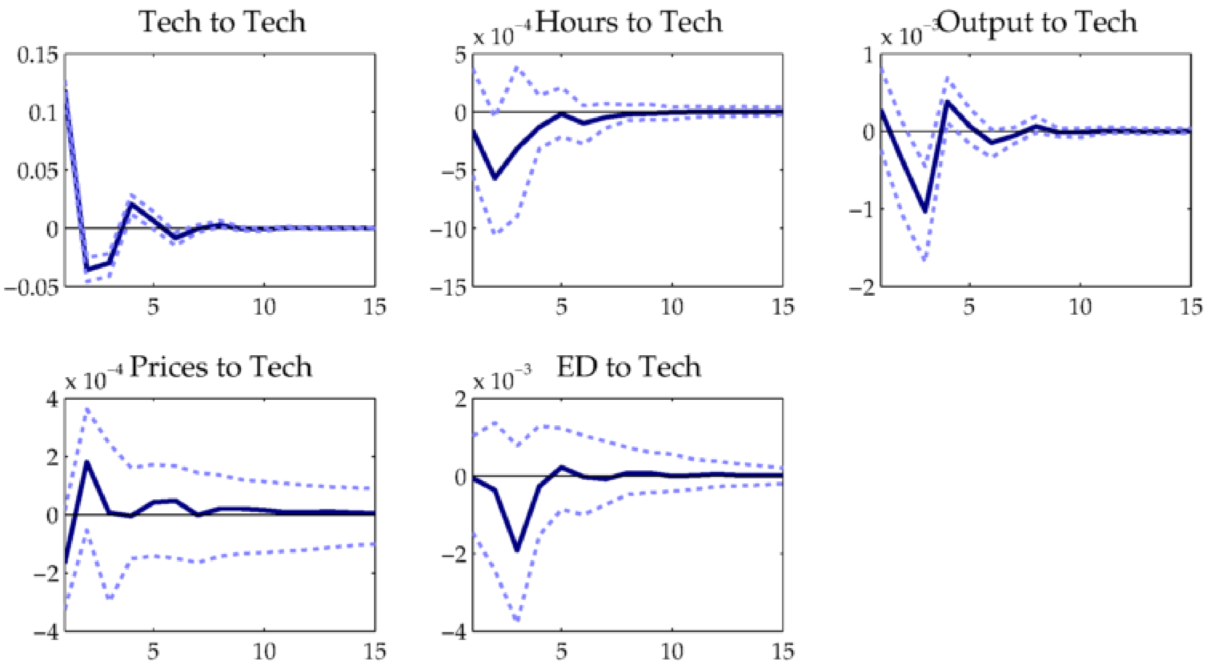


Figure 2.14: Long Run SVAR Impulse Responses to a Technology Shock (Levels)

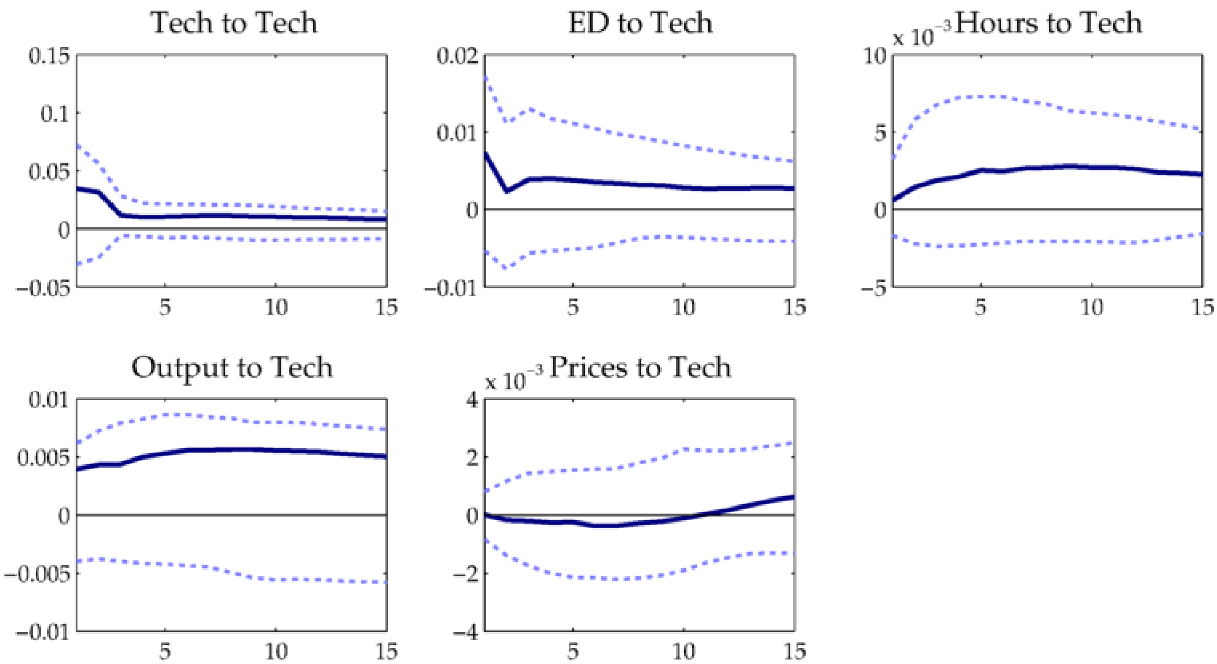


Figure 2.15: Long Run SVAR Impulse Responses to a Technology Shock (First Differenced Hours)

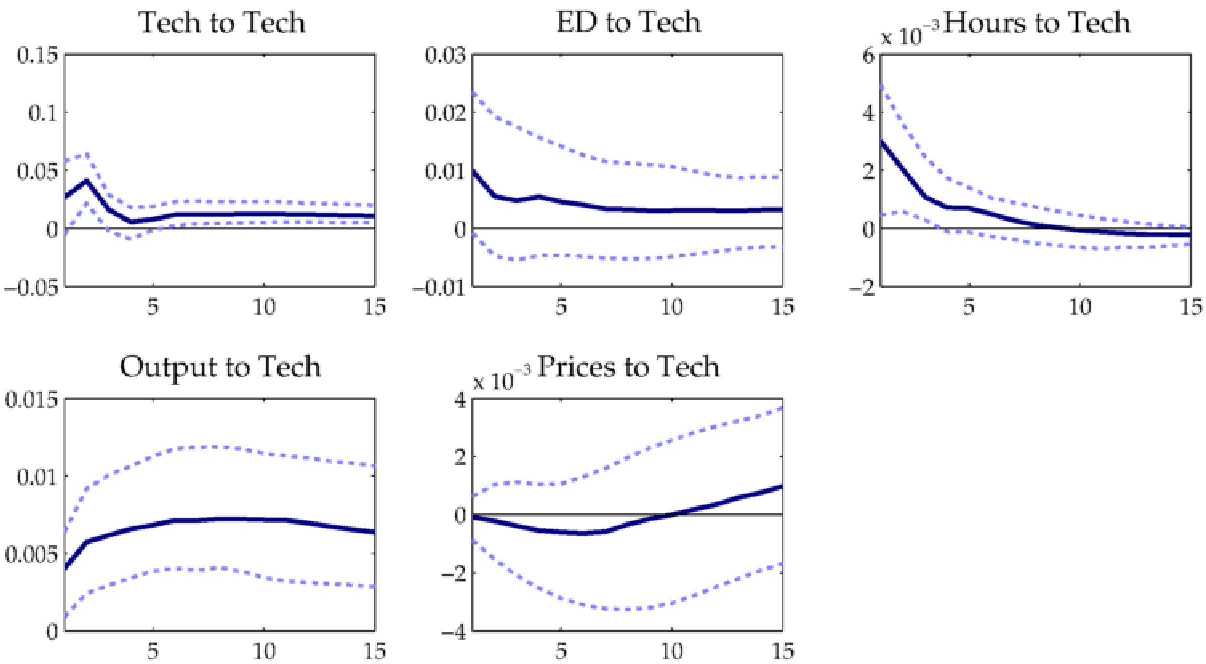


Figure 2.16: Empirical IRFs: Skilled-Specific Technology Shock, TFP Data (Levels)

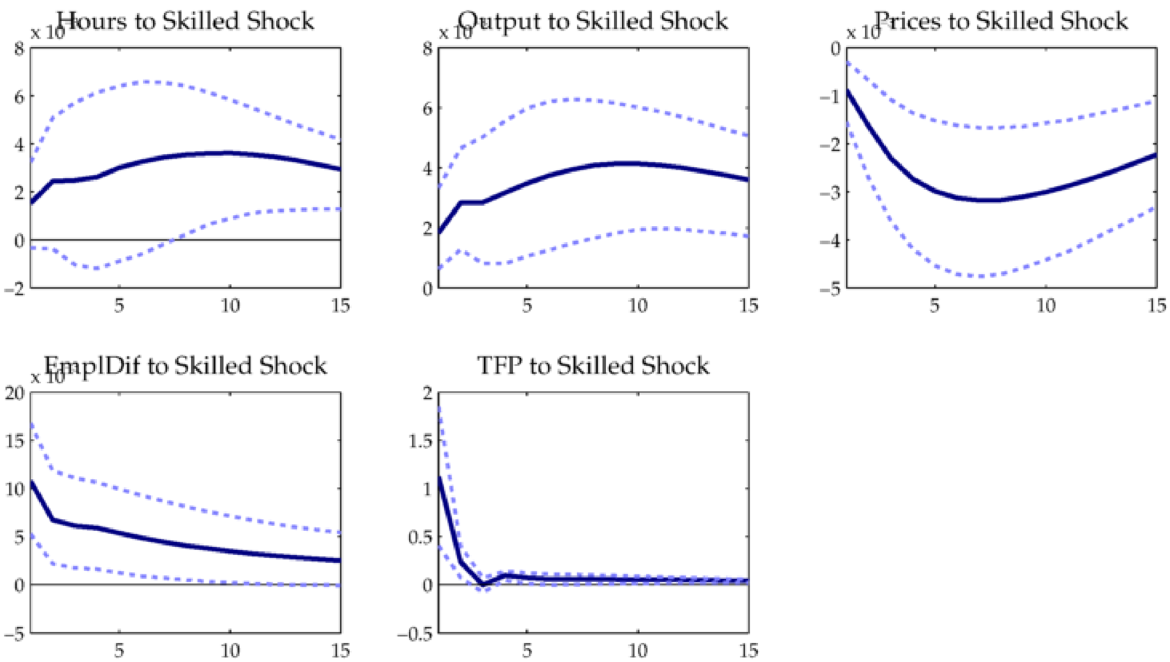


Figure 2.17: Empirical IRFs: Skilled-Specific Technology Shock, TFP Data (First Differenced Hours)

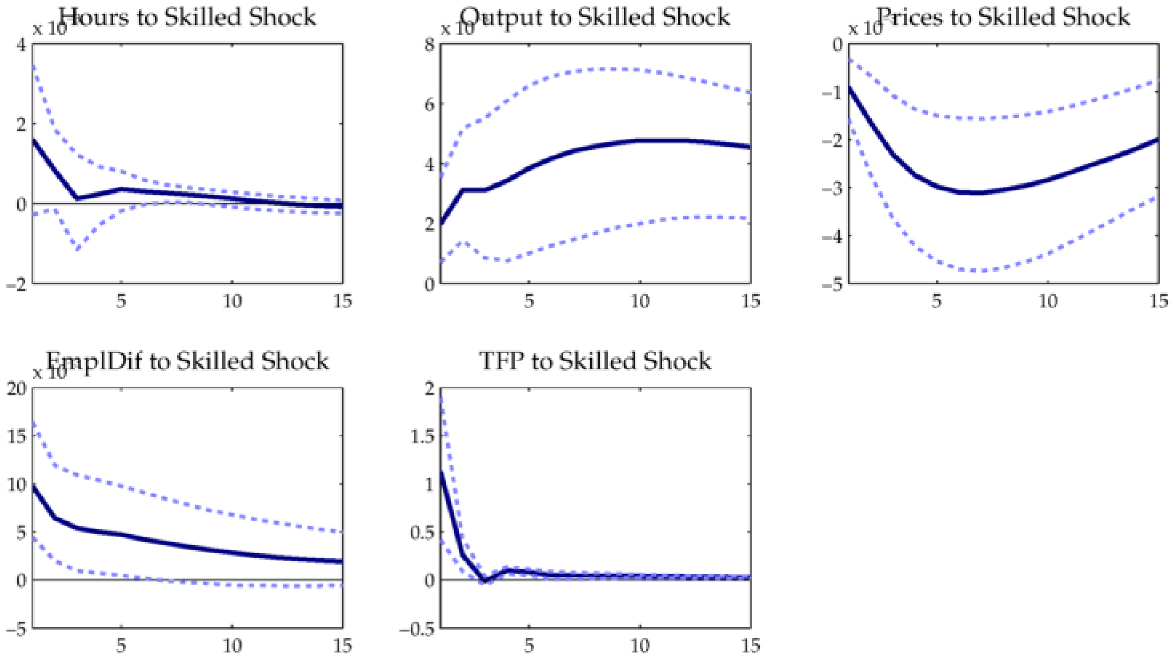


Figure 2.18: Empirical IRFs: Aggregate Technology Shock, TFP Data (Levels)

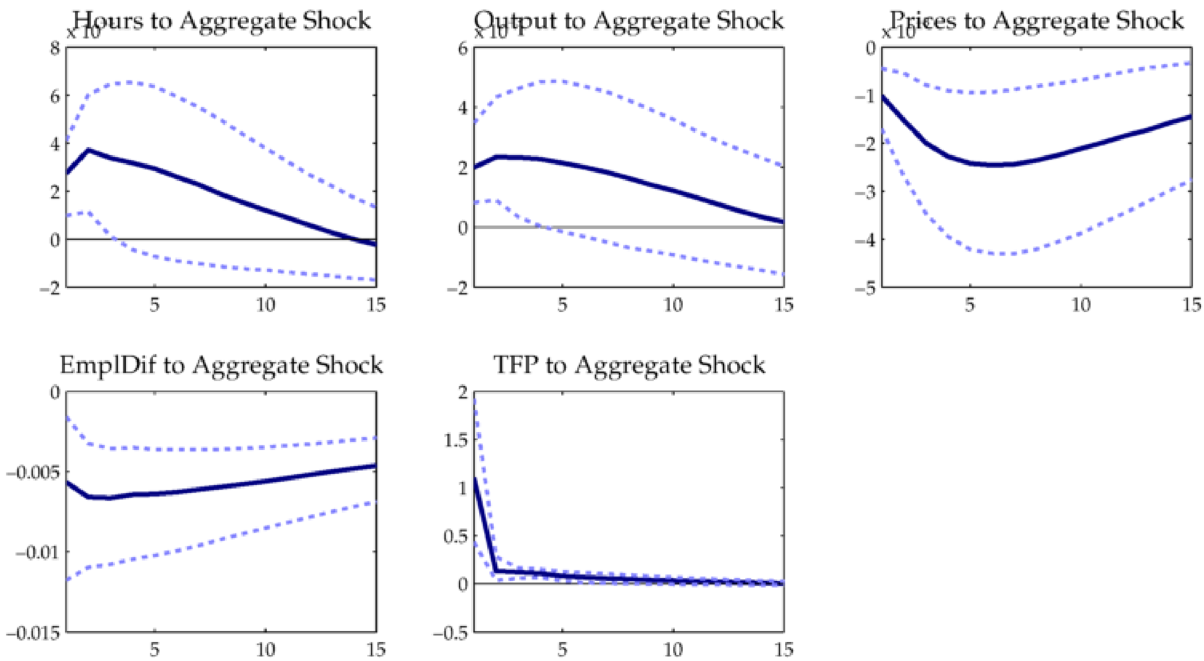
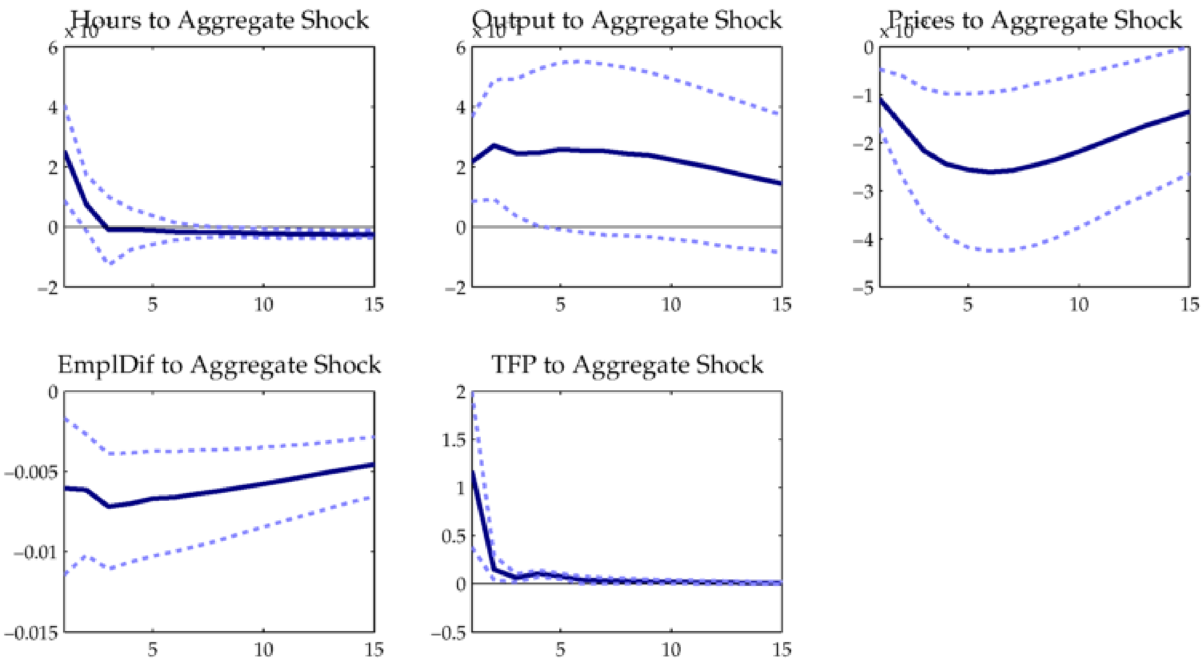


Figure 2.19: Empirical IRFs: Aggregate Technology Shock, TFP Data (First Differenced Hours)



Chapter 3

Effect of an Aggregate Technology Shock in a Factor to Factor VAR

3.1 Introduction

Examining the effects of an aggregate technology shock on economic variables is a very important question. Technological innovations have a tremendous influence on aggregate economic activities. However, there are certain problems associated with estimating the effects of such innovations on the economy especially in the short-run. In this paper, we introduce a novel approach to evaluate the short run effects of technology shocks on the economy. We implement a factor-to-factor augmented vector autoregressive technique (F-FAVAR).We mainly focus on the effects of such a shock on labour market measures given the debate in both the theoretical and empirical macroeconomic literature on the effects of technological innovations on aggregate hours worked.[1]

The factor augmented VAR technique (FAVAR) introduced in [Bernanke et al., 2004] and used to estimate the effects of monetary policy has certain advantage over more traditional VAR methodologies. FAVAR allows us to pick up the effects of the shock of interest on large set of economic variables and does not force a researcher to select only the few "most relevant" economic measures to be included in the VAR. By including factors obtained from a theoretically unbound dataset of economic measures into VAR instead of picking a few measures, a researcher does not have to rigorously justify the choice of variables. As a result, a researcher is able to analyze the effects of the same identified shock on theoretically unbound set of variables.

Following [Bernanke et al., 2004], [Park, 2012] applied a FAVAR to examine the effects of a technology shock. However, unlike a monetary shock that can be represented by the Federal Funds Rate serving as the "impulse variable" in the VAR, the choice of the measure of an aggregate technology shock is more problematic. [Bernanke et al., 2004] includes the Federal Funds rate as the only observable factor along with the latent factors in their VAR. It is, however, much more difficult to come up with an observable factor indicating the changes in aggregate technology.

[Park, 2012] uses the aggregate multifactor productivity as the technological measure in the FAVAR. This measure, as well as any other possible measure, does not perfectly represent the unobserved aggregate technological process as discussed in Chapter 2. TFP are, at most, indirect measures of technological processes that capture their effects on the economy rather than measuring the technology explicitly. None of the TFP measures is safe from non-technological effects being captured. [Rebelo, 2005] indicates that

[1] A curious reader can find the information on the theoretical aspects of the debate in [Rebelo, 2005], while the empirical aspects are discussed in sufficient details in [Chari et al., 2007] and [Oystrakh, 2013].

TFP measured using Solow Residuals can be predicted by military spending and monetary policy. According to [Alexopoulos & Cohen, 2010], both TFP and output per hour variables are affected by various factors other than technological change, while the elimination of the non-technological effects from the TFP measure data is costly and not always successful.

The direct measures of technological process, such as patent and R&D variables, aren't immune to potential criticisms either. The debate found in[Griliches, 1990] about whether the rise in the number of patents is driven by demand, supply or policy factors is still very relevant. While R&D measures have an obvious advantage that a dollar value is attached, they have also several weaknesses. The fact that not every dollar invested in R&D turns out as an invention or an improvement of the current technological state is an obvious limitation of such measure of technology. Moreover, there are certain practical difficulties with measuring the value of R&D especially at frequencies higher than annual.

To circumvent the problem of selecting a variable to represent technological progress in a FAVAR, we propose to augment the FAVAR methodology. When we use FAVAR, we include multiple variables on the response side of the problem and extract factors that can be interpreted as unobserved aggregate economic activity measures, the next step is to repeat the procedure with the variables on the impulse side of the problem. The F-FAVAR methodology, introduced in this study, is a natural extension of the FAVAR methodology. It takes the next step by using factors instead of variables to represent the shock-generating variables besides using the factors to represent the shock-reacting variables.

The advantage of the methodology is that we won't need to select and justify the measure of technology being used. This advantage is especially useful when considering an aggregate technology shock that comes from an unobserved process. Instead of selecting one technological indicator, we include multiple technological indicators both of direct and indirect nature and extract the common factors that represent an unobserved aggregate technological process. Later, these factors are used in a VAR along with the economic activity factors to examine the effects of aggregate technology shocks on the economy.

Hence, the novelty of this study is not just in the implementation of the FAVAR methodology of [Bernanke et al., 2004] to examine the effect of technology shocks on multiple macroeconomic variables. The study goes one step further to extend the FAVAR methodology into a F-FAVAR methodology where the shock variable is also represented by a latent factor. The advantage of this step is that allows avoiding a debate on which of technological measures is the most appropriate to be used in the model. A secondary advantage is the fact that when using the F-FAVAR approach, one does not have to worry about identifying the structural technology shock. All the latent factors used in the model are orthogonal by design and therefore, it is possible to utilize an unrestricted VAR model to determine the responses of the economic variables to a technology shock.

A possible alternative to using a technological latent factor or one of the available measures of technology, is to create a technology index. However, this path would require a researcher to introduce and justify weights assigned to various observed measures of technology.[2] The F-FAVAR on the other hand, would not require such strong assumptions. Another advantage of the F-FAVAR approach is that there is no need to introduce and justify any restrictions in the VAR model in order to identify a technology shock. The technology factor is orthogonal to the macroeconomic factors by construction and hence a shock to such factor is truly a structural technology shock. On the other hand, a technology index does not have to be and most likely is not orthogonal to the latent macroeconomic factors and therefore requires additional assumptions made by a researcher.

There exists an alternative approach to build a technology index by assigning weights empirically instead of using an ad-hoc way. Hence this approach fixes the first problem associated with constructing an index but does not eliminate the fact that a researcher still needs to impose restrictions on the VAR which is not a trivial issue. This approach, described in [Stock & Watson, 1989] and implemented in the context of technology shocks in [Hobijn & Stiroh & Antoniades, 2003] is based on the use of state-space

[2]An example of such index is cross-sectional Bloomberg Innovation Quotient. See `http://www.bloomberg.com/slideshow/2013-02-01/50-most-innovative-countries.html` for more details.

models to determine a common trend growth rate from a group of processes.

There are already indices of technology that are constructed using the state-space model approach. One of them is San Francisco Tech Pulse Index that is constructed using investment into IT sector, consumption of computer hardware and software, employment, and industrial production in the IT sector.[3]. We use this index of technology as a direct measure of technological shocks in a traditional FAVAR model as a robustness check.

The chapter is organized as follows: Section 3.2 briefly summarizes the FAVAR methodology by following [Bernanke et al., 2004] and extends the methodology for F-FAVAR. Section 3.3 introduces the data while leaving a more detailed discussion to the Appendix. We then discuss the results of the F-FAVAR approach. Section 3.4 describes the robustness. Section 3.5 concludes the study.

3.2 Methodology

3.2.1 Two-step FAVAR

We start with a brief introduction of the two-step FAVAR procedure implemented by [Bernanke et al., 2004]. Let X_t denote the set of N endogenous variables with X_{it} denoting an observation of variable $X_i, i = 1 \dots N$ at time $t, t = 1 \dots T$. Note that N can be arbitrary large with no restriction on N being smaller than T.

Let Y_t denote the set of M observable variables, $Y_t \subset X_t, M; N$. Assume that any additional information not captured by Y_t can be represented by unobserved K factors F_t, where $K \ll N$. The joint dynamics are given by:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + \upsilon_t \tag{3.1}$$

Where $\Phi(L)$ is a lag polynomial of order d and υ_t is the error term with mean 0 and covariance matrix Q.

The unobserved factors F_t can be inferred from the informational dataset X_t. X_t is assumed to be related to unobserved and observed factors F_t and Y_t respectively by:

$$X_t' = \Lambda^f F_t' + \Lambda^Y Y_t' + e_t' \tag{3.2}$$

Where Λ^f is a $N \times K$ matrix of loadings and e_t has zero mean but is allowed to have some degree of cross correlation.

The objective is to estimate the system of 3.1 and 3.2. Denote the space spanned by the common components in 3.2 by $C_t = (F_t', Y_t')'$. In the first step of the two step approach, C_t is estimated by extracting the principal components from the information set X_t. The principal components at this point recover the space spanned by both F_t and Y_t. Therefore, it is necessary to uncover $\hat{F}_t$ from the space $\hat{C}_t$ not covered by Y_t.

Depending on the research question and the VAR identification methodology, the way to retrieve $\hat{F}_t$ may be different.[4] To identify a monetary policy shock, [Bernanke et al., 2004] divide the dataset X_t into three groups: the "slow" variables that are assumed to be observed by the monetary authorities when making a decision and that do not react to the shock within the next period. The monetary policy instrument (the federal funds rate) that is the only variable in the set of observed variables Y_t. Finally the remaining part of the dataset X_t consists of the "fast" variables that are assumed to react to the shocks of the slow variables, the monetary policy instruments and to their own shocks within the same period.

Inclusion of more variables into the observed set Y_t is unnecessary and may lead to some complications. Economic variables are often a subject to measurement errors and revisions that may cause

[3]See `http://www.frbsf.org/economic-research/indicators-data/tech-pulse/` for more information.

[4]The description of the [Bernanke et al., 2004] FAVAR identification is introduced due to the fact that a similar approach will be taken in the robustness section 3.4 of this study. However, a different methodology is implemented in the main analysis section.

idiosyncratic disturbance. As it was pointed out in [Park, 2012], all the other relevant variables are observed in the set X_t of "noisy macroeconomic indicators". Moreover, inclusion of more variables into the set Y_t would require unnecessary and most of the time arguable justifications about the choice of one particular variable over another.

After grouping the variables, the following regression equation is estimated:

$$\hat{C}_t = b_{F^s}\hat{F}_t^s + b_Y Y_t + e_t \tag{3.3}$$

Where $\hat{F}_t^s$ stands for the slow moving factors estimated by principal components using the slow subset of X_t. The estimated factors are given by $\hat{F}_t = \hat{C}_t - b_Y Y_t$. This exercise allows to control the part of $\hat{C}_t$ that corresponds to the federal funds rate while still having $\hat{F}_t$ and Y_t correlated since $\hat{F}_t^s$ is correlated with Y_t.

In the second step, Equation 3.1 is formed and estimated recursively which basically implies imposing a Cholesky short run restrictions of the following kind: an innovation to any variable in the VAR has no contemporaneous effect on any other variable preceding it in the VAR ordering. The shock variable treated as the monetary policy innovation is placed last into the VAR system which implies that the latent factors $\hat{F}_t$ do not react to the shock within the same period.

3.2.2 F-FAVAR Approach

Now let's think about the available measures of aggregate technological process. Unfortunately, unlike the federal funds rate, many of them are not safe from measurement errors and are subject to revisions. This is especially true for TFP and R&D variables. Patent variables, on the other hand, suffer from data truncation problems and from the fact that not all the patents are equivalent in terms of their technological significance and need to be weighted according to certain criteria.These facts combined with the observation that there is no perfect or even "best of available" measures of technology motivates the replacement of an observed data Y_t with a latent $F'_{tech,t}$ in 3.2 so that:

$$X'_t = \Lambda^{f_{econ}} F'_{econ,t} + \Lambda^{f_{tech}} F'_{tech,t} + e'_t \tag{3.4}$$

Where $\Lambda^{f_{econ}}$ is a $N \times (K-1)$ matrix of loadings on "economic latent variables", $\Lambda^{f_{tech}}$ is a $N \times 1$ matrix of loadings on the "technological latent variable". e_t has zero mean but allowed to have some degree of cross correlation. $F'_{tech,t}$ is an identified technological factor while $F'_{econ,t}$ are non-technological factors that affect economic variables.

One of the purposes of factor analysis is to reduce the dimension and represent a large dataset in terms of fewer latent factor variables that, presumably, stand for some unobserved process common to the dataset. The idea of the F-FAVAR approach is simple: the unobserved technological process is one of the important forces affecting the economic activity and therefore it is expected to be represented by one of the latent factors affecting the dataset. Thus, instead of adding an observed factor to the K latent factors and then analyzing the effects of a shock to the observed factor on the latent factors and then interpolating the effects on the observables, we take a different route. We identify and compute the factor of interest from the K latent factors and then analyze the effects of a shock to the identified latent factor on other factors and after interpolate them on the economic variables.

Another practical advantage of the F-FAVAR approach is the fact that the latent factors are orthogonal by definition. Given that no other variables besides the K factors are added to the regression, structural shocks can be revealed just from the estimation of equation 3.4 with no further restrictions imposed.

3.2.3 Identification of the Technological Latent Factor

A practical question that arises from the previous discussion is how can we tell which of the K factors stands for the unobserved technological process. As described in [Yong & Pearse, 2013], large datasets that consist of several variables can be reduced by observing groups of variables (i.e. factors). Therefore, factor analysis assembles common variables into descriptive categories. Furthermore, factor loadings provide us with an idea of how much each variable contributes to the given factor. Factors are rotated for better interpretation since unrotated factors are often ambiguous. As a result, a researcher attempts to obtain a structure where each of the observed variables loads "high" on as few factors as possible. Each factor in sch a structure defines a distinct cluster of interrelated variables which facilitates the interpretation of the factors.

This important part of factor analysis seems to be completely missing from the previous FAVAR studies. The researchers would extract a number of factors from economic variables that would define some "economic processes" without attempting to interpret them. The notion of rotating factors towards the interpretable simple structure that is very common to the factor literature seems to be missing in the FAVAR studies. [Bernanke et al., 2004] extract K principal component factors F and imposes a restriction $F'F/T = I_K$, as a result restricting factors against any rotations. This method, as discussed in [Bai & Ng, 2012], identifies space spanned by the columns of F and Λ, where individual columns are not necessarily identified. Therefore, while the total factor space is restricted against rotation, i. e. carries the same set of information, the interpretation of each individual factor may vary with the rotations.

In this study we implement a more economically meaningful technique to identify factors that also satisfy the first set of restrictions introduced in [Bai & Ng, 2012]. This also leads to an exact identification of factors and factor loadings. We start by dividing the observed variables into K groups π. We attempt to make the division as economically meaningful as possible and categorize the variables by what they measure. For example, all the variables that measure output and income will belong to one group, variables that measure the aggregate price level would belong to another. The full list of groups can be found in table 3.2 of appendix 3.6.1. It is important to stress that one of the groups represents the variables that somehow measure unobserved aggregate technological process. It is also important to note that while grouping all the variables into meaningful categories, in the current study, we concentrate on one particular group of variables that measure aggregate technology.

The next step is to transform the $N \times K$ matrix of loadings Λ into a block diagonal matrix that satisfies the first set of [Bai & Ng, 2012] exact identification restrictions.

$$\Lambda = \begin{bmatrix} \pi_1 & 0 & \cdots & 0 \\ 0 & \pi_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \pi_K \end{bmatrix} \tag{3.5}$$

Moreover, the first set of exact identification restrictions determined in [Bai & Ng, 2012] holds under arbitrary permutation of the variable groups. Therefore, matrix $\breve{\Lambda} = \begin{bmatrix} \pi_3 & 0 & \cdots & 0 \\ 0 & \pi_K & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \pi_1 \end{bmatrix}$ is as good as the matrix Λ in 3.5.

In practice, the restrictions-imposing procedure goes as follows, first we estimate the unrotated matrix of loadings Λ_u by the method of principal component. This implies bringing the total communality as close as possible to the total observed variance of the variables [Tryfos, 1998].[5] After estimating the unrotated $N \times K$ matrix of loadings Λ_u, we derive the rotated matrix Λ_r by rotating Λ_u towards the

[5] We define the communality as the part of the observed variance that is explained by the factors, while the variable-specific variance is the part of the observed variance that is not explained by the factors. See [Tryfos, 1998] for more details.

$N \times K$ target matrix Tg, where:

$$Tg = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{3.6}$$

For any given solution of two or more factors there exists an infinite number of alternative orientations of the factors in multidimensional space (rotations) that explain data equally well [Fabrigar et al., 1999]. Thus, instead of sticking to the initial matrix of loading we search the rotation most suitable for our purposes.

We apply orthogonal rotations only in order to preserve the latent variables independence. The transformation of Λ_u into Λ_r is achieved by minimizing the rotation criterion function $c(\Lambda_u(R')^{-1})$ over all feasible matrices R where R is an orthonormal matrix such that $RR' = R'R = I$ so that $\Lambda_u(R')^{-1}$ simplifies to $\Lambda_u R$. The criterion function is the following:

$$c(\Lambda_u R) = \frac{1}{2}|\Lambda_u - Tg|^2 \tag{3.7}$$

After the rotated matrix is achieved, the technological factor is identified as the factor with "high" loadings[6] on the technological variables and "low" loadings on all other variable groups. The remaining $K - 1$ rotated factors can also be economically interpreted for most of the variable groups even though this is not the objective of this study.

The theoretically infinite number of rotations property of the factor analysis is sometimes considered a disadvantage of the method. However, in this study, instead of restricting against the rotations a priori, we make use of them. We search for a rotation that satisfies our research priors and makes most economic sense.

3.3 Results

The study uses 128 variables and indicators that are distributed into 12 groups by the phenomena they measure. A detailed discussion of the variables and the groups can be found in appendix 3.6.1. The variables were stationarized and standardized in order to make factor analysis and VAR analysis meaningful. The observed variables sample size includes 120 quarterly observations.

An important task of factor analysis is to determine the factors' extraction method and the number of factors to be kept. Both of those tasks are subject to debate in the literature.

[Fabrigar et al., 1999] suggests that the maximum likelihood method of factors extraction is superior to the alternative ways unless the multivariate normality is "severely violated". Unfortunately this is the case in this study. The hypothesis of multivariate normal distribution is strongly rejected by multiple tests. Therefore, the maximum likelihood method is inaccessible for this analysis.

The principal component factor extraction method used in this study and the third most popular principal factor (also called principal axis) method often produce very similar results [Fabrigar et al., 1999]. In our case the resulting factors obtained by both cases look very similar in terms of the values of their loadings and variables' communalities.

Now we turn to the next question: how many factors to hold. [Tryfos, 1998] asserts that there is a considerable subjectivity on this matter and in some instances, the number of factors could be hypothesized in advance depending on the research question. For the maximum likelihood factor extraction method there exist several tests that could be used to confirm the optimum number of factors to be kept. However, these tests are unaccessible for the aforementioned reasons.

[6]Given the dimensions of the dataset loadings of 0.32 and higher are considered to be high. [Yong & Pearse, 2013]

The two common ways of determining the number of factors such as a scree plot and a Kaiser criterion, when applied to our data, both suggested 23 factors to be kept. However, both of these methods are imperfect and are subject to criticism.[7] [Bernanke et al., 2004] included 3 factors in their FAVAR analysis and used a 5 factors model for a robustness check. Inclusion of more than 5 factors, according to [Bernanke et al., 2004], did not affect the qualitative nature of the results. [Stock & Watson, 2005] suggest 7 factors to be included for the US postwar macro data. This number exceeds the number of factors that is typically kept in other studies. The remaining factors, even if they are statistically meaningful, may lack any economic meaning.

In this study, given the identification scheme presented in section 3.2 and the presence of 12 variable groups, 12 factors are retained and included in the VAR. Keeping exactly 12 factors significantly simplifies the construction of the target matrix in equation 3.6 and conducting a rotation towards that target. Moreover, 12 factors retained exceed the typical number of factors kept in similar studies, while [Fabrigar et al., 1999] suggest that keeping too few factors is worse than keeping too many.

Retaining 12 factors is preferable to retaining less than 12 factors, given each of the 11 out of the 12 factors have a meaningful economic interpretation.[8] Now we need to justify why we won't keep more than 12 factors as it is suggested by certain criteria. First of all we would need to divide some of 12 variable groups into subgroups in order to identify the technological factors. Moreover, an inclusion of more than 12 factors brings high computational costs at the VAR stage of the analysis and requires a restriction of the number of endogenous lags to one lag. This is due to the high dimensionality it would otherwise create and to software limitations. A model with 12 factors retained explains about 78% of the data variation which is an acceptable amount.

After the 12 latent factors are extracted from the data, the loading matrix is rotated as closely as possible to the block diagonal structure in equation 3.5 while preserving the factors orthogonality. Table 3.6 presents estimated technological measures' loadings used to assemble the latent technology factor. Each of the loadings presents relative importance of a given measure. Slightly higher loadings for the R&D variables can be explained by the fact that these variables dominate the technological group of observed measures.

After the rotated factors are estimated, they are used in an unrestricted VAR. Impulse response-functions to the technological factor are computed for the latent factors first and interpolated to the observable variables using the factors scoring coefficient matrix. 95% confidence intervals are constructed using standard errors from bootstrapped residuals. The resulting impulse-responses to some selected variables are depicted on Figure 3.1.

The results of an aggregate technology shock obtained using the F-FAVAR methodology are similar with the results of other studies including the study of [Park, 2012] that uses a FAVAR methodology. The results are also similar to studies that use long-run restrictions and multi-factor productivity as a technological indicator. Moreover, the results of this chapter complement the results of chapter 2.

The effects of a technology shock on macroeconomic variables obtained using the F-FAVAR approach depicted in Figure 3.1 can be summarized the following way:

1. A positive technology shock increases both potential and real GDP. This result is widely supported by macroeconomic literature and is associated with the rightward shift of the long run aggregate supply curve. Multiple studies starting with [Solow, 1957] examined the effect of improvement in aggregate productivity on output.

2. The unemployment rate first increases marginally and then permanently drops following the improvement in potential GDP.

[7] See [Fabrigar et al., 1999] and [Costello & Osborne, 2005] for details.

[8] One of the variable categories discusses in appendix 3.6.1 includes miscellaneous variables that do not fit any other category. Interpreting the factor, that after rotation towards a target, maximizes loadings of the variables of this particular group is a bad idea.

3. While aggregate wages in the economy rise, average wages react to a technological innovation with an initial drop followed by a gradual increase. As it was discussed in Chapter 2, not all workers equally benefit from a technological innovation, especially if the innovation is skill-specific. In accordance with macroeconomic theory, an increase in the marginal productivity of workers should increase equilibrium wages. After the short run adjustment triggered by structural unemployment, the wage rate of an average worker should start increasing given the increase in the average worker's productivity.

4. The fall in prices after a positive technology shock is in accordance with the theoretical and empirical macroeconomic literature.

5. Figure 3.1 includes three technology and productivity related measures. All of these measures increase following the technology shock.

6. In accordance with the discussion in Chapter 2, the difference in employment of skilled and unskilled workers increase after a positive technology shock.

7. Aggregate hours worked increase initially after a technology shock followed by a decrease in the later periods. Average hours worked, on the other hand, increase steadily throughout the 15-quarter horizon.

3.4 Robustness

We analyze the robustness of the earlier results by conducting the following exercises: First we impose additional restrictions at the VAR stage of our analysis. We also conduct more conventional FAVAR exercises similar to those conducted in [Park, 2012]. We add an observable Y_t variable to the latent factors. This variable will be the technological indicator in the VAR setting. We try various candidates to represent the technological process.

3.4.1 Including an Observable Measure of Technology

In order to apply a conventional FAVAR methodology to examine the effects of a technology shock on economic data we need to select a measure of aggregate technological process and establish a technological shock identification procedure.

There is no consensus about the best observable measure of technology in the literature. In this study we separately use three measures of technology that are common to the literature. The most common way of including technology in the VAR study is to use one of the available measures of factor productivity. In the first FAVAR exercise we use the multifactor productivity variable as the measure of technology. Such a measure was implemented in the study of [Park, 2012]. In the second exercise, we replace the multi-factor productivity with the utilization adjusted total factor productivity introduced in [Fernald, 2012]. The motivation for including of the utilization adjusted total factor productivity measure is simple. Various studies show that unadjusted factor productivity is affected by non-technological process. [Basu et al., 2006] suggest a way of controlling for non technological effects by adjusting the measure for capacity utilization. In the third FAVAR exercise, we include a direct rather than an indirect measure of technology: number of patents granted adjusted by the number of patent claims. This measure was introduced in [Tong & Frame, 1994] and was suggested to be a better way of measuring technology than a raw patent count since not all the patents are equally technologically valuable and the quantity of patent claims is one of the ways of weighting patents by their technological input.[9] Finally, we also use

[9] An alternative way of assigning weights to patents that would indicate their technological significance is to adjust the patent count by the number of citations received. However, the citations received, unlike claims, are subject to the data truncation.

a San Francisco Tech Pulse Index as an aggregate measure of technology. As it was discussed earlier, a technological index is the most obvious alternative to a technological latent factor. Hence, it is necessary to explore its potential in the FAVAR environment.

In order to identify a technological shock we take the following steps. We extract twelve unrotated factors from the dataset X_t. This time we will not apply a rotation towards a target matrix in 3.6. If we were to rotate the latent factors towards that target and add an observable measure of technology as a thirteenth variable in the VAR, we would end up with two collinear measures of technology in the data. Instead we apply first a more conventional *Quartimax* rotation of the factors. The *Quartimax* criterion, unlike the other most common *Varimax* criterion, seeks to maximize the variance of the squared loadings for each variable, and tends to produce factors with high loadings for all variables, while the *Varimax* criterion encourages the detection of factors each of which is related to few variables. [Tryfos, 1998]. The *Quartimax* minimizes the number of factors needed to explain the variation of each variable [Yong & Pearse, 2013]. It is more difficult to interpret such factors given that the procedure seeks for high loadings for all the variables and therefore under different circumstances a *Varimax* criterion that maximizes some of the loadings and minimizes the others, would be preferable.[10] Nevertheless, given the objective of explaining as much of the data variation as possible with twelve factors without a necessity of interpreting each factor separately, the *Quartimax* rotation criterion fits our purposes[11]. It is necessary to mention that *Quartimax* rotation belongs to a class of orthogonal rotation and therefore the independence of the factors is preserved.

When we add an observed factor to the latent factors, constructing impulse-responses from the unrestricted VAR as it was done in section 3.3, is not an option. There is no reason for the observed factor to be independent from the latent factors and therefore it is necessary to impose additional identifying assumptions.

After obtaining the rotated factors we follow the methodology introduced in [Gali, 1999] and implemented for FAVAR analysis in [Park, 2012]. According to this long run identification scheme, the observed technological factor variable is placed first into the VAR and a lower triangular long run restriction matrix is assumed. Therefore, we assume that the technological variable responds only to its own shock in the long run. This assumption is consistent with multiple SVAR studies that examine the effect of a technology shock on economic variables. The SVAR with long run restrictions is estimated using [Blanchard & Quah, 1989] methodology, impulse responses to structural shocks are constructed for the latent factor variables and extrapolated to the observed variables.

It is also possible to apply an identification scheme similar to the one used in [Bernanke et al., 2004] to identify a monetary shock. For that purpose we would need to distinguish between slow and fast variables with respect to our technological indicator, run a regression described in equation 3.3, and then estimate equation 3.1 recursively. Recall that [Bernanke et al., 2004] placed the monetary policy innovation last into the VAR system which implied that the latent factors $\hat{F}_t$ do not react to the shock within the same period. Given that technological innovation requires some degree of economic restructuring, this assumption would apply to a technology shock with even greater certainty than to a monetary policy shock.

Figures 3.2, 3.3, 3.4, and 3.5 depict the impulse responses of some of the observed variables following a technology shock identified using [Gali, 1999] and [Park, 2012] long run restrictions while using multifactor productivity, utilization-adjusted total factor productivity, total number of patent claims, and the San Francisco Tech Pulse index as the observed measure of technology respectively. Figures 3.6, 3.7, 3.8, and 3.9 show the IRF results while applying [Bernanke et al., 2004] short run restrictions instead. The observed measures of technology in the above exercise are the two commonly used productivity

[10][Fabrigar et al., 1999] claims that *Varimax* rotation criterion has been generally regarded as the best orthogonal rotation criterion and has been most widely applied.

[11]In the case of the San Francisco Tech Pulse Index, convergence was not achieved with *Quartimax* and therefore *Varimax* was used instead.

measures in the economic literature, the measure that was used in Chapter 2, and the Tech Pulse Index as a potential alternative to the technological factor. Figures 3.10 to 3.13 compare IRFs derived using the simple FAVAR methodology with long run restrictions with the F-FAVAR approach results. Figures 3.14 to 3.17 compare IRFs derived using the simple FAVAR methodology with short run restrictions with the IRFs derived using F-FAVAR methodology.

When we examine conventional FAVAR results with different observable measures of technology we sometimes find the responses that contradict well-established facts on technological shocks. For example, when FAVAR with [Gali, 1999] and [Park, 2012] long run restrictions are used, potential GDP decreases following a technology shock with all the four technological measures. Real GDP also decreases in three out of four cases (Figures 3.2, 3.3, 3.4, and 3.5). When a conventional FAVAR with [Bernanke et al., 2004] short run restrictions is implemented, potential and real GDP also decrease after a technology shock in most of the cases. The only observable measure of technology that, when used in FAVAR with short-run restrictions, causes the potential and real GDP to increase is the patent measure (Figure 3.8). However, the behaviour of prices in that case contradicts the prior expectations.

We conclude the F-FAVAR results appear more consistent then the results of both FAVARs with short run and FAVAR with long run restrictions and an observed measure of technology. The consistency of the responses is judged by the behaviour of macroeconomic variables for which we have strong beliefs on how they should react to a technology shock. For example, we expect both real and potential GDP to increase and we expect aggregate prices to fall. Furthermore, we expect the observed technology and productivity metrics to react positively to a technology shock. It appears that the technology shock introduced into the system through a technology factor leads to the most consistent results. It is necessary to mention that one should not expect the responses of all the variables in the VAR to correspond perfectly with theory given its reliance on the data. However, the results from all these robustness tests seem to suggest that the F-FAVAR methodology are credible and reliable in spite of all the caveats associated with the VAR methodology.

3.5 Conclusion

In the this study we introduced the F-FAVAR approach to analyze the effects of an aggregate technology shock on macroeconomic variables. The approach allows a researcher to use multiple measures of the impulse process and even a greater quantity of measures of a response process simultaneously to extract robust and theoretically consistent impulse-responses.

The advantage of this approach over the regular SVAR model is the ability to include into the model and to examine a much greater and theoretically unlimited information set and hence to avoid omitted variable bias in the model. The advantage of the F-FAVAR approach over simple FAVAR approach especially when technology shocks are concerned, is the absence of a necessity to select and justify one single observed measure of the impulse process. An additional advantage is the fact that there is no need to impose VAR restrictions that often require very strong and sometimes are arguable assumptions from a researcher. The results of the F-FAVAR methodology also look more plausible than the results of a FAVAR approach that used the San Francisco Tech Pulse Index. Nonetheless it is necessary to mention that the index itself is mainly based on factors affecting only one industry (the IT sector) of the economy and does not encompass all the aspects of aggregate technological growth.

3.6 Chapter 3 Appendices

3.6.1 The Data

Table 3.3 lists all the variables and indicators used in the study. Variables are collected from various sources including NBER, FRED, BLS, R&D Satellite Account, US Census Bureau, Yahoo Finance, and utilization-adjusted quarterly-TFP series for the U.S. Business Sector, produced by John Fernald. The time span of the variables is from 1976 to 2006. Whenever it was possible, quarterly data was used. Variables with higher frequencies were averaged by quarters. Variables with lower frequencies were assigned identical values for each quarter within the time period divided by the number of quarters in the period.

Before presenting the list of variables and economic indicators let us examine the legend of Table 3.3. The first two columns of table 3.3 are the following: variable name and the source of the data. Table 3.1 lists the sources from which the data was retrieved along with their short notations.

Table 3.1: Data Source

Source	Short Notation	Comment
Federal Reserve Economic Data	FRED	The data retrieved from FRED originates from various sources including U.S. Department of Commerce: Bureau of Economic Analysis, Federal Reserve Bank of St. Louis Economic Research Division, U.S. Department of Labor: Bureau of Labor Statistics, U.S. Congress: Congressional Budget Office, U.S. Department of the Treasury: Bureau of the Fiscal Service, Federal Reserve Bank of Chicago, Board of Governors of the Federal Reserve System and others
OECD statistics	OECD	
U. S. Bureau of Labor Statistics	BLS	Retrieved directly from BLS rather than from FRED
A Quarterly, Utilization-Adjusted Series on Total Factor Productivity	Fernald	Various quarterly TFP measures assembled by John Fernald. See [Fernald, 2012] for data description.
Yahoo Finance	Yahoo	
2010 R&D Satellite Account	R&D	
Source: U.S. Census Bureau, Current Population Survey	CB	
NBER Patent Citation Data	NBER	The latest edition of the data was downloaded directly from Bronwyn Hall web page. See [Hall et al., 2001b] for data description

Column 3 of table 3.3 shows the original data frequency (a - annual, q - quarterly, m - monthly), column 4 indicates whether the original data series were seasonally adjusted (yes/no). Column 5 of table 3.3 represents the variable's type (v - value , i - index, % - percent rate). Column 6 contains a natural log indicator (y - natural log was taken from the original value in levels, n - natural log was not taken). Column 7 indicates other transformation applied in order to make data stationary (fd - first difference of the level/log-level was extracted, tr - time trend from the level/log-level was removed, fdtr - first difference of the level/log-level value was extracted and time trend was removed). Note that each of the variables in table 3.3, after passing the manipulations described in the table was standardized by subtracting its mean and dividing by its standard deviation in order to eliminate the scale effect due to the fact that the variables are measured in different units and make the principal components analysis meaningful.[12]

Column 8 of table 3.3 introduces the category that each variable belongs to depending on what the variable measures. Refer to table 3.2 for the groups' description. Finally, column 9 indicates whether a variable is slow fast or coincident with respect to the unobserved technological process (f - fast, s - slow, c - coincident).

Table 3.2: Variables' Categories

Short Notation	Category Description
i	GDP and Income measures
p	Measures of the prices level
b	Economic and business environment measures
r	Interest rates
m	Monetary aggregates
t	Technological measures
e	Employment/unemployment rates, levels, indices, duration; aggregate and average hours worked
w	Salaries, wages, and compensation
d	Employment and wage premia for defined skilled workers †
f	Factors of production measures (costs, profits, shares, capacity utilization)
g	Fiscal policy and government activity measures (federal, state/municipal, and combined levels)
o	Other (Variables not attributed to any of the above groups)

† Refer to table 3.5 for the category "d" measures' definitions.

Table 3.3: **Variables and Indicators**

Variable Name	**Source**	**Freq.**	**Type**	**Seas.**	**Ln**	**Tr-n**	**Cat.**	**Slow**
Gross Domestic Product per Capita	FRED	q	v	a	y	fd	i	s
Real Gross Domestic Product per Capita	FRED	q	v	a	y	fd	i	s
Gross Domestic Product	FRED	q	v	a	y	fd	i	s
Real Gross Domestic Product	FRED	q	v	a	y	fd	i	s
Real Potential Gross Domestic Product	FRED	q	v	a	y	fd	i	s
Gross National Factor Income	FRED	q	v	a	y	fd	i	s
Net National Factor Income	FRED	q	v	a	y	fd	i	s
Continued on next page								

[12]Given that principal components factor procedure implies making the communality as close as possible to the observed variances, if variables had different variables, then those with higher variances would pre-dominate those with the smaller ones. See [Tryfos, 1998] for more details.

Table 3.3 – continued from previous page

Variable Name	Source	Freq.	Type	Seas.	Ln	Tr-n	Cat.	Slow
Business Sector: Real Output Per Hour of All Persons	FRED	q	i	a	y	fd	i	s
Business Sector: Real Output	FRED	q	i	a	y	fd	i	s
Business Sector: Real Output Per Person	FRED	q	i	a	y	fd	i	s
Non-financial Corporations Sector: Real Output	FRED	q	i	a	y	fd	i	s
Non-financial Corporations Sector: Real Output Per Hour	FRED	q	i	a	y	fd	i	s
Non-financial Corporations Sector: Real Output Per Person	FRED	q	i	a	y	fd	i	s
Industrial Production Index	FRED	m	i	a	y	fd	i	s
Industrial Production: Manufacturing (NAICS)	FRED	m	i	a	y	fd	i	s
Real Personal Income	FRED	m	v	a	y	fd	i	s
Business Sector: Implicit Price Deflator	FRED	q	i	a	y	fdtr	p	s
Non-financial Corporations Sector: Implicit Price Deflator	FRED	q	i	a	y	fdtr	p	s
Consumer Prices Index: All Items	OECD	q	i	a	y	fd	p	s
Consumer Price Index for All Urban Consumers: All Items	FRED	m	i	a	y	fd	p	s
Producer Price Index: Finished Goods	FRED	m	i	a	y	fd	p	s
GDP Deflator (Nominal GDP/Real GDP)	FRED	q	v	a	y	fdtr	p	s
Non-financial Corporate Business Profits Before Tax	FRED	q	v	a	y	fd	b	s
Change in Real Private Inventories	FRED	q	v	a	n	l	b	f
Business Sector: Unit Labor Cost	FRED	q	i	a	y	fd	f	s
Non-financial Corporations Sector: Profits	FRED	q	i	a	y	fd	b	s
Manufacturing Employment: Future Tendency	OECD	q	i	a	n	l	b	f
Manufacturing Capacity Utilization (Survei)	OECD	q	%	a	y	l	f	s
Manufacturing Confidence Indicator	OECD	q	i	a	n	l	b	f
Real Investment	FRED	q	v	a	y	fd	b	s
Commercial and Industrial Loans, All Commercial Banks	FRED	m	v	a	y	fd	b	f
ISM Manufacturing: PMI Composite Index	FRED	m	i	a	y	l	b	f
ISM Manufacturing: New Orders Index	FRED	m	i	a	y	l	b	f
Continued on next page								

Table 3.3 – continued from previous page

Variable Name	Source	Freq.	Type	Seas.	Ln	Tr-n	Cat.	Slow
Chicago Fed National Activity Index	FRED	m	i	n	n	l	b	f
Chicago Fed National Financial Conditions Index	FRED	m	i	n	n	tr	b	f
Chicago Fed National Activity Index: Production and Income	FRED	m	i	n	n	l	b	f
Chicago Fed Nat. Activity Index: Sales, Orders & Inventories	FRED	m	i	n	n	l	b	f
Long-Term Interest Rates	OECD	q	%	n	y	fd	r	f
Short-Term Interest Rates	OECD	q	%	n	y	fd	r	f
Effective Federal Funds Rate	FRED	m	%	n	y	fd	r	f
1-Year Treasury Constant Maturity Rate	FRED	m	%	n	y	fd	r	f
10-Year Treasury Constant Maturity Rate	FRED	m	%	n	y	fd	r	f
5-Year Treasury Constant Maturity Rate	FRED	m	%	n	y	fd	r	f
3-Month Treasury Bill: Secondary Market Rate	FRED	m	%	n	y	fd	r	f
M1 Money Stock	FRED	m	v	a	y	fd	m	f
Real M2 Money Stock	FRED	m	v	a	y	fd	m	f
Monetary Base; Total	FRED	m	v	n	y	fd	m	f
Real M1 Money Stock	FRED	m	v	a	y	fd	m	f
Business Sector: Hours of All Persons	FRED	q	i	a	y	fd	e	s
Business Sector: Employment	FRED	q	i	a	y	fd	e	s
Business Sector: Average Weekly Hours	FRED	q	i	a	y	fd	e	s
Non-financial Corporations Sector: Employment	FRED	q	i	a	y	fd	e	s
Non-financial Corporations Sector: Average Weekly Hours	FRED	q	i	a	y	fd	e	s
Non-financial Corporations Sector: Hours Worked	FRED	q	i	a	y	fd	e	s
Average Weekly Hours Manufacturing*	FRED	m	v	a	y	fd	e	s
Civilian Employment	FRED	m	v	a	y	fd	e	s
Civilian Labor Force	FRED	m	v	a	y	fd	e	s
All Employees: Manufacturing	FRED	m	v	a	y	fd	e	s
All Employees: Total Non-Farm	FRED	m	v	a	y	fd	e	s
Average (Mean) Duration of Unemployment	FRED	m	v	a	y	l	e	s
Civilian Unemployment Rate	FRED	m	%	a	y	tr	e	s
All Employees: Government	FRED	m	v	a	y	fd	e	s
All Employees: Private	BLS	m	v	a	y	fd	e	s

Continued on next page

Table 3.3 – continued from previous page

Variable Name	Source	Freq.	Type	Seas.	Ln	Tr-n	Cat.	Slow
Average Weekly Hours*	BLS	m	v	a	y	tr	e	s
Aggregate Weekly Hours*	BLS	m	v	a	y	fd	e	s
Unemployment Level: Job Losers	FRED	m	v	a	y	l	e	s
Unemployment Rate: Job Losers	FRED	m	%	a	y	tr	e	s
Median Duration of Unemployment	FRED	m	v	a	y	l	e	s
(Seas) Employment-Population Ratio	BLS	m	%	a	y	fd	e	s
Business Sector: Compensation Per Hour	FRED	q	i	a	y	fd	w	s
Business Sector: Compensation	FRED	q	i	a	y	fd	w	s
Business Sector: Real Compensation Per Hour	FRED	q	i	a	y	fd	w	s
Non-financial Corporations Sector: Hourly Compensation	FRED	q	i	a	y	fd	w	s
Non-financial Corporations Sector: Real Hourly Compensation	FRED	q	i	a	y	fd	w	s
Total Private Average Hourly Earnings*	BLS	m	v	a	y	fdtr	w	s
Total Private Real Average Weekly Earnings*	BLS	m	v	a	y	fd	w	s
Total Private Average Weekly Earnings*	BLS	m	v	a	y	fd	w	s
Total Private Real Average Hourly Earnings*	BLS	m	v	a	y	fd	w	s
Total Private Aggregate Weekly Payrolls*	BLS	m	v	a	y	fd	w	s
Business Sector: Labor Share	FRED	q	i	a	y	fd	f	s
Non-financial Corporations Sector: Labor Share	FRED	q	i	a	y	fd	f	s
Non-financial Corporations Sector: Unit Profits	FRED	q	i	a	y	fd	f	s
Non-financial Corporations Sector: Total Unit Costs	FRED	q	i	a	y	fd	f	s
Non-financial Corporations Sector: Unit Labor Costs	FRED	q	i	a	y	fd	f	s
Utilization of capital and labor	Fernald	q	d %	n	n	tr	f	s
Business Sector: Capital input	Fernald	q	d %	n	n	tr	f	s
Business Sector: Capital's Share of Income	Fernald	q	%	n	n	n	f	s
Capacity Utilization: Total Industry	FRED	m	%	a	y	tr	f	s
Real Gov-t Consumption Expenditures & Gross Investment	FRED	q	v	a	y	fd	g	s
Federal Government Budget Surplus or Deficit	FRED	q	v	a	n	tr	g	s
Government Total Expenditures	FRED	q	v	a	y	fdtr	g	s

Continued on next page

Table 3.3 – continued from previous page

Variable Name	Source	Freq.	Type	Seas.	Ln	Tr-n	Cat.	Slow
Federal Government Total Public Debt	FRED	q	v	a	y	fdtr	g	s
Government Current Receipts	FRED	q	v	a	y	fd	g	s
Federal Government Non-defense Gross Investment	FRED	q	v	a	y	fd	g	s
Business Sector: TFP	Fernald	q	d %	n	n	l	t	c
Utilization-Adjusted TFP	Fernald	q	d %	n	n	l	t	c
Business Sector: Labor Productivity	Fernald	q	d %	n	n	l	t	c
Output Per Hour	OECD	q	d %	n	n	l	t	c
Real National Defense Consumption Expenditures: R&D	FRED	q	v	a	y	fd	t	c
San Francisco Tech Pulse	FRED	m	i	a	y	fdtr	t	c
Real Total R&D Investment	R&D	a	v	n	y	fd	t	c
Real Private R&D Investment	R&D	a	v	n	y	fd	t	c
Real IT R&D Investment	R&D	a	v	n	y	fd	t	c
Real Public R&D Investment	R&D	a	v	n	y	fd	t	c
Real Academia R&D Investment†	R&D	a	v	n	y	fd	t	c
Real Total R&D Performance	R&D	a	v	n	y	fd	t	c
Real Private R&D Performance	R&D	a	v	n	y	fd	t	c
Real Public R&D Performance	R&D	a	v	n	y	fd	t	c
Real Academia R&D Performance†	R&D	a	v	n	y	fd	t	c
Total Number of Patents Granted	NBER	d	v	n	y	fd	t	c
Total Number of Claims in the Patents Granted	NBER	d	v	n	y	fd	t	c
Claims to Granted Patents Ratio	NBER	d	v	n	y	fd	t	c
Private Non-Farm Net Multifactor Productivity	BLS	a	v	n	y	fd	t	c
Skilled to Unskilled Employment Ratio, Scheme C‡	BLS	m	v	a	y	fd	d	s
Skilled to Unskilled Employment Difference, Scheme C‡	BLS	m	v	a	n	fd	d	s
Skilled to Unskilled Employment Ratio, Scheme A‡	CB	a	v	n	y	fd	d	s
Skilled to Unskilled Employment Difference, Scheme A‡	CB	a	v	n	n	fd	d	s
Skilled to Unskilled Employment Ratio, Scheme B‡	CB	a	v	n	y	fd	d	s
Skilled to Unskilled Employment Difference, Scheme B‡	CB	a	v	n	n	fd	d	s
Skilled to Unskilled Wage Ratio, Scheme A‡	CB	a	v	n	y	fd	d	s
Skilled to Unskilled Wage Difference, Scheme A‡	CB	a	v	n	y	fd	d	s

Continued on next page

Table 3.3 – continued from previous page

Variable Name	Source	Freq.	Type	Seas.	Ln	Tr-n	Cat.	Slow
Skilled to Unskilled Wage Ratio, Scheme B‡	CB	a	v	n	y	fd	d	s
Skilled to Unskilled Wage Difference, Scheme B‡	CB	a	v	n	y	fd	d	s
Real National Defense Consumption Expenditures	FRED	q	i	a	y	fd	o	s
Housing Starts: Total: New, Privately Owned	FRED	m	v	a	y	fd	o	f
Spot Oil Price: West Texas Intermediate	FRED	m	v	n	y	fd	o	f
New Private Housing Units Authorized by Building Permits	FRED	m	v	a	y	fd	o	f
S&P 500 Dividend Adjusted Close Price (GSPC)	Yahoo	m	v	n	y	fd	b	f

* Production, Nonsupervisory Employees
† Refer to table 3.4 for variables' definitions
‡ Refer to table 3.5 for variables' definitions

Some variables in table 3.3 require special attention. The following two tables will define the academia R&D variables and the employment/wage premia variables respectively. Table 3.4 will show how the academic R&D investment and performance variables were generated. Table 3.5 explains three different approaches taken in order to distinguish skilled from unskilled workers when constructing skilled employment and skilled wage premia variables.

Table 3.4: Definition of Academia R&D Variables

Variable	Description
Real Academia R&D Investment	The variable combines real private investment in R&D of universities and colleges with real public investment in R&D of universities and colleges. Source: US 2010 R&D Satellite Account, Table 2.2
Real Academia R&D Performance	The variable combines real private output of R&D of universities and colleges with the real private output of federally funded R&D centers' in universities and colleges with real public output of R&D of universities and colleges and with real public output of federally funded R&D centers' in universities and colleges. Source: US 2010 R&D Satellite Account, Table 3.2

Table 3.5: Definition of Skilled and Unskilled Workers

Scheme	Skilled	Unskilled
A	Workers with a bachelour's degree or higher	All other non-institutional workers of age 18+
B	Workers with a college diploma or higher	All other non-institutional workers of age 18+
C	Workers with the following occupations: information, financial activities, business services, professional, health, education	Total non-farm employment minus skilled as defined in scheme C

Table 3.6: Rotated Technology Factor Loadings

Measure of Technology	Loading
Business Sector: TFP	0.3209
Utilization-Adjusted TFP	0.3875
Business Sector: Labour Productivity	0.3890
Output per Hour	0.3178
Real National Defense Consumption Expenditures: R& D	0.2944
San Francisco Tech Pulse	0.1901
Total Number of Patents Granted	0.3327
Total Number of Claims in Patents Granted	0.3441
Claims to Granted Patents Ratio	0.1499
Real Total R&D Investment	0.8237
Real Private R&D Investment	0.6403
Real IT R&D Investment	0.3907
Real Public R&D Investment	0.5241
Real Academia R&D Investment	0.4484
Real Total R&D Performance	0.8203
Real Private R&D Performance	0.7630
Real Public R&D Performance	0.5634
Real Academia R&D Performance	0.4211
Private Non-Farm Net Multifactor Productivity	0.3212

3.6.2 Figures

Figure 3.1: F-FAVAR Impulse-Responses to Technological Shock, Bootstrapped 95% CI

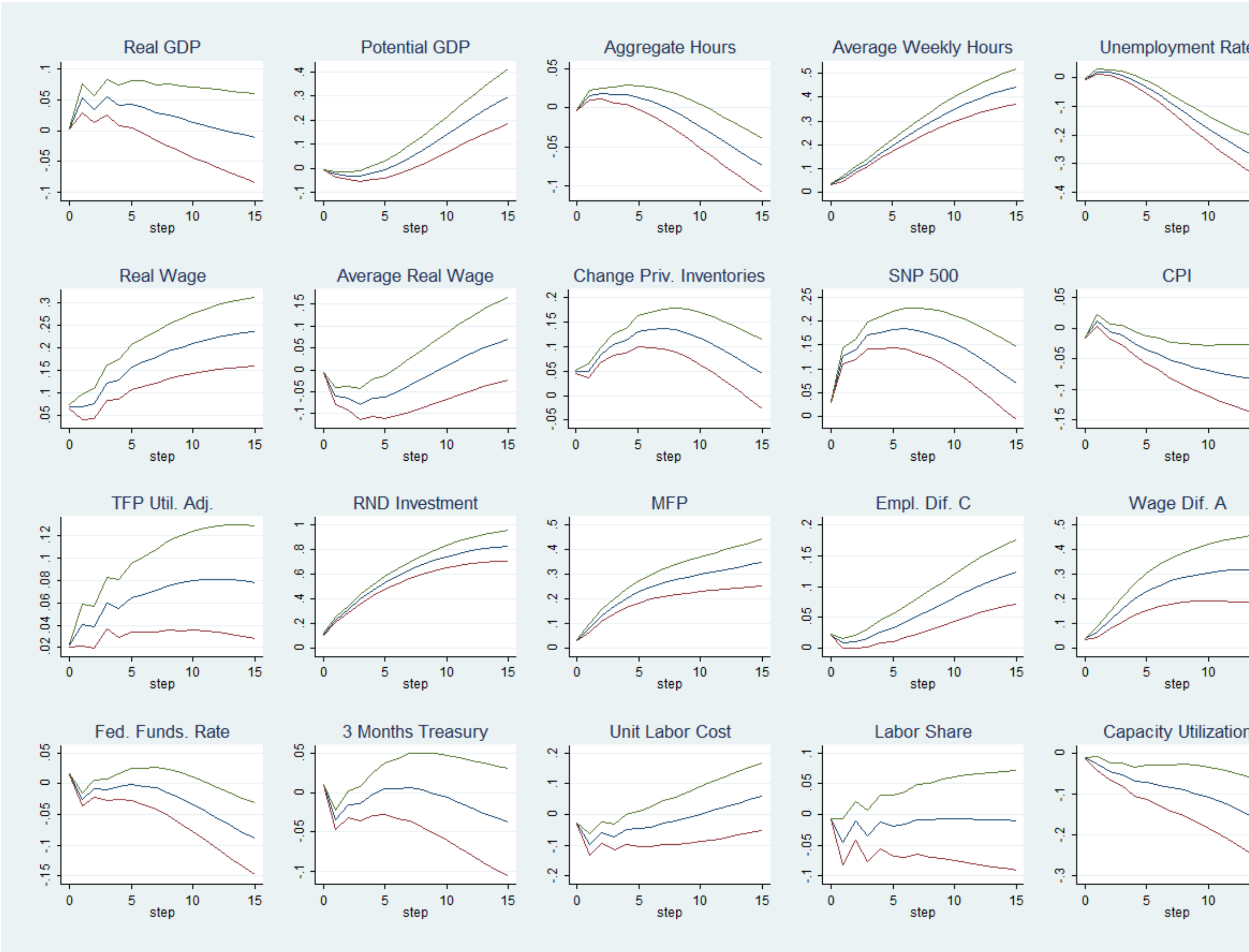


Figure 3.2: FAVAR Impulse-Responses to Technological Shock Using Multifactor Productivity Measure and Long Run SVAR Restrictons, Bootstrapped 95% CI

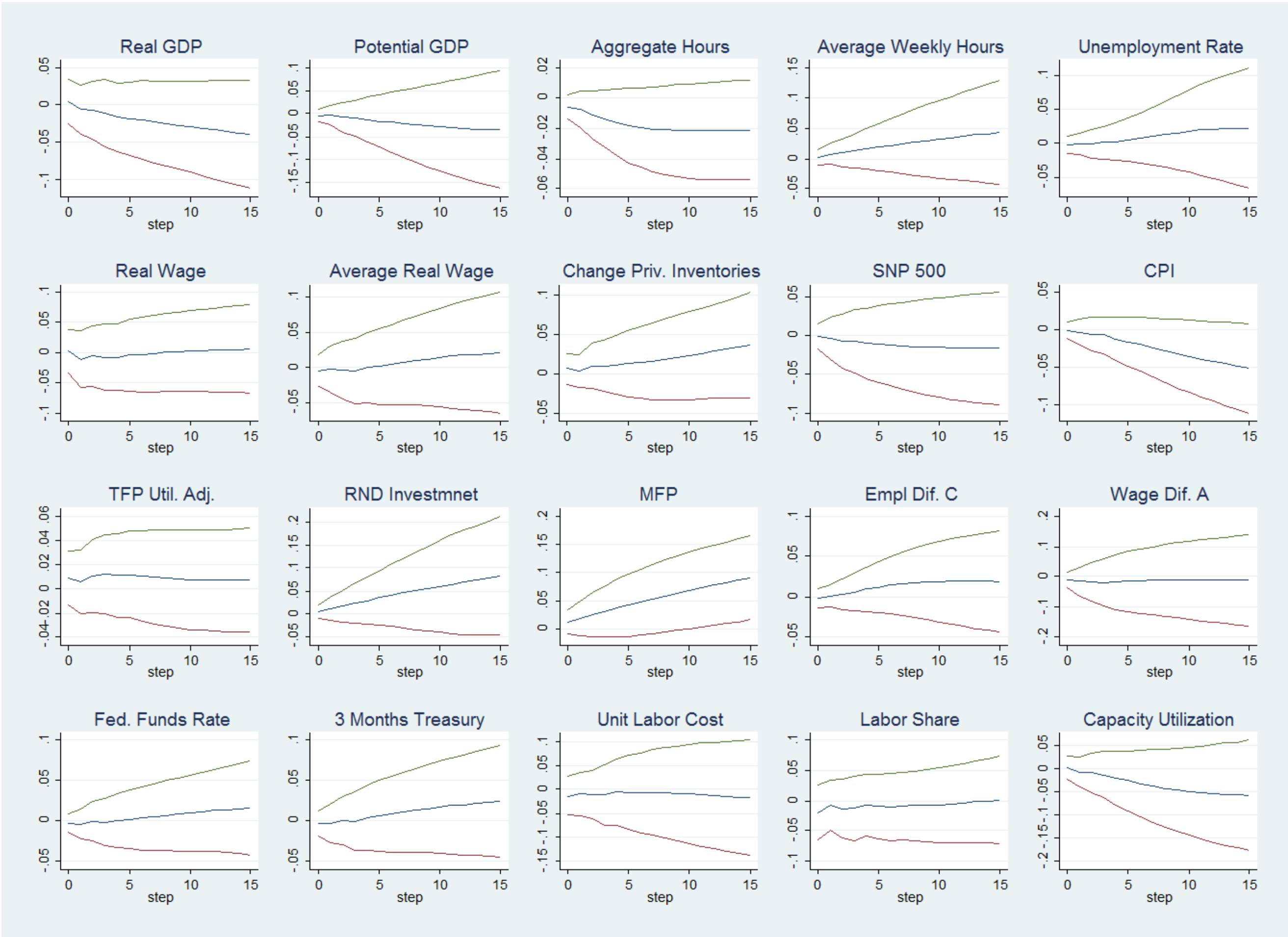


Figure 3.3: FAVAR Impulse-Responses to Technological Shock Using Utilization Adjusted Total Factor Productivity Measure and Long Run SVAR Restrictons, Bootstrapped 95% CI

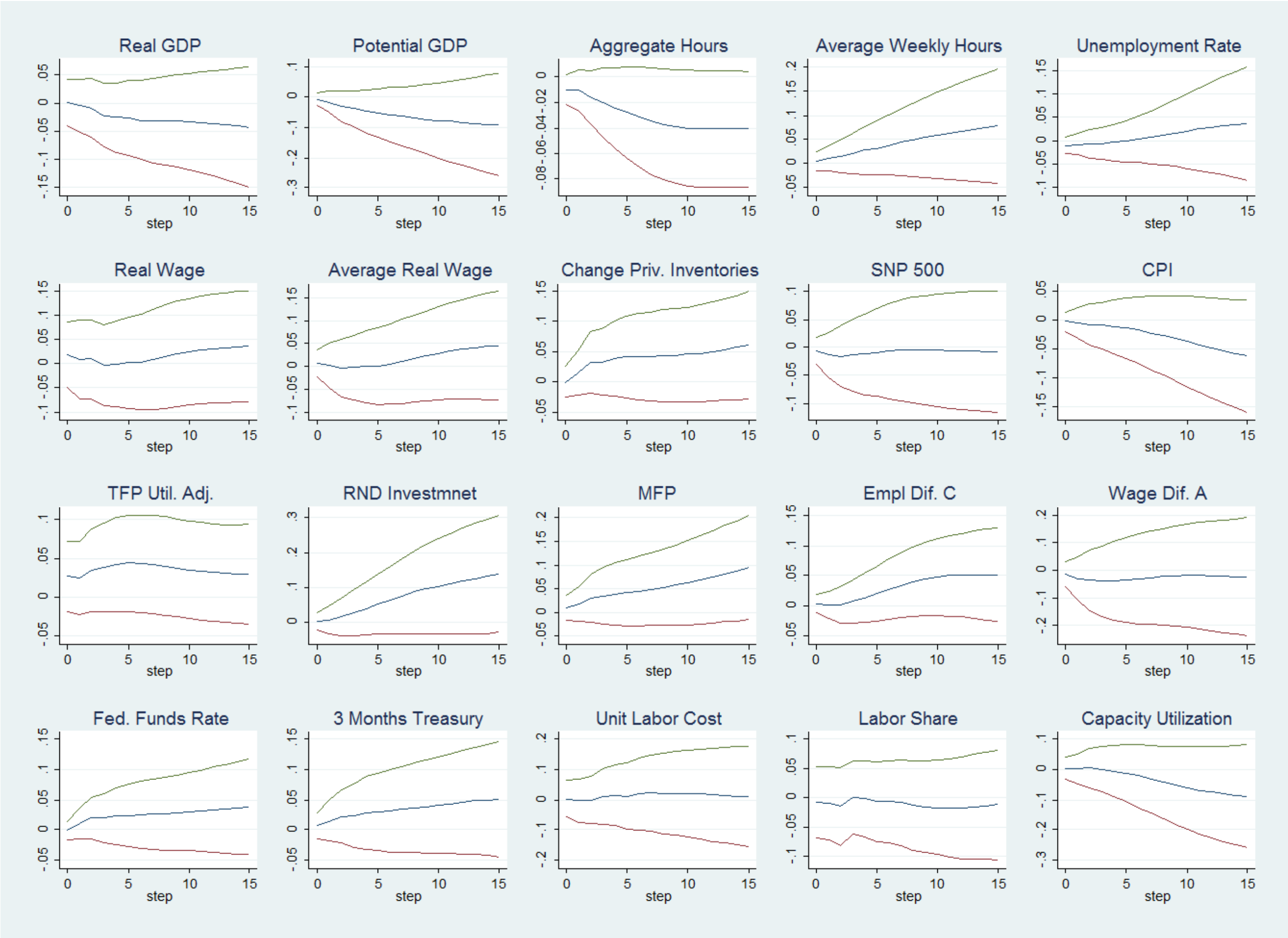


Figure 3.4: FAVAR Impulse-Responses to Technological Shock Using Claim Adjusted Patent Measure and Long Run SVAR Restrictons, Bootstrapped 95% CI

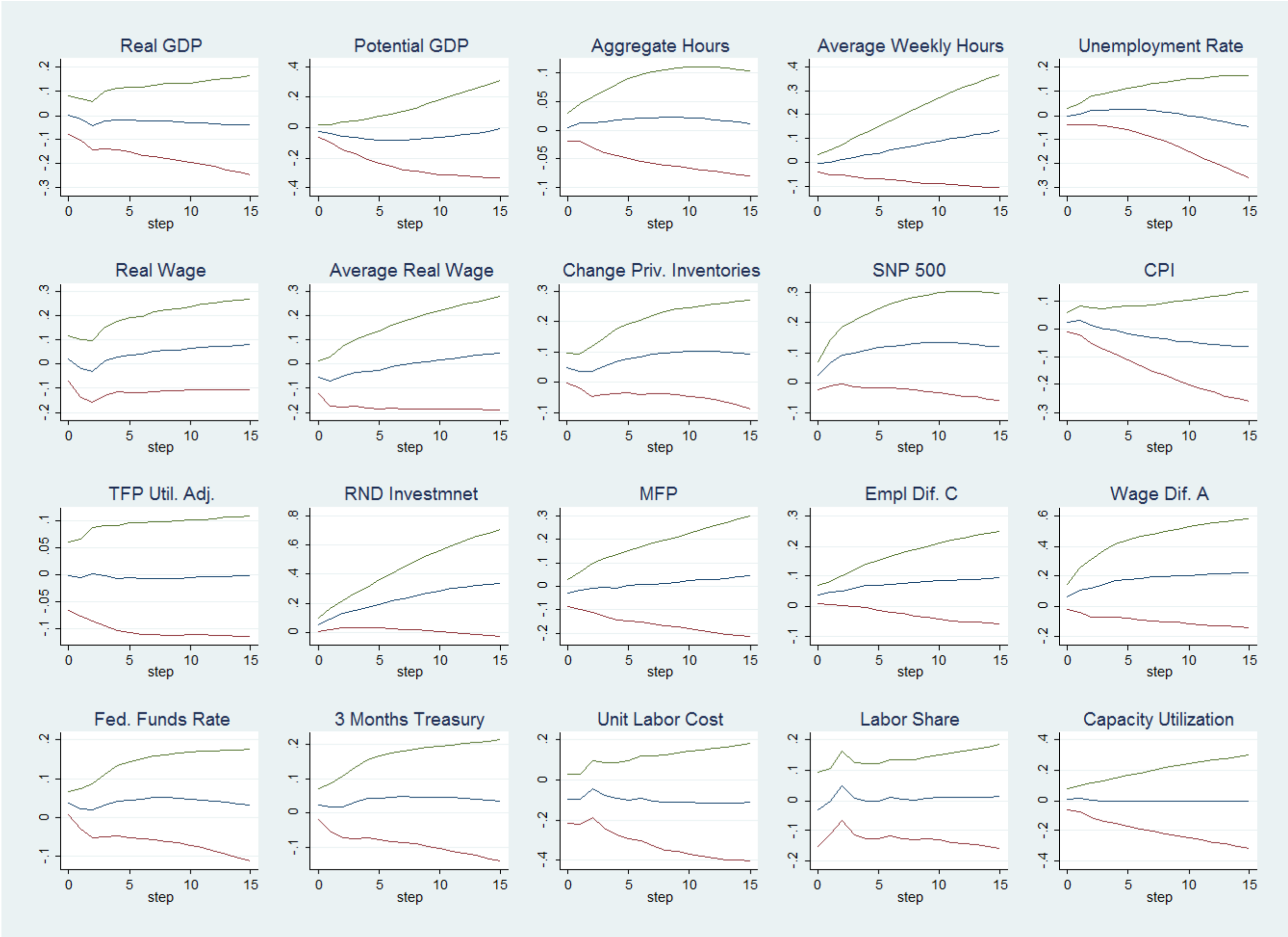


Figure 3.5: FAVAR Impulse-Responses to Technological Shock Using San-Francisco Tech Pulse and Long Run SVAR Restrictons, Bootstrapped 95% CI

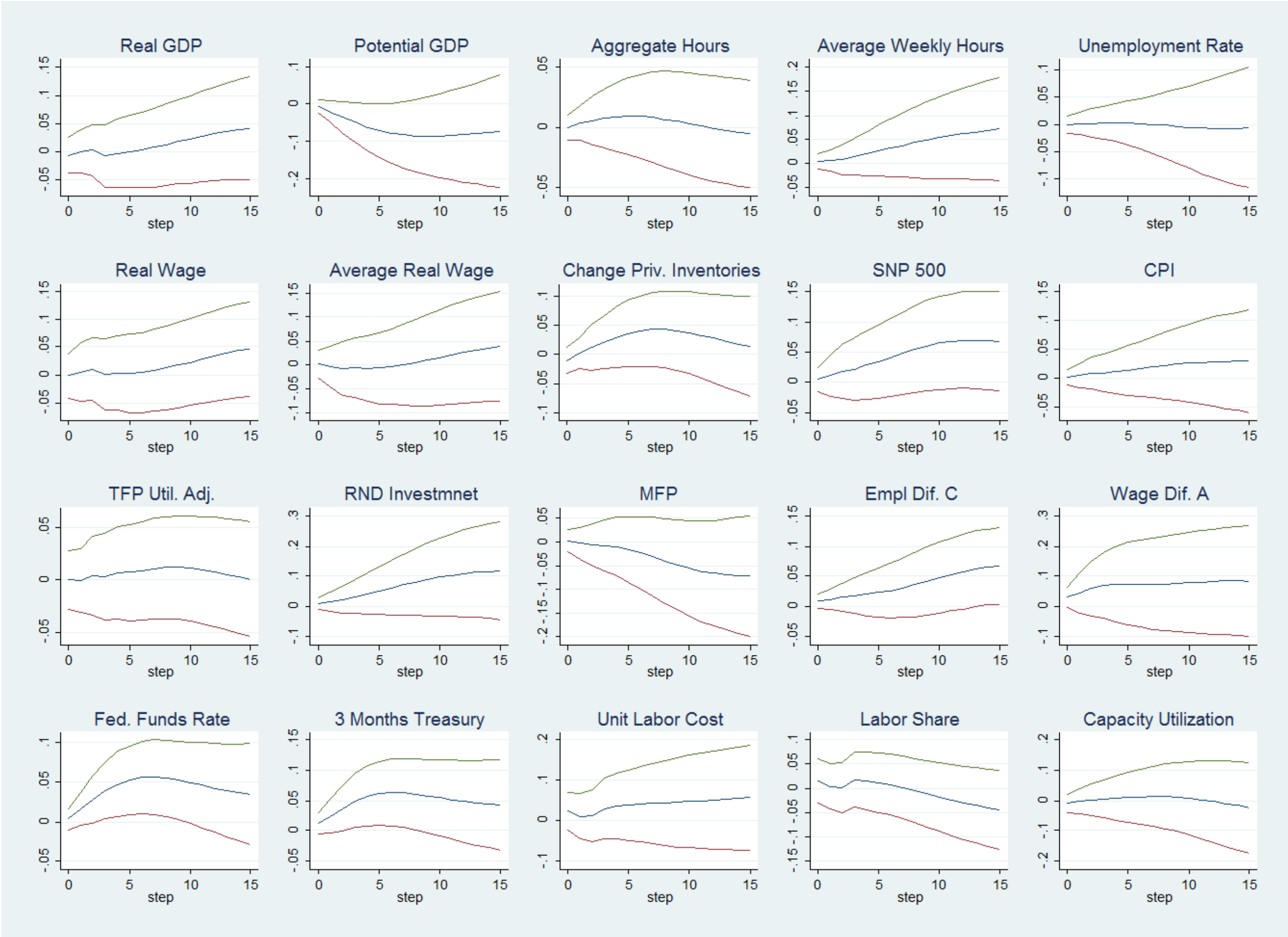


Figure 3.6: FAVAR Impulse-Responses to Technological Shock Using Multifactor Productivity Measure and Short Run SVAR Restrictons, Bootstrapped 95% CI

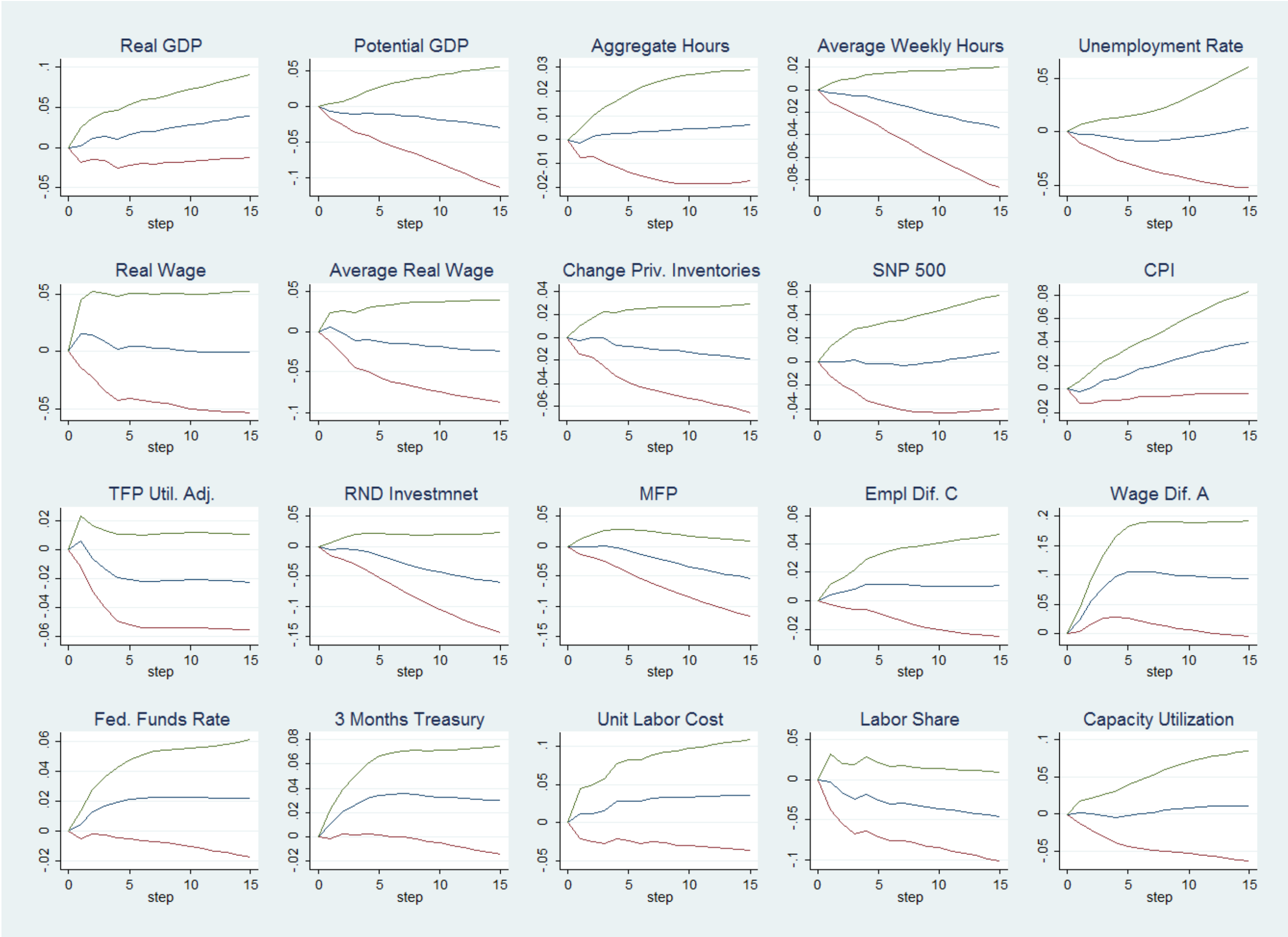


Figure 3.7: FAVAR Impulse-Responses to Technological Shock Using Utilization Adjusted Total Factor Productivity Measure and Short Run SVAR Restrictons, Bootstrapped 95% CI

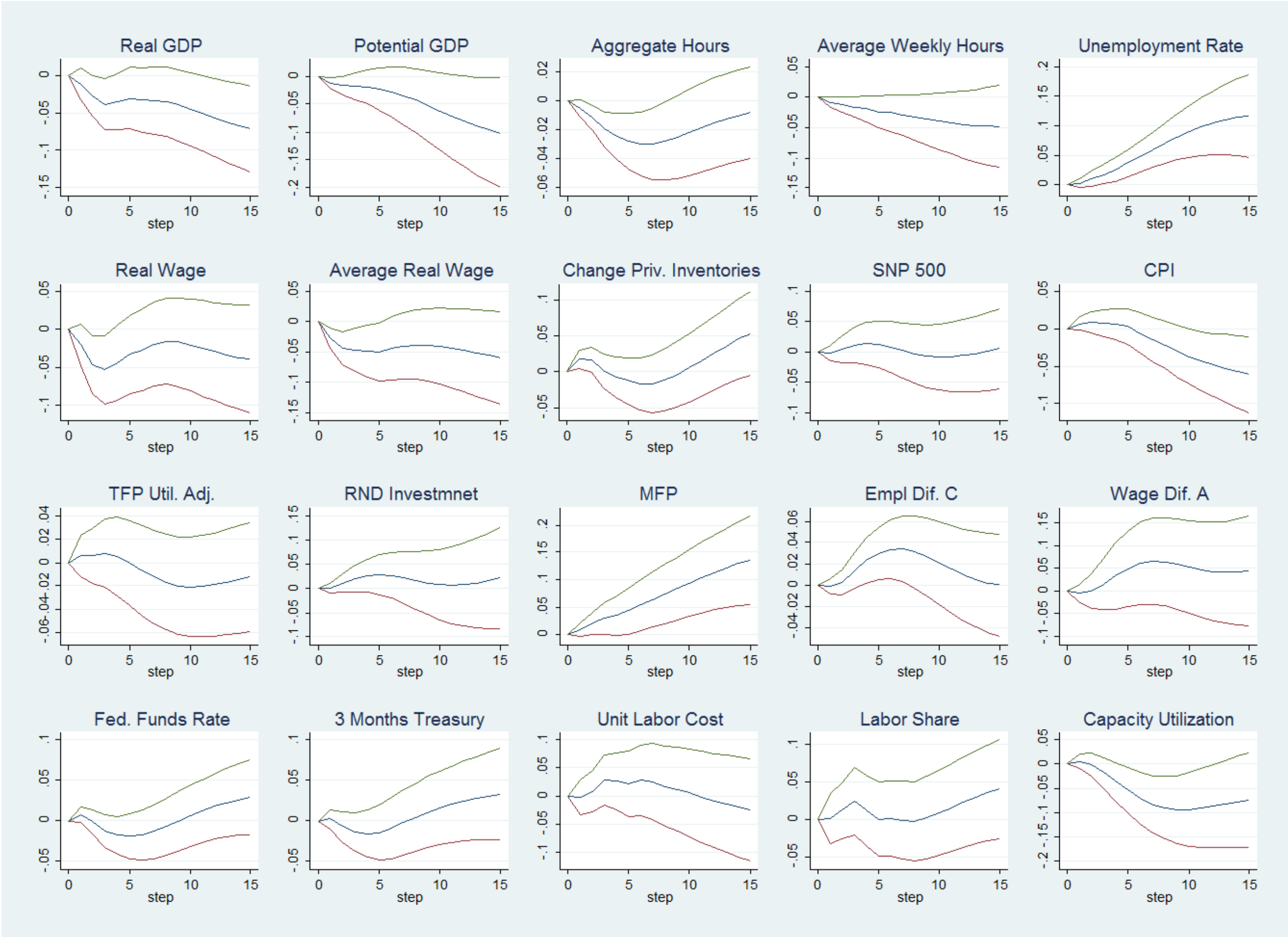


Figure 3.8: FAVAR Impulse-Responses to Technological Shock Using Claim adjusted Patent Measure and Short Run SVAR Restrictons, Bootstrapped 95% CI

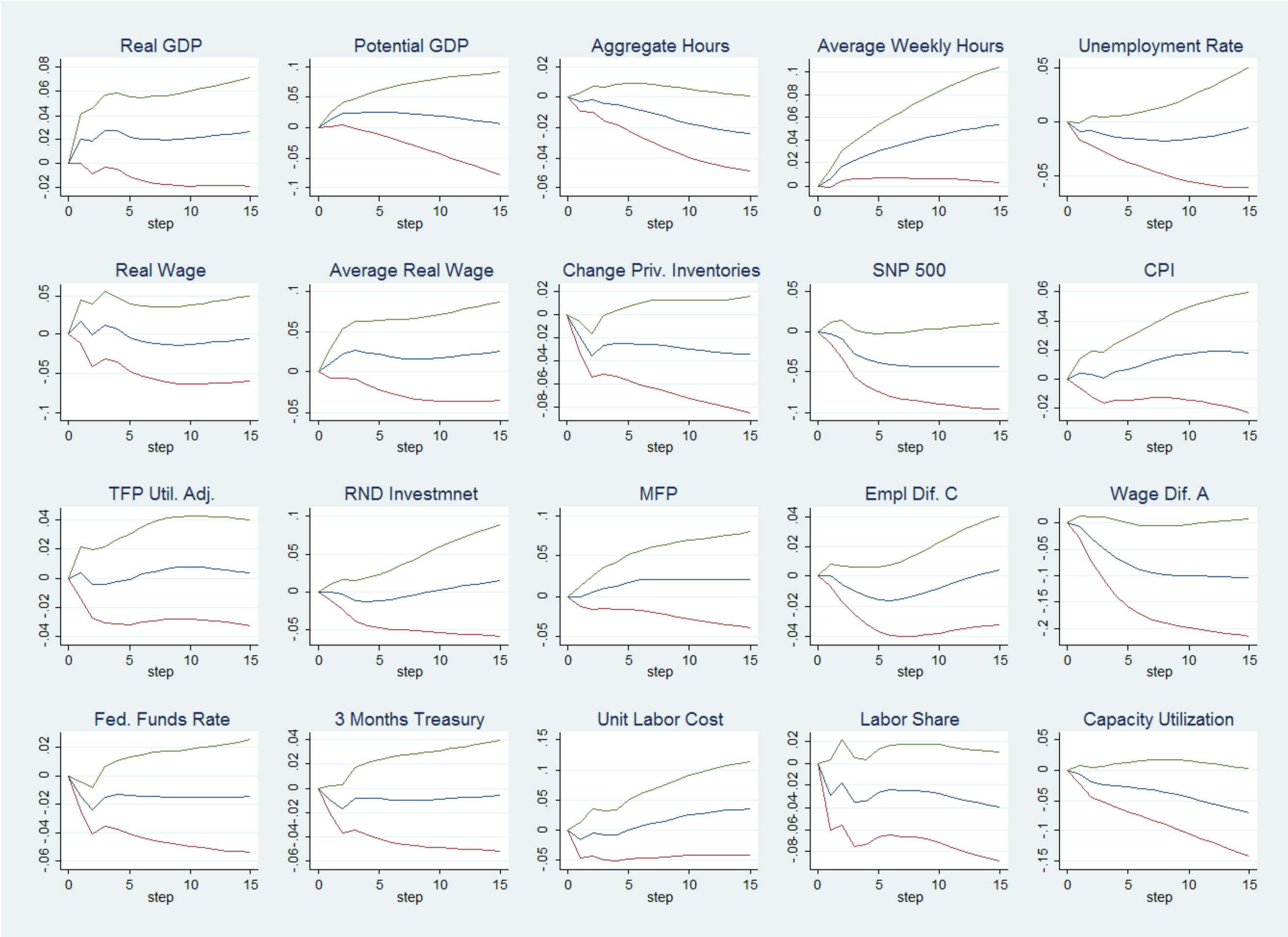


Figure 3.9: FAVAR Impulse-Responses to Technological Shock Using San Francisco Tech Pulse and Short Run SVAR Restrictons, Bootstrapped 95% CI

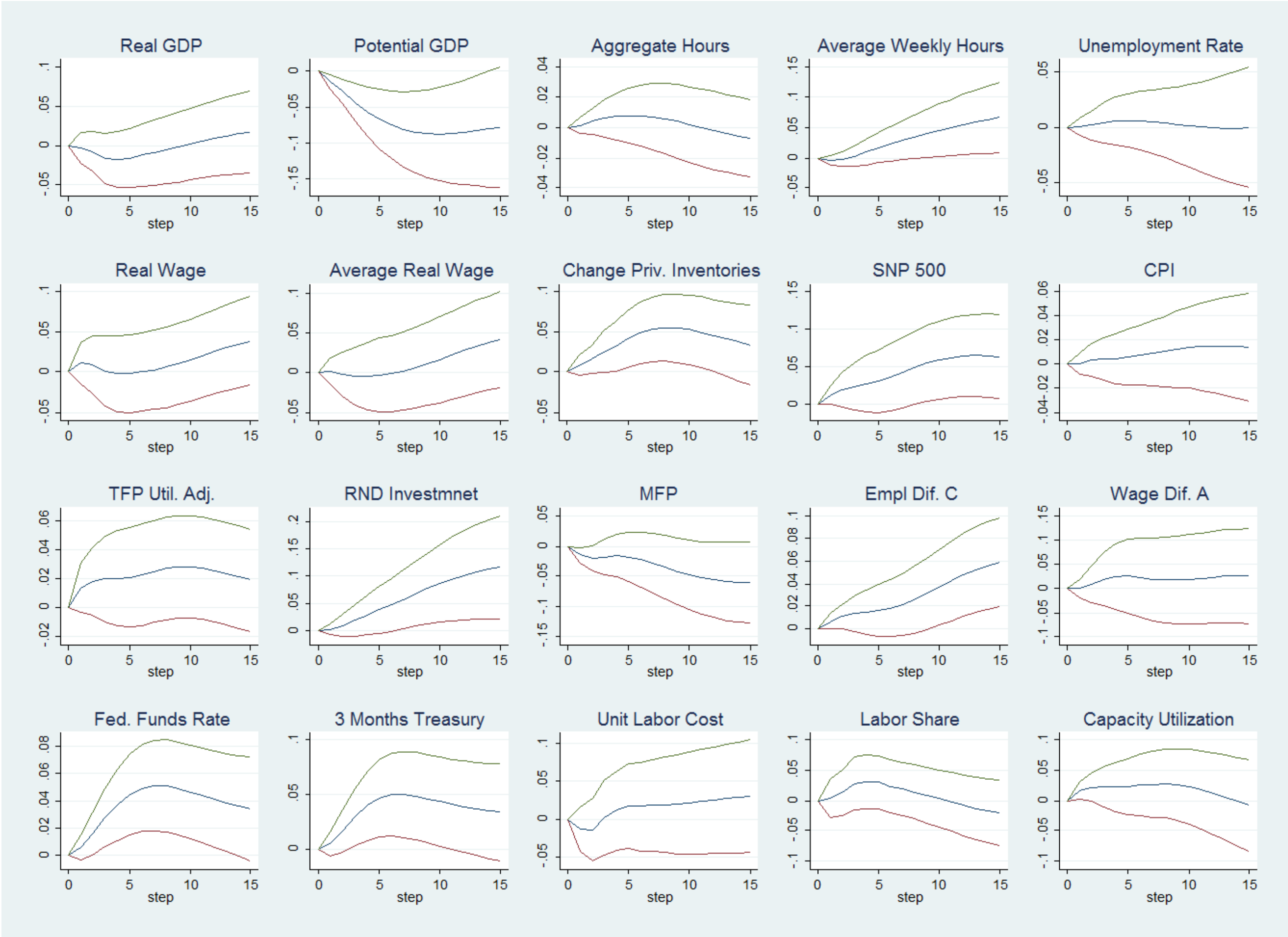


Figure 3.10: F-FAVAR - MFP Measure FAVAR with Long Runs Restrictions Comparison

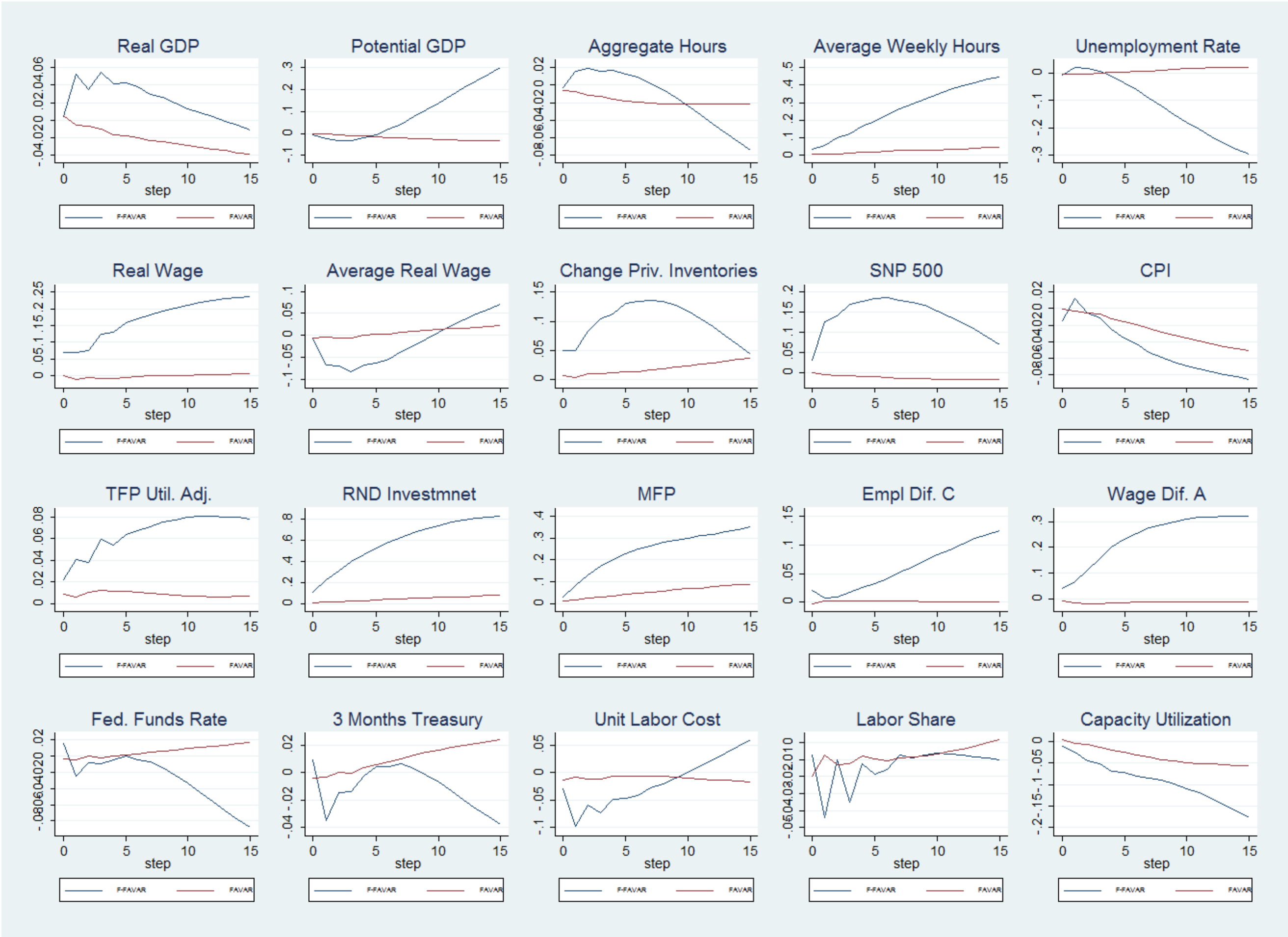


Figure 3.11: F-FAVAR - Utilization-Adjusted TFP Measure FAVAR with Long Runs Restrictions Comparison

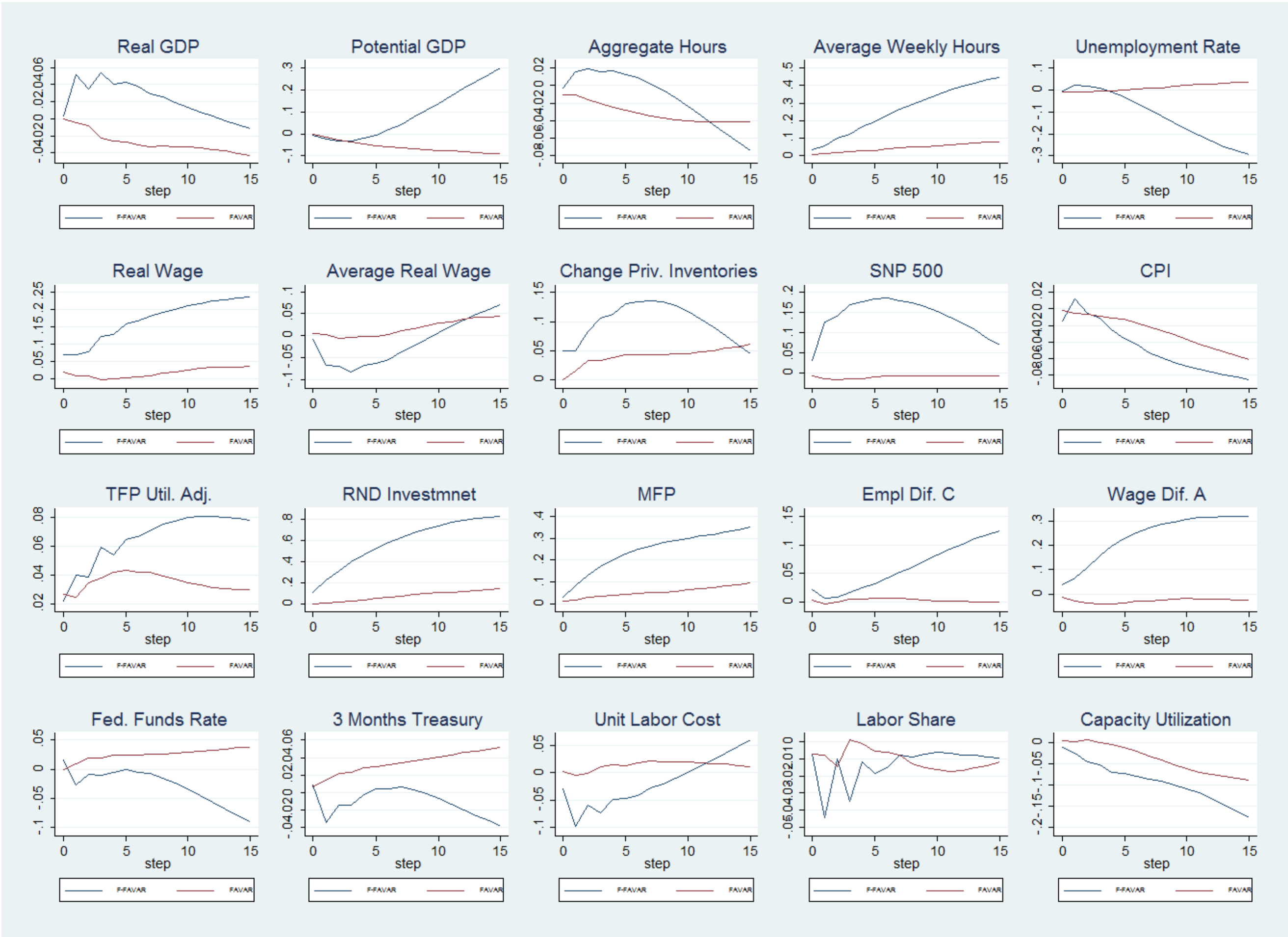


Figure 3.12: F-FAVAR - Claim-Adjusted Patent Measure FAVAR with Long Runs Restrictions Comparison

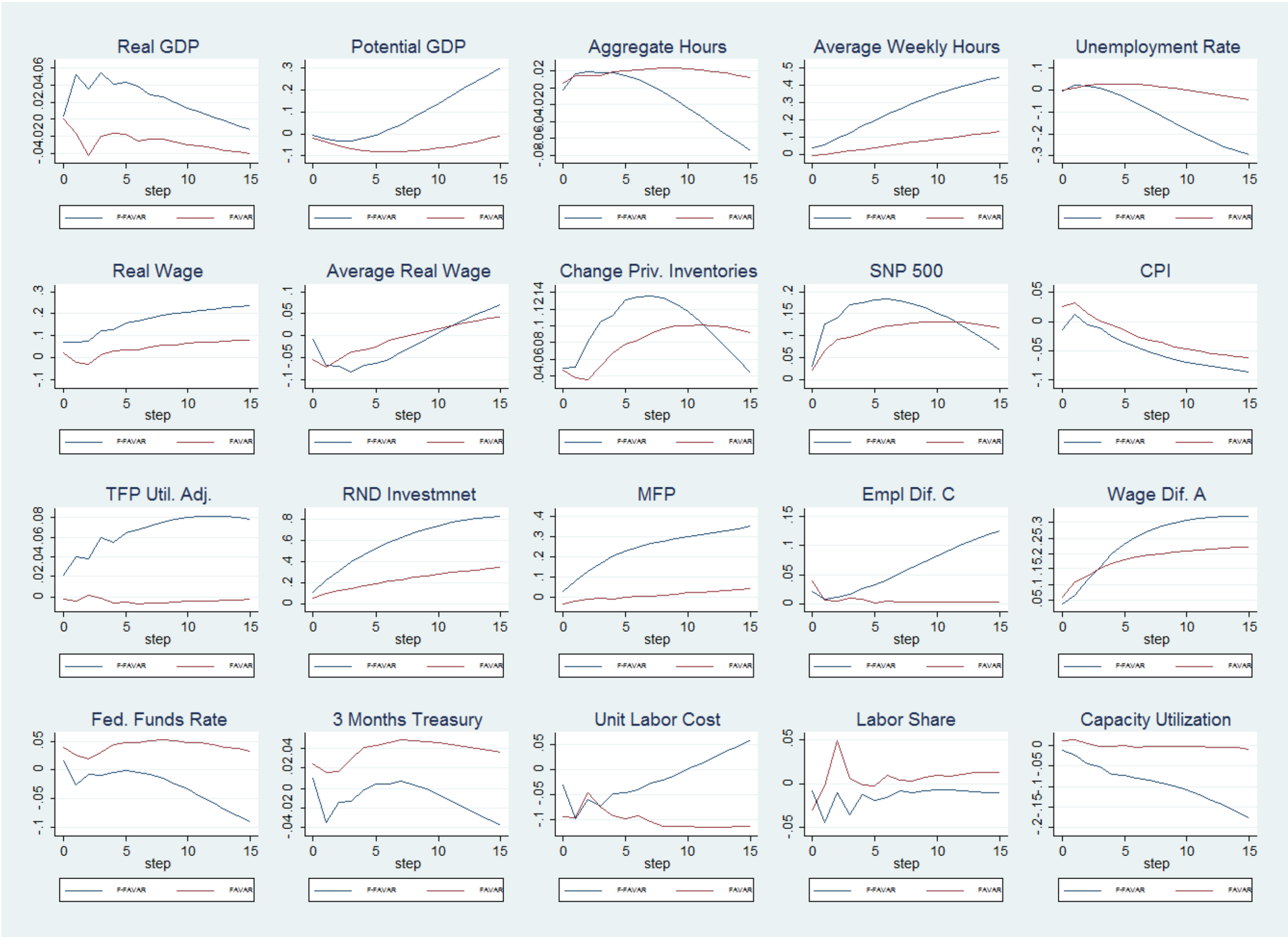


Figure 3.13: F-FAVAR - San Francisco Tech Pulse Measure FAVAR with Long Runs Restrictions Comparison

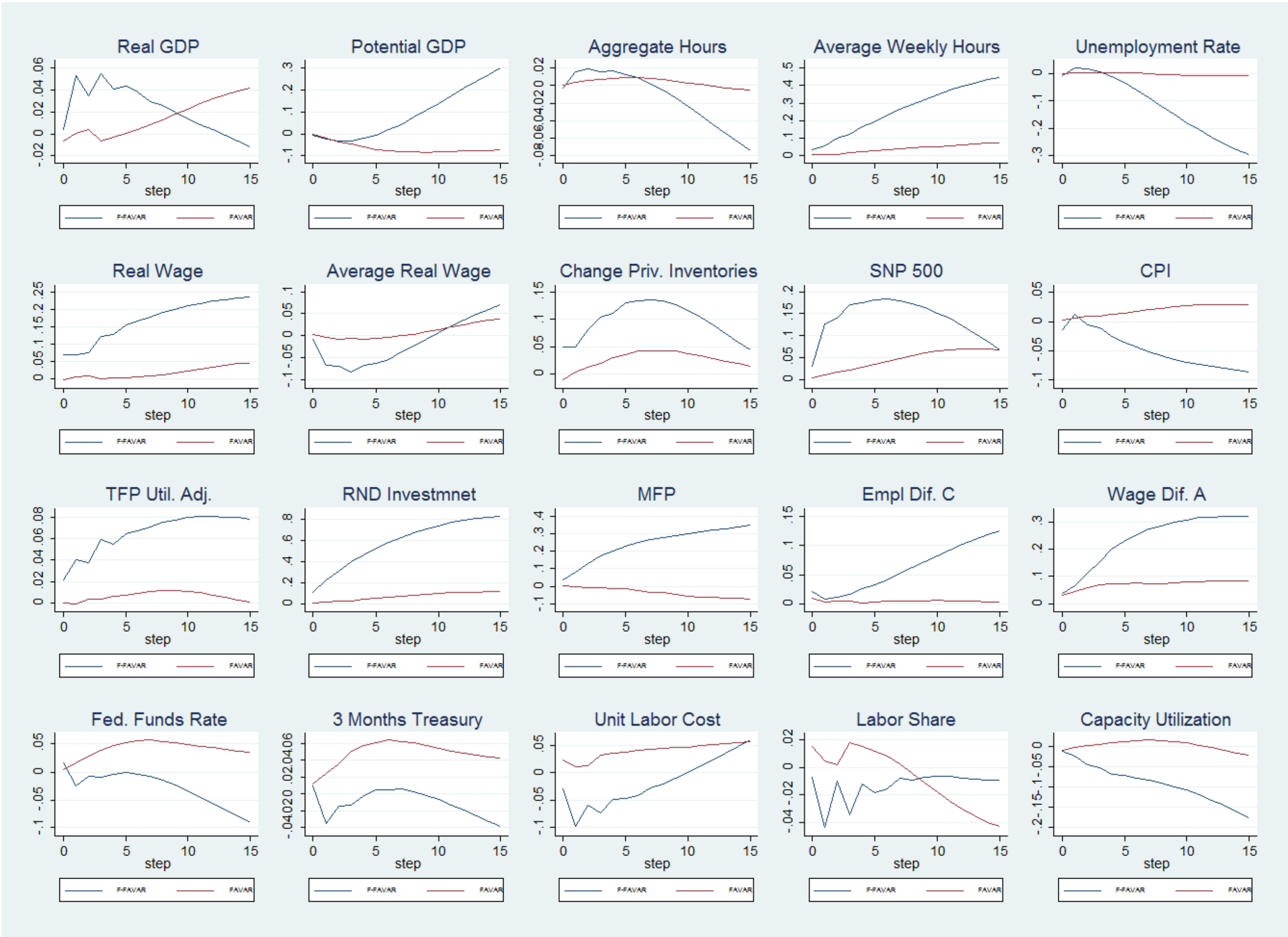


Figure 3.14: F-FAVAR - MFP Measure FAVAR with Short Run Restrictions Comparison

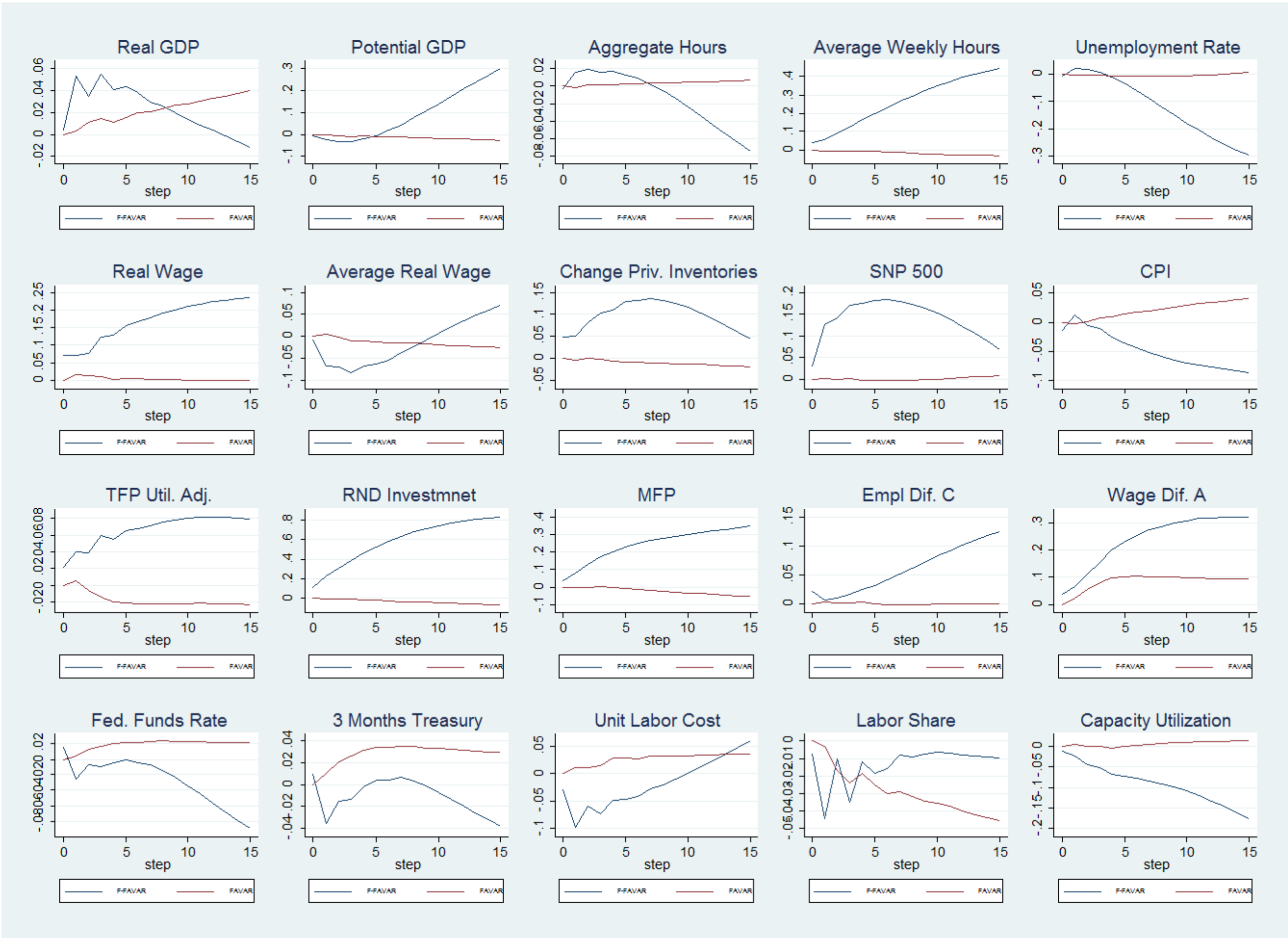


Figure 3.15: F-FAVAR - Utilization-Adjusted TFP Measure FAVAR with Short Run Restrictions Comparison

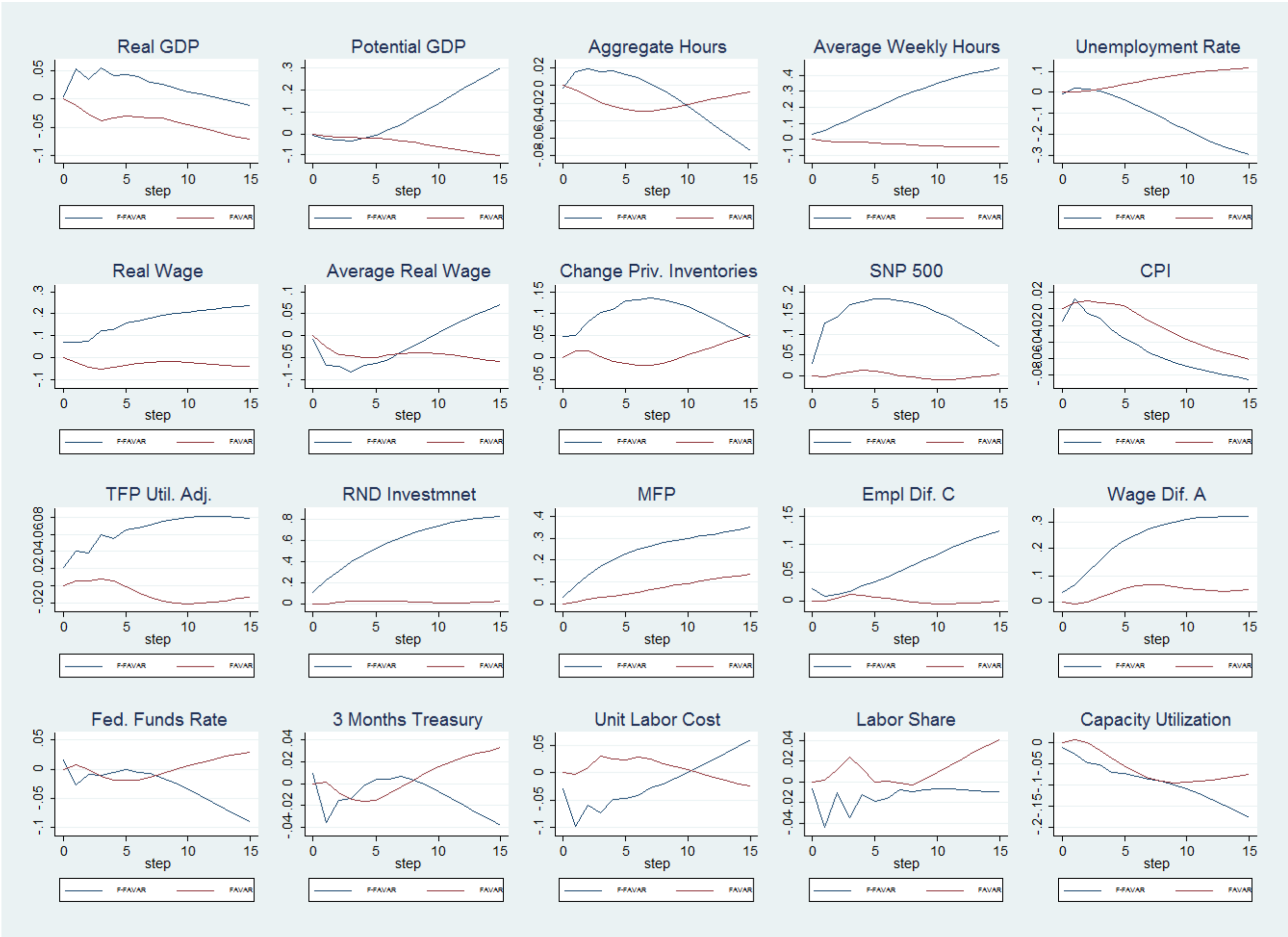


Figure 3.16: F-FAVAR - Claim-Adjusted Patent Measure FAVAR with Short Run Restrictions Comparison

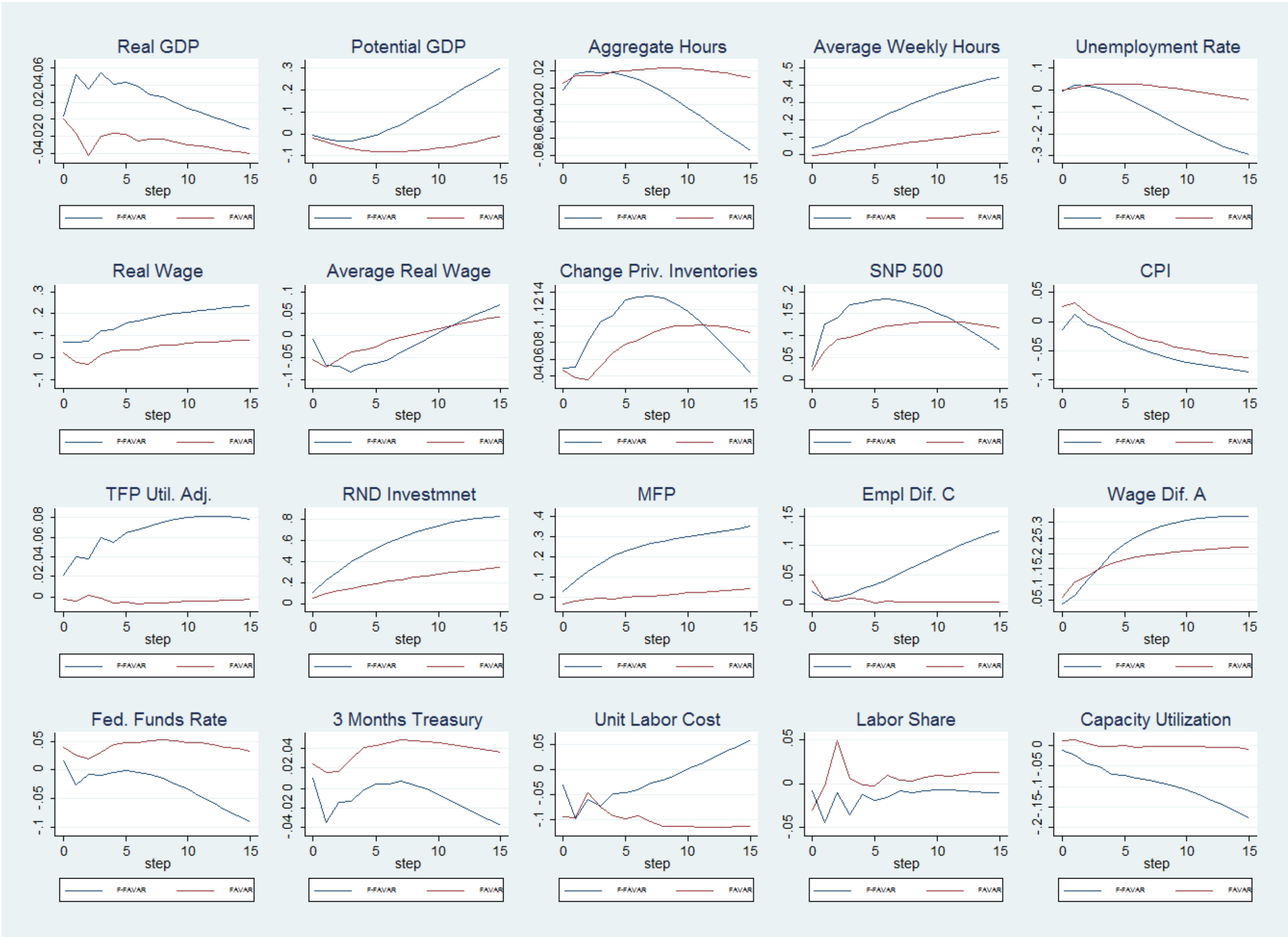
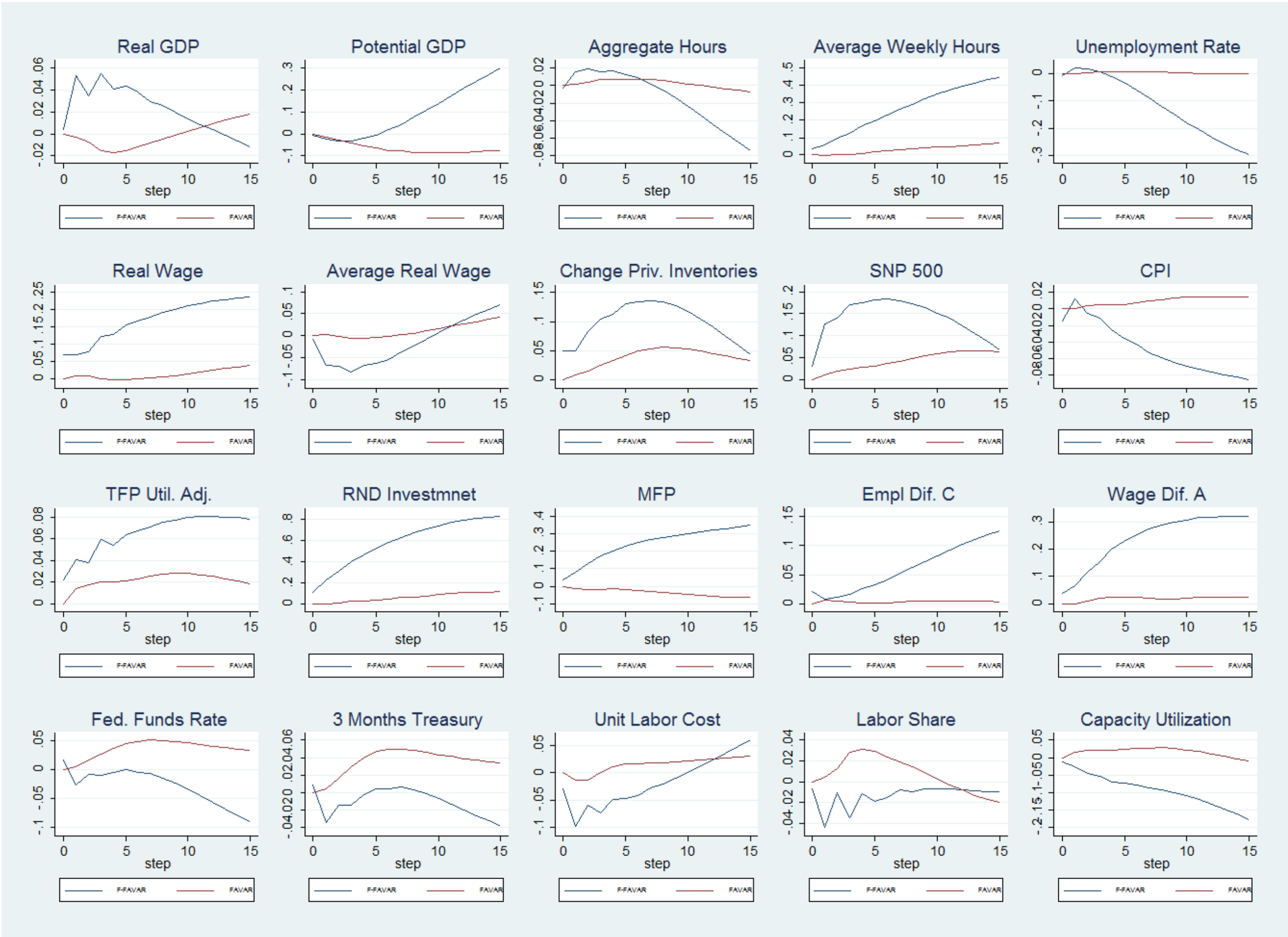


Figure 3.17: F-FAVAR - San Francisco Tech Pulse Measure FAVAR with Short Run Restrictions Comparison



3.7 Thesis Conclusion

In my PHD dissertation, I examined the effect of technological innovations on the economy. The first chapter was dedicated to uncovering and analyzing general purpose technologies in the patent data. I develop a way to identify and trace the full life cycle of general purpose technologies. I show that previous studies were able to identify only a particular segment of a GPT. Moreover, the methodology introduced in the previous studies over-identified the number of co-existing active GPTs in any historical period. On the other hand, developed in my first chapter identifies only a technology that is commonly accepted as a GPT and trace its evolution and diffusion across technological fields. The microcomputer general purpose technology identified in the study is found to behave in line with the GPT theory. This chapter also introduces a useful tool that allows us to examine the evolution and diffusion of a given technology over time and over rounds of research activity.

In the second and third chapters, I analyze how a technological innovation affects economic variables in the short run. The second chapter investigates the response of hours worked following a technology shock. A unique combination of a technology measure and an empirical identification methodology is used. The measure of technology comes from patent claims. Instead of using short-run or long-run restrictions in the VAR, we use sign restrictions. The identification for the SVAR comes from a theoretical model that distinguishes between skilled and unskilled workers. We find that a technology shock leads to an increase in aggregate hours worked. Unlike previous studies where the results depended on the specification of the variables, our results are robust to these changes.

In the third chapter, I extend the FAVAR methodology to analyse of the effects of technological shocks on a large set of economic variables. Unlike some shocks for which there exists an impulse variable that can be included into FAVAR, there is no such variable for the aggregate technological process. Multiple measures of technology exist and each have their own advantages and disadvantages. As a result, the selection of a given technological measure requires a justification. The F-FAVAR methodology introduced in the third chapter allows us to derive a latent factor technological variable from multiple technology measures. We examine the effects of a shock to this variable on numerous economic and business variables. In general the F-FAVAR methodology could be applied to any VAR analysis in order to include a larger set of data than the one that can be processed by regular VARs. The results of the F-FAVAR study generally support the results of the second chapter. A positive technology shock leads to an increase of hours worked. Moreover, the responses of other economic variables to a positive technology shock are largely in accordance with the theory which is not always achievable when a VAR framework is used. The F-FAVAR approach is found to out-perform more conventional FAVAR approach with different observable measures of technology and structural restrictions in terms of the intuitiveness of the results.

Bibliography

[Aghion & Howitt, 1998] Aghion, P. & Howitt, P. (1998). On the macroeconomic effects of major technological change. In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 5 (pp. 121–144). Cambridge, Ma: Massachusetts Institute of Technology.

[Aghion et al., 2002] Aghion, P., Howitt, P., & Violante, G. L. (2002). General purpose technology and wage inequality. *Journal of Economic Growth*, 7, 315–345.

[Alexopoulos & Cohen, 2010] Alexopoulos, M. & Cohen, J. (2010). *Volumes of Evidence - Examining Technical Change Last Century Through a New Lens*. Working Paper 392, University of Toronto, Department of Economics, Toronto, ON.

[Allen, 1996] Allen, S. G. (1996). *Technology and Wage Structure*. Working Paper 5534, NBER, Cambridge, MA.

[Archibugi & Pianta, 1996] Archibugi, D. & Pianta, M. (1996). Measuring technological change through patents and innovation surveys. *Technovation*, 16(9), 451–468.

[Bai & Ng, 2012] Bai, J. & Ng, S. (2012). Principal components estimation and identification of static factors. *Journal of Econometrics*, 176, 18–29.

[Basu & Fernald, 2007] Basu, S. & Fernald, J. (2007). Information and communications technology as a general-purpose technology: Evidence from US industry data. *German Economic Review*, 8(2), 146–173.

[Basu et al., 2006] Basu, S., Fernald, J. G., & Kimball, M. S. (2006). Are technology improvements contractionary? *American Economic Review*, 96(5), 1418–1448. The Published Version.

[Beaudry & Schiffauerova, 2011] Beaudry, C. & Schiffauerova, A. (2011). Impacts of collaboration and network indicators on patent quality: The case of canadian nanotechnology innovation. *European Management Journal*, 29, 364–376.

[Belenzon, 2006] Belenzon, S. (2006). *Knowledge Flow and Sequential Innovation: Implications for Technology Diffusion, R&D and Market Value*. Oxford University and CEP Discussion Paper 0721, Centre for Economic Performance.

[Bernanke et al., 2004] Bernanke, B. S., Boivin, J., & Eliasz, P. (2004). *Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach*. Working Paper 10220, NBER, Cambridge, MA.

[Blanchard & Quah, 1989] Blanchard, O. J. & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, 79(4), 655–673.

[Bocola et al., 2011] Bocola, L., Hagedorn, M., & Manovskii, I. (2011). Identifying technology shocks in models with heterogeneous inputs. Preliminary and Incomplete.

[Bresnahan, 2007] Bresnahan, T. F. (2007). Creative destruction in the PC industry. In F. Malebra & S. Brusconi (Eds.), *Perspectives on Innovation*, number 4 (pp. 105 – 141). Cambridge University Press.

[Bresnahan, 2010] Bresnahan, T. F. (2010). General purpose technologies. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the Economics of Innovations*, number 18 (pp. 761–790). Cambridge University Press.

[Bresnahan & Trajtenberg, 1995] Bresnahan, T. F. & Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65(1), 83–108.

[Brock & Laws, 2012] Brock, D. C. & Laws, D. A. (2012). The early history of microcircuitry: An overview. *IEEE Annals of the History of Computing*, 34(1), 7–19.

[Canova, 2007] Canova, F. (2007). Bayesian vars. In *Methods for Applied Macroeconomic Research* chapter 10, (pp. 351–394). Princeton University Press.

[Canova & De Nicolo, 2002] Canova, F. & De Nicolo, G. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, 49, 1131–1159.

[Cantner & Vannuccini, 2012] Cantner, U. & Vannuccini, S. (2012). *A New View of General Purpose Technologies*. Technical Report 2012-054, JENA Economic Research Papers.

[Carlaw & Lipsey, 2006] Carlaw, K. I. & Lipsey, R. G. (2006). GPT driven endogenous growth. *The Economic Journal*, 116, 155–174.

[Carlaw & Lipsey, 2011] Carlaw, K. I. & Lipsey, R. G. (2011). Sustained endogenous growth driven by structured and evolving general purpose technologies. *Journal of Evolutionary Economics*, 21(4), 563–593.

[Chari et al., 2007] Chari, V. V., Kehoe, P. J., & McGrattan, E. R. (2007). *Are Structural VARs with Long-Run Restrictions Useful in Developing Business Cycle Theory?* Research Department Staff Report 364, Federal Reserve Bank of Minneapolis.

[Christiano et al., 2003a] Christiano, L. J., Echenbaum, M., & Vigfusson, R. J. (2003a). *The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology*. International Finance Discussion Papers 790, Board of Governors of the Federal Reserve System.

[Christiano et al., 2003b] Christiano, L. J., Echenbaum, M., & Vigfusson, R. J. (2003b). *What Happens After a Technology Shock?* Working Paper 9819, NBER, Cambridge, MA.

[Christiansen, 2008] Christiansen, L. E. (2008). *Do Technology Shocks Lead to Productivity Slowdowns? Evidence from Patent Data*. Working Paper 08/24, IMF.

[Costello & Osborne, 2005] Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10(7).

[David, 1990] David, P. A. (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *The American Economic Review*, 80(2), 355–361. Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association (May, 1990).

[David & Wright, 1999] David, P. A. & Wright, G. (1999). *General Purpose Technologies and Surges in Productivity: Historical Reflections on the Future of the ICT Revolution*. Technical Report 31, University Of Oxford. Discussion Papers in Economic and Social History.

[Debola & Neri, 2007] Debola, L. & Neri, S. (2007). What does a technology sshock do? a VAR analysis with model-based sign restrictions. *Journal of Monetary Economics*, 54, 512–549.

[Duguet & MacGarvie, 2007] Duguet, E. & MacGarvie, M. (2007). How well do patent citations measure flows of technology? evidence from french innovation surveys. *Economics of Innovation and New Technology*, 14(5), 375–393.

[Fabrigar et al., 1999] Fabrigar, L. R., MacCallum, R. C., Wegener, D. T., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299.

[Faust, 1998] Faust, J. (1998). The robustness of identified VAR conclusions about money. In *Carnegie-Rochester Conference Series on Public Policy*, volume 49 (pp. 207–244). North-Holland: Elsevier.

[Fernald, 2012] Fernald, J. G. (2012). *A Quarterly, Utilization-Adjusted Series on Total Factor Productivity*. Working Paper 2012-19, Federal Reserve Bank of San Francisco.

[Fernandez-Villaverde, 2009] Fernandez-Villaverde, J. (2009). *The Econometrics of DSGE Models*. Working Paper 14677, NBER, Cambridge, MA.

[Francis & Ramey, 2002] Francis, N. & Ramey, V. A. (2002). *Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revised*. Working Paper 8726, NBER, Cambridge, MA.

[Freeman & Soete, 1997] Freeman, C. & Soete, L. (1997). *The Economics of Innovation*. The MIT Press, 3 edition.

[Fry & Pagan, 2007] Fry, R. & Pagan, A. (2007). *Some Issues in Using Sign Restrictions for Identifying Structural VARs*. NCER Working Paper Series 14, National Centre for Economic Research.

[Fry & Pagan, 2010] Fry, R. & Pagan, A. (2010). *Sign Restrictions in Structural Vector Autoregressions: A Critical Review*. CAMA Working Paper Series 22/2010, Centre for Applied Macroeconomic Analysis, Barry Dr, Acton ACT 0200, Australia.

[Gali, 1999] Gali, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *The American Economic Review*, 89(1), 249–271.

[Gali et al., 2011] Gali, J., Smeth, F., & Wouters, R. (2011). *Unemployment in an Estimated New Keynesian Model*. Working Paper 17084, NBER, Cambridge, MA.

[Goldin & Katz, 2007] Goldin, C. & Katz, L. F. (2007). Long-run changes in the wage structure: Narrowing, widening, polarizing. In *Brookings Papers on Economic Activity*, number 2 (pp. 135 – 167).

[Gress, 2010] Gress, B. (2010). Properties of the USPTO patent citation network: 1963 - 2002. *World Patent Information*, 32, 3–21. This article has been developed from a presentation by the author at the International Patent Information Conference, IPI-ConfEx, in March 2007, in Sorrento, Italy.

[Griliches, 1990] Griliches, Z. (1990). *Patent Statistics as Economic Indicators: A Survey Part I*. Working Paper 3301, NBER, Cambridge, MA.

[Grossman & Helpman, 1990] Grossman, G. M. & Helpman, E. (1990). *Trade, Knowledge Spillovers, and Growth*. Working Paper 3485, NBER, Cambridge, MA.

[Hall et al., 2001a] Hall, B. H., Jaffe, A., & Trajtenberg, M. (2001a). *Market Value and Patent Citations: A First Look*. Working Paper E01-304, UC Berkeley. This is a revision of a paper prepared for the Conference in Commemoration of Zvi GrilichesâĂŹ 20 Years as Director of the NBER Program on Productivity and Technological Progress, Cambridge, Massachusetts, March 5 and 6, 1999.

[Hall et al., 2001b] Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001b). *The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools*. Working Paper 8498, NBER, Cambridge, MA.

[Hall & Trajtenberg, 2004] Hall, B. H. & Trajtenberg, M. (2004). *Uncovering GPTs With Patent Data*. Working Paper 10901, NBER, Cambridge, MA.

[Hall & Ziedonis, 2001] Hall, B. H. & Ziedonis, H. (2001). The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *RAND Journal of Economics*, 32(1), 101–128.

[Hall, 1988] Hall, S. S. (1988). The age of electricity. In E. L. Newhouse (Ed.), *Inventors and Discoverers Changing Our World* (pp. 44–79). Washington, D. C.: National Geographic Society.

[Hansen, 1985] Hansen, G. D. (1985). Indivisible labor and the business cycle. *Journal of Monetary Economics*, 16, 309 – 327.

[Harris, 1998] Harris, R. G. (1998). Internet as a gpt. In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 6 (pp. 146–166). Cambridge, Ma: Massachusetts Institute of Technology.

[Helpman & Trajtenberg, 1998a] Helpman, E. & Trajtenberg, M. (1998a). Diffusion of general purpose technologies. In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 4 (pp. 85–119). Cambridge, Ma: Massachusetts Institute of Technology.

[Helpman & Trajtenberg, 1998b] Helpman, E. & Trajtenberg, M. (1998b). A time to sow and a time to reap: Growth based on general purpose technologies. In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 3 (pp. 55 – 84). Cambridge, Ma: Massachusetts Institute of Technology.

[Hobijn & Stiroh & Antoniades, 2003] Hobijn, B. & Stiroh, K. J. & Antoniades, A. (2003). *Taking the Pulse of the Tech Sector: A Coincident Index of High-Tech Activity*. Current Issues in Economics and Finance, 9(10), Federal Reserve Bank of New York.

[Jovanovich & Rousseau, 2005] Jovanovich, B. & Rousseau, P. L. (2005). General purpose technologies. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth*, volume 1B chapter 18, (pp. 44–79). Elsevier, 1 edition.

[Katz & Margo, 2013] Katz, L. F. & Margo, R. A. (2013). *Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective*. Working Paper 18752, NBER, Cambridge, MA.

[Kilian, 2012] Kilian, L. (2012). Structural vector autoregressions. University of Michigan.

[Kim & Marschke, 2004] Kim, J. & Marschke, G. (2004). Accounting for the recent surge in U.S. patenting: Changes in R&D expenditures, patent yields, and the high tech sector. *Economics of Innovation and New Technology*, 13(6), 543–558.

[King et al., 1988] King, R. J., Plosser, C. I., & Rebelo, S. T. (1988). Production growth and business cycles: The basic neoclassical model. *Journal of Monetary Economics*, 21, 195–232.

[King & Rebelo, 2000] King, R. J. & Rebelo, S. T. (2000). *Resuscitating Real Business Cycles*. Working Paper 7534, NBER, Cambridge, MA.

[Kortum & Lerner, 1999] Kortum, S. & Lerner, J. (1999). What is behind the recent surge in patenting? *Research Policy*, 28(1), 1–22.

[Lach, 1995] Lach, S. (1995). Patents and productivity growth at the industry level: A first look. *Economics Letters*, 49, 101–108.

[Lanjouw et al., 1998] Lanjouw, J. O., Pakes, A., & Putnam, J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. *The Journal of Industrial Economics*, 46(4), 405–432.

[Lanjouw & Schankerman, 1999] Lanjouw, J. O. & Schankerman, M. (1999). *The Quality of Ideas: Measuring Innovation with Multiple Indicators*. Working Paper 7345, NBER, Cambridge, MA.

[Lanjouw & Schankerman, 2004] Lanjouw, J. O. & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, 114, 441–465.

[Lin & Chen, 2005] Lin, B. & Chen, J. S. (2005). Corporate technology portfolios and R& D performance measures: A study of technology intensive firms. *R& D Management*, 35(2), 157–170.

[Lipsey & Carlaw, 2004] Lipsey, R. G. & Carlaw, K. I. (2004). Total factor productivity and the measurement of technological change. *The Canadian Journal of Economics*, 37(4), 1118–1150.

[Lipsey et al., 1998] Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (1998). What requires explanation? In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 1 (pp. 15 –54). Cambridge, Ma: Massachusetts Institute of Technology.

[Lipsey et al., 2005] Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). *Economic Transformations: General Purpose Technologies and Long Term Economic Growth*. New York: Oxford University Press.

[Marco et al., 2015] Marco, A. C., Carley, M., Jackson, S., & Myers, A. F. (2015). *The USPTO Historical Patent Data Files: Two Centuries of Innovations*. USPTO Economic Working Paper 2015-1, USPTO, Alexandria, VA.

[Marquis et al., 2011] Marquis, M. H., Trehan, B., & Tantivong, W. (2011). *The Wage Premium Puzzle and the Quality of Human Capital*. Working Paper WP 2011-06, Federal Reserve Bank of San Francisco.

[McCandless, 2008] McCandless, G. (2008). *The ABCs of RBCs: An Introduction to Dynamic Macroeconomic Models*. Cambridge, Massachusetts: Harward University Press.

[Menz & Ott, 2011] Menz, N. & Ott, I. (2011). *On the Role of General Purpose Technologies within the Marshall-Jacobs Controversy: The Case of Nanotechnologies*. Working Paper 18, KIT - University of the State of Baden-Wuerttemberg and National Laboratory of the Helmholtz Assoc.

[Michelacci & Lopez-Salido, 2007] Michelacci, C. & Lopez-Salido, D. (2007). Technology shocks and job flows. *Review of Economic Studies*, 74, 1195–1227.

[Mokyr, 1990] Mokyr, J. (1990). *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press.

[Moser & Nicholas, 2004] Moser, P. & Nicholas, T. (2004). Was electricity a general purpose technology? evidence from historical patent citations. *The American Economic Review*, 94(2), 388–394.

[Nikulainen & Kulvik, 2009] Nikulainen, T. & Kulvik, M. (2009). *How General Are the General Purpose Technologies*. Working Paper 1208, ETLA - The Research Institute of Finnish Economy.

[Aghion, P. & Howitt, P., 1992] Aghion, P. & Howitt, P. (1992). A Model of Growth Through Creative Destruction *Econometrica*, 60(2), 323–351.

[Oliner & Sichel, 2000] Oliner, S. D. & Sichel, D. E. (2000). The resurgence of growth in the late 1990s: Is information technology the story? *Journal of Economic Perspectives*, 14(4), 3–22.

[Oystrakh, 2013] Oystrakh, M. (2013). *The Effect of Technological Innovations on Hours Worked*. PhD thesis, University of Waterloo, Waterloo, ON. Chapter 2: Effect of Technological Innovations on Hours Worked.

[Oystrakh, 2014] Oystrakh, M. (2014). General purpose technology: Nature, classification, examples. Unpublished Manuscript. University of Waterloo, Department of Economics.

[Park, 2012] Park, K. (2012). Employment responses to aggregate and sectoral technology shocks. *Journal of Macroeconomics*, 34, 801–821.

[Peersman & Straub, 2004] Peersman, G. & Straub, R. (2004). *Technology Shocks and Robust Sign Restrictions in a Euro Area SVAR*. Working Paper 373, European Central Bank.

[Prasad, 1996] Prasad, E. (1996). Skill heterogeneity and the business cycle. *The Canadian Journal of Economics*, 29(4), 910–929.

[Rafiquzzaman & Whewell, 1998] Rafiquzzaman, M. & Whewell, L. (1998). *Recent Jumps in Patenting Activities: Comparative Innovative Performance of Major Industrial Countries, Patterns and Explanations.* Working Paper 27, Industry Canada Research Publication Program, Ottawa, ON.

[Rebelo, 2005] Rebelo, S. T. (2005). Real business cycle models: Past, present and future. *The Scandinavian Journal of Economics*, 107(2), 217–238.

[Reid, 1988] Reid, T. R. (1988). Computers & chips. In E. L. Newhouse (Ed.), *Inventors and Discoverers Changing Our World* (pp. 246–273). Washington, D. C.: National Geographic Society.

[Rogers, 1983] Rogers, E. M. (1983). *Diffusion of Innovations*. New York: The Free Press, 3 edition.

[Rogerson, 1988] Rogerson, R. (1988). Indivisible labor, lotteries and equilibrium. *Journal of Monetary Economics*, 21, 3 – 16.

[Romer, 1990] Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economys*, 98(5.2), S71–S102.

[Rosenberg, 1998] Rosenberg, N. (1998). Chemical engineering as a general purpose technology. In E. Helpman (Ed.), *General Purpose Technologies and Economic Growth*, number 7 (pp. 167 – 192). Cambridge, Ma: Massachusetts Institute of Technology.

[Rothwell et al., 2013] Rothwell, J., Lobo, J., Strumsky, D., & Muro, M. (2013). *Patenting Prosperity: Invention and Economic Performance in the United States and Its Metropolitan Areas*. Metropolitan policy program research paper, Brookings Institution, Washington D.C.

[Rubio-Ramirez et al., 2005] Rubio-Ramirez, J. F., Waggoner, D., & Zha, T. (2013). *Markov-Switching Structural Vector Autoregressions: Theory and Application*. Federal Reserve Bank of Atlanta Working Paper Series, 2005-27

[Schaller, 1997] Schaller, R. R. (1997). Moore's law: Past, present and future. *Spectrum, IEEE*, 34, 52–59. The sixth issue of this article was used, the date of the issue: August 2002.

[Sedeen, 1988] Sedeen, M. (1988). The power of steam. In E. L. Newhouse (Ed.), *Inventors and Discoverers Changing Our World* (pp. 14–43). Washington, D. C.: National Geographic Society.

[Shea, 1998] Shea, J. (1998). *What Do Technology Shocks Do?* Working Paper 6632, NBER, Cambridge, MA.

[Solow, 1957] Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3), 312–320.

[Stock & Watson, 1989] Stock, J. H. & Watson, M. W. (1989). *New Indexes of Coincident and Leading Economic Indicators*. NBER Macroeconomic Annual 1989, 4, MIT PRESS.

[Stock & Watson, 2005] Stock, J. H. & Watson, M. W. (2005). *Implications of Dynamic Factor Models for VAR Analysis*. Working Paper 11467, NBER, Cambridge, MA.

[Strumsky et al., 2012] Strumsky, D., Lobo, J., & Van der Leeuw, S. (2012). Using patent technology codes to study technological change. *Economics of Innovation and New Technology*, 21(3), 267–286.

[Summer, 1985] Summer, M. (1985). *Computers: Concepts and Uses*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

[The BBC English Dictionary, 1992] The BBC English Dictionary.

[The Oxford English Dictionary, 1989] The Oxford English Dictionary.

[Tong & Frame, 1994] Tong, X. & Frame, J. D. (1994). Measuring national technological performance with patent claims data. *Research Policy*, 23(2), 133–141.

[Tryfos, 1998] Tryfos, P. (1998). Factor analysis. In *Methods for Business Analysis and Forecasting: Text and Cases* chapter 14. John Wiley and Sons, Inc.

[Turner, 2004] Turner, J. L. (2004). In defense of the patent friendly court hypothesis: Theory and evidence. Unpublished manuscript, University of Georgia, Department of Economics.

[Uhlig, 2005] Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52, 381–419.

[United State Patent Act, 1836,] United State Patent Act, 1836. Technical report.

[van Zeebroeck & van Pottelsberghe de la Potterie, 2011a] van Zeebroeck, N. & van Pottelsberghe de la Potterie, B. (2011a). Filing strategies and patent value. *Economics of Innovation and New Technology*, 20(6), 539–561.

[van Zeebroeck & van Pottelsberghe de la Potterie, 2011b] van Zeebroeck, N. & van Pottelsberghe de la Potterie, B. (2011b). The vulnerability of patent value determinants. *Economics of Innovation and New Technology*, 20(3), 283–308.

[van Zon et al., 2003] van Zon, A., Fortune, E., & Kronenberg, T. (2003). *How to Sow and Reap as You Go: a Simple Model of Cyclical Endogenous Growth*. Research Memorandum Series 026, Merit-Infonomics, 6200 MD Maastricht The Netherlands.

[Webster's Dictionary, 1997] Webster's Dictionary.

[Wen, 2006] Wen, Y. (2006). *Demand Shocks and Economic Fluctuations*. Working Paper WP 2006-011A, Federal Reserve Bank of St. Louis.

[Weste & Harris, 2009] Weste, N. H. E. & Harris, D. M. (2009). *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley, 4 edition.

[Yong & Pearse, 2013] Yong, G. & Pearse, S. (2013). A beginnerâĂŹs guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79–94.

[Youtie et al., 2008] Youtie, J., Iacopetta, M., & Graham, S. (2008). Assessing the nature of nanotechnology: Can we uncover an emerging general purpose technology? *The Journal of Technology Transfer*, 33(3), 315–329.