The Cumulative Cost of Regulations

Bentley Coffey* Patrick McLaughlin† Pietro Peretto‡

November 25, 2015

Abstract

We estimate the effects of federal regulation on value-added to GDP for a panel of 22 industries in the U.S. over a period of 35 years (1977 – 2012). The structure of our linear specification is explicitly derived from the closed-form solutions of a [multi-sector] Schumpeterian model of endogenous growth. We allow regulation to enter that specification in a flexible manner. Our estimates of the model's parameters are then identified from covariation in some standard sector-specific data joined with RegData 2.2, which measures the incidence of regulations on industries based on a text analysis of federal regulatory code. With the model's parameters fitted to real data, we confidently conduct counter-factual experiments on alternative regulatory environments. Our results show that economic growth has been dampened by approximately 0.8 percent per annum since 1980. Had regulations been held constant at levels observed in 1980, our model predicts that the economy would be nearly 25 percent larger (i.e., regulatory growth since 1980 cost GDP $4T in 2012, or about $13k per capita).

Keywords: regulation; economic growth; macroeconomic performance; endogenous growth; Schumpeterian growth; RegData; costs of regulation; total cost of regulation; federal regulation; costs of entry; productivity; total factor productivity; labor productivity

JEL Codes: E20; L50; 040

---

* Visiting Assistant Professor, Sanford School of Public Policy, Duke University

† Corresponding author; Senior Research Fellow, Mercatus Center at George Mason University; Email: pmclaughlin@mercatus.gmu.edu

‡ Professor, Department of Economics, Duke University

The authors thank Oliver Sherouse, Scott Eastman, and Laura Stanley for their research assistance. Some very useful comments were provided by participants at several Duke University workshops, as well as seminars at the University of South Carolina and at the Mercatus Center at George Mason University.

1 Introduction

In theory, regulations have long represented an important policy tool for addressing market failure or advancing other goals of policymakers. However, economists have suspected for decades that the legislative and regulatory processes may undermine regulations' effectiveness as positive policy tools—for example, by offering opportunities for regulations to serve the purposes of special interests rather than the public interest (Stigler 1971). Even an otherwise virtuously conceived regulation has long raised concerns for skeptics because such regulations may still result in adverse consequences that are hard to anticipate in a complex and dynamic economy (Peltzman 1975). Our understanding of these forces, potentially opposing in their effects on the economy, have laid the groundwork for a large and growing literature on the causes and consequences of regulation.

Much of this literature focuses on economic growth. Relatively recently, multinational indexes, such as the World Bank's Doing Business project or the OECD Indicators of Product Market Regulation Database have permitted a first generation of estimations of the effect of regulation on economic growth, generally finding that macroeconomic growth can be considerably slowed by lower quality regulatory regimes.[1] Djankov et al (2006), for example, uses the World Bank's Doing Business Index to examine a large panel of countries' regulations, finding that a country's improvement from the worst (first) to the best (fourth) quartile of business regulations leads to a 2.3 percentage point increase in annual GDP growth.[2]

While informative, these studies' reliance on broad regulatory indexes, such as the World Bank's Doing Business Index and the similar, OECD-produced Product Market Regulation database, necessitates tradeoffs. First, these indexes generally cover relatively short time spans, while changes in regulatory policy often require several years, if not decades, to implement. Second, they are not comprehensive. Instead, they focus on a

[1] For examples, see Djankov et al (2002) and Botero et al (2004).

[2] Several other studies use similar World Bank and/or OECD-produced panel data on regulation across countries; see Aghion et al (2010), Loayza et al (2005), Nicoletti and Scarpetta (2003), and Gorgens et al (2003).

few areas of regulation, and then typically only on whether regulations exist in these areas—not how complex or burdensome they are. Finally, they often rely upon the opinions of either the creators of the index (in the cases where the index only incorporates a few areas of regulation, the creators select the areas) or of "country experts" who rate how regulated a country is in questionnaires and surveys.

Subsequently, metrics based on actual regulatory text have permitted Dawson and Seater (2013) to use time-series measures of US federal regulation in a single-country, macroeconomic growth setting. Using a before-after identification strategy on covariation between aggregate data and a page count of federal regulatory code, the authors conclude that regulatory accumulation is responsible for slowing economic growth in the United States by an average of 2 percent per year between 1949 and 2005. We add to this literature in several ways.

First, we build a formal model of economic growth from microeconomic foundations. We explicitly model how firms' decisions to invest in improving their productivity ultimately drives economic growth; by allowing our measure of regulations to flexibly enter into our econometric model, we rely on the data to tell us how much regulation distorts these investment decisions of firms and thus hamper long-run economic growth. Our model is built on microeconomic foundations with enough complexity in detail and flexibility in specification to simultaneously capture the aggregate effect of regulatory accumulation as a drag on the long-run macroeconomy. Simultaneously, our model accommodates a rich mix of short-run outcomes in output, [net] entry, and investment in particular industries which can be spurred by some types of regulations.

We have constructed our theoretical model from a well-established model of endogenous growth (Peretto and Connolly 2007).[3] This style of endogenous growth model has consistently received strong support in empirical investigations of competing models of economic growth (Madsen and Ang

[3] Although it may be less apparent on the surface, the structure of the endogenous growth model in Peretto and Connolly (2007) is deeply connected to the structure in Peretto (2007), where the related model was built to study a more intimately related topic: the dynamic deadweight loss due to corporate taxation. However, the structure in Peretto (2007) is less amendable to generalizing those closed-form solutions to a multi-sector structure.

2011).[4] To conduct our analysis, we have dramatically extended the model to a truly multi-sector economy where each heterogeneous industry's growth is governed by a set of linear equations that can be influenced by regulatory shocks. To our knowledge, ours is the first multi-sectoral endogenous growth model with closed form solutions. We then estimate our model using industry data from the Bureau of Economic Analysis (BEA) and the Census Bureau (Census) in combination with RegData 2.2, further described below.

A key insight of endogenous growth models in general is that the effect of government intervention on economic growth is not simply the sum of static costs associated with individual interventions. One recent proposal regarding the phenomenon of regulatory accumulation reframed this insight with respect to regulation, pointing out that when regulations are created in reaction to major events, "[N]ew rules are [placed] on top of existing reporting, accounting, and underwriting requirements. [...] For each new regulation added to the existing pile, there is a greater possibility for interaction, for inefficient company resource allocation, and for reduced ability to invest in innovation. The negative effect on U.S. industry of regulatory accumulation actually compounds on itself for every additional regulation added to the pile" (Mandel and Carew 2013).

For its part, the federal government requires regulatory agencies to perform regulatory impact analyses of significant, new regulations. In theory, these analysts estimate the marginal impacts—positive or negative—that the new rule would precipitate. However, these regulatory impact analyses are only performed for a very small portion of new regulations, they rarely consider interactive or cumulative effects, and their overall quality falls far short of the best practices specified in the executive orders and OMB guidance documents related to the process (Ellig and McLaughlin 2011).

More rigorous studies of regulations sometimes originate in the academy, where scholars tend to focus on the empirical effect of some particular regulation[s] on a limited scope of industries. For example, researchers have given considerable attention to the employment effects on manufacturing

[4] For more evidence, see Madsen (2010) or an earlier analysis in Madsen (2008).

industries of particular environmental regulations from the Clean Air Act.[5] Another case that has been heavily studied is the deregulation of the transportation and telecommunication industries during the mid-1980s; findings broadly concur with deregulation resulting in a significant surge in investment for the United States and the U.K. relative to Italy, France, and Germany (Alessina, et al. 2005). However, compared to the myriad of regulations that actually affect the economy, these [and most other] regulatory studies focus on interventions that are relatively limited in scope or on economic outcomes related to a narrowly defined sector. Such studies may not be able to see beyond the immediate negative (or positive) effects that regulation imposes on the decisions of firms within the industry.

Second, in addition to contributing to endogenous growth theory, this study is the first in this literature to utilize a novel panel dataset that measures sectoral regulation over time: RegData 2.2 (McLaughlin and Sherouse 2015). This rich data source offers a multi-sector panel that quantifies federal regulation by industry from 1970 to 2014. RegData's metrics of regulation are created using text analysis programs run on the entire corpus of federal regulations in effect in each year; its primary improvement over other datasets is its association of regulations with affected industries in an industry-specific data series spanning multiple decades. This industry-specific dataset allows us to utilize panel data methods to identify the effects of regulation more confidently than if we only had national data. By using these data, we also can produce industry-specific estimates of the effects of regulation on the key decisions of firms: investment, output, and [net] entry. We combine industry regulation data from RegData 2.2 with measures of industry inputs and outputs taken from other data sources to identify estimates for enough of the parameters in our model to conduct valid counterfactual experiments.

Most importantly, our careful combination of modelling and data have enabled us to estimate the effects of regulation on investment in an endogenous growth context. In endogenous growth theory, innovation is not an exogenous gift from the gods but rather the result of costly effort expended by firms to realize gains. The growth generated by that entrepreneurship

[5] See Greenstone (2002) or Morgenstern et al. (2002) for heavily cited examples or Walker (2013) for a more recent example.

can be thwarted by misguided public policy. By deflecting firm resources away from the investments that maximize the stream of profits, and towards regulatory compliance, regulations can theoretically slow the real growth of an industry. Indeed, each additional [binding] regulation should be seen as a [binding] constraint upon the firm's profit maximization problem. Yet, some regulations can actually succeed in enhancing growth. For instance, if a regulation does manage to efficiently correct a market failure, then some deadweight loss in output will be recovered by that regulation. Alternatively, a regulation requiring the use of certain inputs like safety equipment can increase investment in the stock of that capital. Naturally, these regulations can cause a select set of industries to grow faster, at least in the short-run.[6]

These countervailing effects of regulation can now be handled in an empirical study due to RegData providing separate observations for multiple sectors over a relatively long timeframe. We show that the average industry does indeed experience a slower growth rate as a result of the accumulation of regulations that target the industry. However, a handful of industries do appear to grow faster because of regulation. Nonetheless, our results are dominated by the harms inflicted by regulation's distortion of firms' investment decisions. On net, regulations slows the growth of the entire economy by an average of 0.8 percent per annum. Because economic growth is an exponential process, this seemingly small figure grows powerfully over time into a truly dramatic difference in the level of GDP per capita. Had regulations been held constant at levels observed in 1980, our model predicts that the economy would be nearly 25% percent larger. In other words, the growth of regulation since 1980 has cost the U.S. roughly $4 trillion in GDP (which amounts to nearly $13k per person) in 2012 alone.

Of course, our estimate says little about the benefits of regulation, aside from those that are captured in GDP. Some regulations may lead to benefits, such as improvements in environmental quality, which are well-known to be mainly missing from GDP measurements. Other regulations, such as

[6] For example, the output of corn farmers has arguably increased in response to US regulations requiring that gasoline contains a minimum percentage of ethanol, which is primarily derived from corn.

regulations designed to prevent monopolistic practices, or even those that are designed to improve human health, may only be captured in GDP to a limited degree. For example, if regulations decrease employee absenteeism because they reduce asthma-inducing air pollution, we would expect a marginal increase in GDP. Nonetheless, we caution the reader that this study is not a weighing of the costs and benefits of regulation. It is an examination of how regulatory accumulation in specific sectors of the economy affects the growth path of those sectors. From our findings, we can deduce how the effects of regulation can change the national economy's growth path.

The remainder of this study proceeds as follows. In Section 2, we describe how we measure regulation, as well as the other data that we utilize in our estimation. Our theoretical model is outlined in Section 3. Section 4 details our estimation methodology. We explain our results in Section 5. Section 7 concludes our discussion.

2 Data

Because RegData 2.2 is fairly new, we devote some time to explaining it here. The Code of Federal Regulations (CFR) is published annually and contains all regulations issued at the federal level. A regulation may be in effect for up to one year prior to actual publication in the CFR, but all regulations are ultimately published in the CFR. Among other things, RegData 2.2 searches the entirety of the CFR in each year from 1970 to 2014 for a vocabulary of industry-specific terms and phrases that indicate that an industry is targeted by the regulatory text. We have used RegData 2.2's industry regulation index as our metric of regulation.

RegData 2.2 quantifies two essential elements of regulatory text. First, it quantifies regulatory restrictions relevant to specific industries in the Code of Federal Regulations (CFR), the set of documents published annually that contain all regulations that are in effect each year. Restrictions are those words used in legal language to either obligate or prohibit an action (e.g. "shall"). Version 2.2 of RegData relies upon machine-learning algorithms to classify chunks of text in the CFR according to their relevance to specific industries, as defined by the North American Industry Classification System (NAICS). The RegData 2.2 program was trained to identify the relevance of regulatory text to specific industries using some selected documents from the Federal Register.

RegData 2.2 quantifies two essential elements of regulatory text. First, it quantifies regulatory restrictions relevant to specific industries in the Code of Federal Regulations (CFR), the set of documents published annually that contain all regulations that are in effect each year. Restrictions are those words used in legal language to either obligate or prohibit an action (e.g. "shall").[7] Version 2.2 of RegData relies upon machine-learning algorithms to classify chunks of text in the CFR according to their relevance to specific industries, as defined by the North American Industry Classification System (NAICS). The RegData 2.2 program was trained to identify the relevance of regulatory text to specific industries using selected documents from the Federal Register.[8]

Documents were analyzed using a vocabulary of 10,000 words learned from the training documents. RegData 2.2 uses Logit-based classifications that were first made available in RegData 2.1 and have outperformed competing classification schemes.[9] This classification methodology yields a set of probability scores ranging from 0 to 1 for each CFR part – a legal division

[7] RegData specifically searches for a subset of all restrictions, consisting of the strings, "shall," "must," "may not," "prohibited," and "required." While this subset is not comprehensive of all ways in which a restriction can be created with legal language, it is probably representative of the restrictiveness of regulatory text (Al-Ubaydli and McLaughlin 2015).

[8] The Federal Register is a daily publication of the federal government which includes rules, proposed rules, presidential documents, and a variety of notices of current or planned government activity. Some of these documents are specifically labeled with relevant NAICS codes, and the language they employ is similar to that of the CFR. Training documents for each 3-digit NAICS industry were obtained by searching the Federal Register API for an exact match for the word "NAICS" and the code for the 3-digit code and each 4, 5, and 6-digit code it contains. Additionally, the exact names of the 3-digit industries and their children industries were used to identify documents. These searches yielded approximately 24,500 documents which were associated with at least one NAICS 3-digit industry. Industries with fewer than 20 positive training documents were excluded from analysis.

[9] RegData 2.1—the iteration just previous to the version we use—explored several well-established methods of classification: Support Vector Machines (SVM) with a linear kernel, Logistic Regression (Logit), Random Forests, and K-Nearest Neighbors. All such classification can be carried out using the a single toolkit: scikit-learn (Pedregosa, et al. 2011). Cross-validation tests and other evaluation scores indicated that the Logit model outperformed all other models, albeit by a relatively narrow margin (McLaughlin and Sherouse 2015).

of text that typically houses a regulatory program. Since the CFR is published annually, RegData offers probability scores for each CFR part in each year from 1970 to 2014. A probability score reflects the probability that a given part is relevant to a given industry.[10]

[FIGURE 1]

Figure 1 compares the growth of the stock of federal regulations, in aggregate as measured by RegData as well as the next best alternative that has been used in the literature: counting pages in the Code of Federal Regulations (Dawson and Seater 2013). Figure 2 graphs the timepath of regulation across our 22 industries, which reveals considerable within and between variation from trend:

[FIGURE 2]

We also gathered some fairly standard data from BEA and Census. All three data sources organize industries according to the North American Industry Classification System (NAICS). We have assembled data at the 3-digit NAICS level. However, the BEA data combines certain industries together (e.g., BEA combines industries 113, 114, and 115), so we have correspondingly combined regulation data from each industry to match the BEA's system.[11] Once the full dataset was assembled, it contained 22 industries, which are listed in Table 1.

[TABLE 1]

Our assembled dataset is summarized in Table 2. All variables are reported in levels, except for our measure of regulation. As specified in our estimation equations, regulation is logged.

[TABLE 2]

[10] As in the first edition of RegData, each part's probability score is then multiplied by the number of restrictions contained in the same part and then summed across all parts for each agency (Al-Ubaydli and McLaughlin 2015).

[11] There were similar concerns that we had to address for translating between NAICS and SIC codes. Also, the Census data that we gathered did not contain the full set of industries covered by the BEA's data for value-added to GDP.

Given the panel nature of our dataset, and the absence of a plausible instrument, we perform Granger causality testing focused on the relation between regulation and each of the six different dependent variables that we ultimately use in our regressions: the four types of real investment (equipment, IP, structures, and fixed capital), real value-added to GDP per firm, and people per firm. We find regulation Granger causes investment in equipment for 12 of the industries included; regulation Granger causes investment in IP for 15 of the industries; regulation Granger causes investment in structures for 9 of the industries; and regulation Granger causes investment in fixed capital for 15 of the industries. Regulation Granger causes real value-added to GDP per firm for 15 of the industries as well, and it Granger causes people per firm for 11 of the industries. The industries for which we fail to reject the null hypothesis are not the same across the four tests. There is no industry for which regulation does not Granger cause at least one of the dependent variables. For the average industry, regulation Granger causes 3.5 of the dependent variables—a fact that makes us relatively confident that the failure to reject is a result of noise rather than a reverse causal or non-causal relationship.

3 Model

This section builds a multi-sector endogenous growth model beginning with microeconomic foundations.

3.1 Overview

We consider an economy populated by identical individuals who supply labor services from their unit of time endowment in a competitive labor market and spend their earnings on consuming final goods. Their decisions yield an optimal path of expenditures and savings via freely borrowing and lending in a competitive market for financial assets at the prevailing interest rate. The households' income consists of wage income and returns on asset holdings. Government policy (i.e. regulation and spending) can convey benefits to the representative household but those benefits are modeled to be separable in the household's preferences.

The production side of the economy consists of a final goods sector, commodity producers in J industries, and the intermediate good suppliers within those J industries. The final goods sector has a representative firm that produces final goods from the J commodity outputs from the J different industries. Likewise, each industry's commodities are produced by a representative firm that aggregates intermediate goods. Also within each industry, there is a continuum of identical firms that make differentiated intermediate goods from labor and invest in gaining additional knowledge to reduce their production costs. As each firm invests in knowledge, it contributes to the pool of public knowledge that benefits the quality of all intermediate goods firms in the industry. The engine for endogenous growth in this economy is the increasing returns from this public knowledge and the variety of intermediate goods that grow with the proliferation of intermediate goods firms.

3.2 Households Primitives

The population, P(t) of identical individuals is exponentially growing at rate ω from an initial level of P_0. At time t, each individual maximizes the discounted stream of their dynastic utility:

$$\int_{t}^{\infty} e^{-(\rho-\omega)(s-t)} [\ln c(s) + \lambda \ln(1-l(s)) + u(G,R,D,P)]ds \tag{1}$$

Where the time arguments to functions has been suppressed for brevity, ρ is their subjective discount rate, c is per capita consumption, l is the fraction of time allocated to work (1-l is leisure), λ governs preference for leisure, and u represents any (dis)utility received from combining public debt (D) with regulations (R) as well as government spending (G) for the population (P). The individual faces the following flow budget constraint:

$$a = [r(1-\tau_A) - \omega]a + (1-\tau_L)Wl - (1+\tau_C)p_Y c \tag{2}$$

Where a is the individual's asset holdings, r is the rate of return to savings (and the [after-tax] rate of return to equity), τ_A is the tax rate on assets, W is the wage rate, τ_L is the tax rate on labor income, τ_C is the tax rate on consumption, and the price of the final goods that the individual consumes is p_Y. From this objective and budget constraint, we derive the household's labor supply and Euler equation in the Appendix.

3.3 Final Goods Producers Primitives

The representative firm produces final goods (Y) with a CES technology:

$$Y = \left[\sum_{j=1}^{J} \psi_j Y_j^{\frac{\xi-1}{\xi}} \right]^{\frac{\xi}{\xi-1}} \tag{3}$$

Exhibiting constant returns to scale (i.e. $\Sigma\psi_j = 1$), this final goods sector is perfectly competitive and generates no profits. It serves only to assemble final goods from the commodity outputs of J different industries.

3.4 Industry Commodities Producers Primitives

Each industry has a representative firm that produces its commodity (Y_j) from a mass of differentiated intermediate goods (X_{ij}):

$$Y_j = N_j \left[\int_0^{N_j} \frac{1}{N_j} X_{ij}^{\frac{\chi-1}{\chi}} di \right]^{\frac{\chi}{\chi-1}} \tag{4}$$

Where N_j is the mass of intermediate goods firms for industry j and χ is the elasticity of substitution. The representative commodity producer for industry j maximizes its profits:

$$\Pi_j = p_{Yj} Y_j - \int_0^{N_j} p_{Xij} X_{ij} \, di \tag{5}$$

Because this sector is perfectly competitive, the industry commodity producer earns zero profit. The industry commodity producer's demand for the intermediate good is implicitly given by:

$$p_{xij} X_{ij} = p_{Yj} Y_j \left[\frac{p_{xij}^{1-\chi}}{\int_0^{N_j} p_{xij}^{1-\chi} di} \right] \tag{6}$$

3.5 Intermediate Goods Producers Primitives

The typical intermediate firm produces its differentiated good with the following technology:

$$X_{ij} = Z_{ij}^{\zeta}[L_{Xij} - \phi_j] \tag{7}$$

Where Z_{ij} is the labor-enhancing knowledge specific to the firm, L_{Xij} is the labor employed in producing X_{ij}, and ϕ_j is a fixed labor cost. The firm accumulates knowledge according to the technology:

$$Z_{ij} = \kappa K_j \mathrm{L}_{Zij} \tag{8}$$

Where L_{Zij} is the quantity of labor invested in knowledge accumulation and K_j is the stock of public knowledge in the industry. Public knowledge, which is non-rival and non-excludable, accumulates via spillovers:

$$K_j = \max_i Z_{ij} \tag{9}$$

In equilibrium, under the assumption of within-industry symmetry of identical intermediate goods firms, the stock of firm-specific knowledge will end up equaling the stock of public knowledge over the industry (i.e. $Z_{ij} = K_j$ $\forall$i). The firm's pre-tax flow of profits is its flow of revenue net contemporaneous production costs, knowledge accumulation investment (I_{Zij}), and its fixed cost (F_{ij}):

$$\Pi_{ij} = p_{Xij}X_{ij} - WZ_{ij}^{-\zeta}X_{ij} - \underbrace{WL_{Z_{ij}}}_{\mathrm{I}_{Zij}} - \underbrace{W\phi_j}_{\mathrm{F}_{ij}} \tag{10}$$

The firm maximizes its value, which equals the present value of their stream of after-tax profits:

$$\mathrm{V}_{ij}\ t\ = \int_t^{\infty} e^{-\int_t^s [r\ \nu\ + \delta_j] d\nu}\ 1 - \tau_\pi\ \Pi_{ij}\ s\ ds \tag{11}$$

where τ_π is the tax rate on corporate income, r is the interest rate (return to savings), and δ_j parameterizes the hazard of a death shock. The firm maximizes V_{ij} subject to the knowledge accumulation process (8), a given initial stock of knowledge Z_{ij} 0 > 0, and a non-negativity constraint on the growth of knowledge. The solution to this problem, detailed in the Appendix, yields the (maximized) value of the firm given exogenous influences.

3.6 Government Primitives

The government sets tax rates, determines its expenditure on investment (as well as consumption), and regulates. Because the government's behavior is purely exogenous, we have not engaged in modeling it.[12] Nonetheless, it is noteworthy that our limited model of the government does not necessarily reduce it to a useless drain on the economy. Government spending, as well as regulation, can be net beneficial at some level depending on its magnitude in the household's separable utility function where it directly enters.

3.7 Solution for Equilibrium

The model primitives result in supply and demand equations for the goods markets, the labor market, and the asset market. Bringing in market clearing conditions, we attain a solution for the equilibrium as detailed in the Appendix. In this section, we simply describe the 3 equations in the solution that motivate our empirical specification.

---

[12] We could describe a budget constraint for the government that is consistent with our model. Suppose that any deficit gets supplied by net imports (M) and accumulates into debt (D) held by foreign investors willing to inject real wealth into the economy in return for that principal wealth plus interest (at the market rate) financed by future tax receipts:

$$D = [rD + Mp_Y] + [G - M]p_Y + WL_G - \left[\tau_L WL + \tau_C C + \tau_A rPa + \sum_{j=1}^{J} N_j \tau_\pi \Pi_{ij}\right]$$

Where G is the quantity of final goods purchased by the government, rD is the repayment to the (foreign) owners of the debt, and L_G is the units of labor hired by the government (which presumably comes from some underlying Leontief production function that requires the government to hire at least some labor in order to process public investment and consumption, although it may well hire more labor than required).

Real expenditure by firms on investment per establishment, equal to the labor allocated to knowledge accumulation times the wage rate, is a linear function of the real interest rate (which we define as the nominal interest rate on long-term Treasury Bills net the growth in nominal wages) and real output per establishment:

$$\frac{I_j}{N_j W} = \left[-\frac{\delta_j}{\kappa_j}\right] + \left[-\frac{1}{\kappa_j}\right]\left(r_{Zj} - \frac{W}{W}\right) + \left[\zeta\left(\frac{\chi-1}{\chi}\right)\right]\frac{p_{Yj}Y_j}{N_j W} \tag{12}$$

The state equation that drives the dynamics of the model is fairly simple: the change in real output per establishment is a linear function of its level, as well as real investment per establishment and an interaction between the real exchange rate with the level of real output per establishment:

$$\begin{aligned}\left(\frac{p_{Y_j}Y_j}{N_j W}\right) = &\left[-\psi_j\left(\frac{1-\tau_\pi}{\theta W^*}\right)\left(\frac{1}{\chi} - \zeta\left(\frac{\chi-1}{\chi}\right)\right)\right] \\ &+ \left[-\psi_j\kappa_j\right]\left(\frac{I_j}{N_j W}\right) \\ &+ \left[-\psi_j\zeta\left(\frac{\chi-1}{\chi}\right)\right. \\ &\left.- \frac{1-\tau_\pi}{\theta\kappa\psi_j}\frac{W^*}{}(\delta_j - \kappa_j\phi_j)\right]\left(\frac{p_{Yj}Y_j}{N_j W}\right) \\ &+ \left[-\frac{1-\tau_\pi}{\theta\kappa\psi_j}\frac{W^*}{}\right]\left(\left(r - \frac{W}{W}\right) \times \left[\frac{p_{Yj}Y_j}{N_j W}\right]\right)\end{aligned} \tag{13}$$

According to the model, as detailed in the appendix, the number of people per establishment should follow the exact same process.

4 Estimation Methodology

Our structural model can be estimated via a few simple linear regressions, where the specification for each is explicitly derived from the model. To operationalize (12) into a regression, we replace the square bracketed terms into composite parameters and include a disturbance term. In our data, we actually have 4 different measures of investment: equipment, intellectual property, structures, and fixed capital. Some of that investment would map to the knowledge accumulation investment in the model while other investment would map to the creation of additional establishments (for which the model specifies a different variable to represent that effort); rather than trying to reproduce that mapping in an ad hoc fashion, we have simply used all four of these measures and let the data identify the contribution of each. Hence, this yields 4 different investment regressions that have essentially the same form.

Taking an agnostic approach about exactly how investment changes with regulation, we want a specification that can flexibly capture how regulations tend to affect the parameters governing firms' investment decisions as a sort of reduced-form of the deeper underlying mechanisms at work. Hence, we include regulation in the most flexible form that is practical for our data; we include the log of the current level of regulation, as well as its lag, as independent regressors and interacted with each of the other regressors specified by the theoretical model:

$$\begin{aligned}\left[\frac{I_{jst}}{N_{jt}W_t}\right] = {} & \beta_{0js} + \beta_{1js}\left[r_t - \frac{W_{t+1} - W_t}{W_t}\right] + \beta_{2js}\left[\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\right] \\ & + \beta_{3js}\log R_{jt} + \beta_{4js}\log R_{jt-1} \\ & + \beta_{5js}\left[\left(r_t - \frac{W_{t+1} - W_t}{W_t}\right) \times \log R_{jt}\right] \\ & + \beta_{6js}\left[\left(r_t - \frac{W_{t+1} - W_t}{W_t}\right) \times \log R_{jt-1}\right] \\ & + \beta_{7js}\left[\left(\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\right) \times \log R_{jt}\right] \\ & + \beta_{8js}\left[\left(\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\right) \times \log R_{jt-1}\right] + \varepsilon_{jst}\end{aligned} \tag{14}$$

The t subscript, indicating the year, has been resurrected here; we also introduced a new subscript (s) to indicate each of the 4 types of investment. Note that the econometric model is specified in real terms (units of labor) from the model, which does not include inflation. However, by normalizing all prices by the average nominal wage rate for the nation (i.e. total annual labor earnings divided by number of full-time and part-time workers) all of these measures are now in real terms -- both in terms of units of labor and with inflation effectively removed. Our estimations and counterfactuals are then conducted using these real units but the results can then easily be transformed back into nominal terms and then into the more conventional measure of real terms as desired (e.g. using BEA's GDP deflator).

As mentioned in the modeling section, the state equation that drives the dynamics of the model naturally lends itself to an AR(1) in the real output per firm. We proceed in operationalizing (15) in a manner similar to the investment equation; we operationalize this into a regression by replacing the square bracketed terms into composite parameters, include a disturbance term, include each of the investment measures, and interact each term with both the log of regulations and its lag:

$$
\begin{aligned}
\left[\frac{p_{Yjt+1}Y_{jt+1}}{N_{jt+1}W_{t+1}}\right] &= \alpha_{0j} + \sum_s \alpha_{1js}\left[\frac{I_{jst}}{N_{jt}W_t}\right] + \alpha_{2j}\left[\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\right] \\
&+ \alpha_{3j}\left[\left(r_t - \frac{W_{t+1}-W_t}{W_t}\right)\times\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\right] + \alpha_{4j}\log R_{jt} \\
&+ \alpha_{5j}\log R_{jt-1} + \sum_s \alpha_{6js}\left[\left(\frac{I_{jst}}{N_{jt}W_t}\right)\times\log R_{jt}\right] \\
&+ \sum_s \alpha_{7js}\left[\left(\frac{I_{jst}}{N_{jt}W_t}\right)\times\log R_{jt-1}\right] \\
&+ \alpha_{8j}\left[\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\times\log R_{jt}\right] \\
&+ \alpha_{9j}\left[\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\times\log R_{jt-1}\right] \\
&+ \alpha_{10j}\left[\left(r_t - \frac{W_{t+1}-W_t}{W_t}\right)\times\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\times\log R_{jt}\right] \\
&+ \alpha_{11j}\left[\left(r_t - \frac{W_{t+1}-W_t}{W_t}\right)\times\frac{p_{Yjt}Y_{jt}}{N_{jt}W_t}\times\log R_{jt-1}\right] \\
&+ \varepsilon_{0jt}
\end{aligned}
\tag{15}
$$

As previously mentioned, the number of people per establishment follows the same process. Hence, we estimate a nearly identical AR(1), with these same covariates, for people per establishment (where real GDP per establishment on the left hand side, as well as its appearance on the right hand side, gets replaced by people per establishment and its lag).

In theory, each of these industry-specific regressions could simply be estimated by OLS. However, because we have only 36 data points for each regression (34 once we lose observations to construct leads and lags), OLS runs the risk of over-fitting the data when the regressions have between 9 parameters (for the investment equations) and 22 parameters (for the AR(1) equations). The risk with such over-fitting is that out-of-sample predictions on these linear equations can explode beyond reasonable bounds. Indeed, we have included simple OLS results in our plots (as red lines) and we do witness a few such explosions. A simple pooling of all of the data would yield more degrees of freedom but would be pooling industries that are very heterogeneous, as well as somewhat diluting our inter-industry identification strategy.

Instead, we make the assumption that each industry's parameter is drawn from a normal distribution across all industries (called a hyper-distribution). This sort of hierarchical model structure represents a compromise between completely pooling the data and running strictly separate regressions. Hierarchical models enable a partial pooling of information across industries – when an industry's own data is too noisy then the parameter values to which OLS would have overfit get shrunk back towards the grand mean (a parameter of the hyper distribution, so called a hyper-parameter) so that outlying values for a parameter become unlikely.

Summing up the size of the estimation problem, the estimates will include not only 22 industry-specific estimates of the 9×4 + 21×2 parameters (plus 22 estimates of each of 6 disturbance variances to yield a grand total of 1,848 parameters); the estimates also include an estimate of 156 hyper-parameters: mean and standard deviation of the normal hyper-distribution of a parameter's values across industries. With a grand total of 2,004 parameters coupled in the likelihood through the hyper-distribution, MLE would practically be [computationally] infeasible. The easiest way to fit such a hierarchical model is via Bayesian Markov Chain Monte Carlo (MCMC) with the hyperdistribution moved from the specification of the likelihood to the specification of the prior.

Hence, we take a Bayesian approach to estimation whereby draws from a posterior distribution (of uncertainty over parameter values having observed the data) are generated from sampling the product of the likelihood (of observing the data given the parameter values) and the prior (distribution of uncertainty over parameter values) in accordance to Bayes' rule defining conditional probabilities. Bayesian estimation is equivalent to classical estimation when the priors are flat—meaning that the prior is a constant so that the posterior is just the likelihood (MLE finds the maximum of that likelihood and invokes a central limit theorem to get a normal distribution of uncertainty around that point estimate, while Bayesians would just use the likelihood itself as the uncertainty distribution with the mode, mean or median serving as equally good point estimates). True to our classical roots, we employ flat priors so that the only prior information that we impose upon this estimation is that the industry-specific parameters are drawn from a common normal hyper-distribution (whose hyper-parameters are also given flat priors so that we are completely agnostic on the degree of similarity between these industry parameter values).

Once we have estimated these equations, we can make predictions on GDP by simply dividing real GDP per establishment by the number of people per establishment and then scaling the resulting rational function by the population (and the nominal wage if we desire to have GDP in nominal terms). Because the residuals are orthogonal to the regressors by construction, then presumably those unobserved shocks would be unchanged by the counterfactual and can be added to the resulting estimate to retain those features in the data (e.g., the financial crisis) when comparing the factual to the counterfactual.

5 Results

Prior to discussing the results of our counterfactual experiment, we examine the goodness of fit for our econometric models. Ultimately, our purpose is to construct a forecast for the counterfactual of a different regulatory regime: one that "freezes" regulations at some point in the past, similar to the leading approach in the literature (Dawson and Seater 2013). Hence, what matters most is not hypothesis testing on individual parameters and providing an insightful interpretation of such results; indeed there are too many parameters to realistically review within the confines of a single academic article (see tables 4 through 9 for summary statistics across the Monte Carlo uncertainty draws on the hyper-parameters as our attempt to report such a vast quantity of parameters). Instead, what matters for our purposes is a good fit to the within-sample data and a reasonable argument for the external validity of our out-of-sample forecasting, which is justified by our theoretical model.

[TABLES 4 – 9]

The goodness of fit is apparent in figures 3 through 8. Each figure presents the results for an estimation equation: four equations for which each type of investment serves as the outcome variable, one AR(1) equation for real output per establishment as the outcome variable, and one AR(1) equation for persons per establishment as the outcome variable.

[FIGURES 3 – 8]

Each figure is divided into 22 facets – one for each industry in our dataset. For each facet, the observed outcome variable over time is plotted as black dots. A jagged blue line plots the mean prediction for the outcome variable over time across 1,000 uncertainty draws from the posterior distribution on the tuple of parameters estimated for our Bayesian hierarchical model. A jagged red line plots that same prediction when OLS is used to estimate the parameters (without a hyperdistribution in a hierarchical structure) as a basis for comparison. In the top corners of each facet is the R^2 summary statistic in color corresponding to the respective method used to generate the predictions. In summary, figures 3 through 8 show goodness of fit plots; each figure has observed data plotted as black dots, Bayesian predictions plotted as the blue line, OLS predictions plotted as the red line, and the R^2 summary statistic for each of the two methods.

We find the fits to be acceptable across the board. In general, the blue fits of our Bayesian model appear less noisy, which is attributable to the hierarchical structure that effectively down-weights the temptation to overfit when the underlying signal in the data is less informative and instead shrinks the prediction towards the grand mean. The R^2 often gets higher for OLS because it is not constrained by the joint distribution of the industry-specific parameters and hence free to over-fit the data. The fits are substantially better when examining the dynamic estimation equations for real output per establishment and people per establishment. This should be expected because these dynamic estimation equations have an AR(1) structure and an economist's ex ante expectation here would be for considerable persistence in these outcome variables. Nonetheless, we are still pleased to have such tight fits here because our counterfactual forecast ultimately produces predictions for GDP as a rational function that put real output per establishment in the numerator and people per firm in the denominator. With stochastic disturbance terms that are normal and mean zero, this implies that the noise in their ratio will have the much fatter tails of the Cauchy distribution, so obtaining a low variance of the stochastic terms in these dynamic equations is highly desirable.

The counterfactual predictions themselves are presented in figure 9 through figure 17. Figure 9 displays both the actual time path of regulation and our counterfactual experiment on regulation where we initiate a freeze in the first year of the Reagan administration (i.e. regulation in 1981 is fixed at 1980 levels) and sustain that freeze through the last year in our study (near the end of Obama's first term). We have selected so broad of a window within our study in an effort to be more comparable with the leading estimate in the literature, where that counterfactual experiment was a regulatory freeze from the late 1940s until the early 21st century (Dawson and Seater 2013). The results of our Granger Causality tests for the exogeneity of regulation, in that most tend lean in our direction but are limited by the sample size and test's power, are qualitatively similar to those in the existing literature (Dawson and Seater 2013). Hence, we remain comfortable with the assumption that regulation is exogenous.

[FIGURE 9]

Our counterfactual simulation then proceeds 1 year at a time. First, we predict the amount of each of the 4 types of investment from that year's real output per establishment, the real interest rate (shocks to which have been assumed to be strictly exogenous), and our measure of the amount of regulation (which, for the counterfactual, were frozen at the levels observed in 1980). To each prediction, we add in the shocks (i.e. the residuals from the regressions), which are orthogonal by construction but help avoid false comparisons between our counterfactual and the observed data by capturing major business cycle events (like the great recession) that would not be produced by the model. Then we perform a similar prediction of the next year's persons per establishment and real output per capita from the current year, the current year's investment (all 4 types), regulation, and real interest rate. Once we have made all of the predictions of outcome variables in an industry in a year, we repeat this process for the next year and then repeat this in an outer loop across all industries. This results in the counterfactual prediction for each of the 4 types of investment (figure 10 through figure 13) and the counterfactual predictions for real output per establishment and persons per establishment (figure 14 and figure 15).

[FIGURES 10-15]

These figures are also separated into 22 different facets to enable better examination of the each industry. The black jagged line plots the observed time path for the outcome variables, the blue jagged line plots the mean prediction from our Bayesian model, and the red jagged line is the OLS prediction which we have retained as a basis for comparison. In general, the counterfactual predictions appear reasonable with the Bayesian predictions tending to be milder than the OLS predictions. However, there are a few industries that have a few variables that explode for the OLS forecasting – either implausibly strong exponential growth that dwarfs the observed data into a flat line or an exponential dive below 0. We interpret these findings as the hazards of using OLS with so few degrees of freedom in a given industry. Essentially, this sort of undesirable behavior justifies our use of the more complicated Bayesian estimation technique because estimates for individual industries will be tied to the grand mean across industries, making it unlikely for an individual industry's predictions to dramatically diverge from the rest of the pack.

Figure 16 displays the prediction for value-added to GDP for each industry that results from our dividing of the counterfactual prediction for real output per establishment by establishments per person (and then scaled up by population, which is assumed to be exogenous). Again, the results appear reasonable except for those few industries that suffer from the OLS predictions exploding outside of the reasonable range. To get the aggregate effect across these 22 industries, we can simply add up our counterfactual predictions, as depicted in figure 17. Note that the OLS explosions take their toll on the aggregate predictions in red, which suffer from an implosion followed by an explosion towards the end of the simulation. Hence, we focus on our Bayesian results to get some insight on the effects of a counterfactual regulatory freeze.

[FIGURES 16 & 17]

In aggregate, we find about 0.8 of a percentage point in the (real) growth rate is lost from the additional regulations passed since 1981. To the extent that we are comfortable treating these 22 industries as representative of the rest of the economy, our result would scale up to a similar loss in aggregate GDP: 0.8 of a percentage point. Over the course of 30 years, this grows into nearly $4 trillion in GDP (nearly $13k per capita). Although our finding is slightly less than one-half of the result in the literature (Dawson and Seater 2013), it is still within a reasonable range. We caution that our results are for the costs of regulation to the measurable economy (net any benefits to that measurable economy) but that does not imply that none of the regulations passed since 1981 have been net beneficial. Indeed, many regulations exist to generate non-market benefits that would probably not appear in our measurable economy. Nevertheless, this suggests that a wide scale review of regulations could be greatly beneficial if it sunsetted redundant regulations and supplanted command-and-control regulations with simpler market-based mechanisms.

6 Conclusion

Though the topic of regulation and economic growth has been widely studied, most studies focus on a narrow set of regulations, industries, or both. Such designs cannot estimate the cumulative effect of regulation, even though accumulation of increasing complexity of regulatory constraints is clearly a salient characteristic of the regulatory regime in the U.S. The analysis of individual regulations is analogous to the choice to throw a rock into a stream. Throwing one single rock may seem like a good idea, because now no one will trip on it. But as more and more rocks are thrown into the stream and accumulate, eventually the stream's flow is diverted or damned to a halt. Similarly, a single regulation can appear net beneficial when examined on its own—indeed, government agencies typically claim that all, or nearly all, of their regulations create positive net benefits—but still have a net negative effect on economic growth by virtue of being piled on top of (and interacting with) other regulations. Although government agencies do not have much of an incentive to count these cumulative effects in their costs, it is only fair to acknowledge that economists have not provided these agencies with the sort of tools that they would need to quantify such costs of their regulations.

We model this cumulative effect in the context of endogenous growth. We are able to estimate our model with a panel dataset of 22 industries observed annually from 1977 to 2012. Specifically, we examine not only the direct effect of regulation on the output per establishment and [net] entry of additional establishments but also the indirect effect of regulation on four different types of investment that in turn affect the output per establishment and [net] entry of additional establishments. Our dataset uses BEA and Census measures of these industries, combined with novel measures of regulation by industry provided by RegData's text-analysis-based quantification of regulations impacting industries over time. Our results suggest that regulation has been a considerable drag on economic growth, on the order of 0.8 percentage points. Our counterfactual simulation predicts that the economy would be about 25% larger than it currently is, if regulations had been frozen at levels observed in 1980. This amounts to about $4 trillion, or $13,000 per capita.

The empirical investigation of the relationship between all federal regulations and a broad panel of industries is unprecedented and has only recently been made possible by the advent of RegData. Further research along these lines are also possible with RegData. For example, RegData permits the user to separate regulations according to the agency that created them. Researchers could therefore examine the effects of regulations created by labor, environmental, or financial regulatory bodies in isolation or in comparison to other sorts of regulations.

References

Aghion, P., Y. Algan, P. Cahuc, and A. Shleifer. 2010. “Regulation and Distrust.” *Quarterly Journal of Economics* 125: 1015-1049.

Alessina, A., S. Ardagna, G. Nicoletti, and F. Schiantarelli. 2005. “Regulation and Investment.” *Journal of the European Economic Association* 3: 791-825.

Al-Ubaydli, O., and P. McLaughlin. 2015. “RegData: a numerical database on industry-specific regulations.” *Regulation and Governance.*

Botero, J., S. Djankov, R. La Porta, F. Lopez-de-Silanes, and A. Schleifer. 2004. “The Regulation of Labor.” *Quarterly Journal of Economics* 125: 1339-1382.

Dawson, J., and J. Seater. 2013. “Federal Regulation and Aggregate Economic Growth.” *Journal of Economic Growth* 18: 137-177.

Djankov, S, C McLiesh, and R Ramalho. 2006. “Regulation and Growth.” *Economic Letters.*

Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. “The Regulation of Entry.” *Quarterly Journal of Economics* 117: 1-37.

Ellig, J., and P. McLaughlin. 2011. “The Quality and Use of Economic Analysis in 2008.” *Risk Analysis* 32: 855-880.

Glaeser, E., and A. Shleifer. 2003. “The Rise of the Regulatory State.” *Journal of Economic Literature* 41: 401-425.

Gørgens, T., M. Paldam, and A. Würtz. 2003. *How does public regulation affect growth?* University of Aarhus.

Greenstone, M. 2002. “The Impacts of Environmental Regulations on Industrial Activity.” *Journal of Political Economy* 110: 1175-1217.

Loayza, N. 2005. “The Impact of Regulation on Growth and Informality: Cross-Country Evidence.” *World Bank Policy Research Working Paper No. 3623.*

Madsen, J. 2008. “Semi-endogenous versus Schumpeterian growth models: Testing the knowledge production function using international data.” *Journal of Economic Growth* 13: 1-26.

Madsen, J. 2010. “The anatomy of growth in the OECD since 1870.” *Journal of Monetary Economics* 57: 753-767.

Madsen, J., and J. Ang. 2011. “Can Second-generation Endogenous Growth Models Explain Productivity Trends and Knowledge

Production in Asian Economies.” *The Review of Economics and Statistics* 93: 1360-1373.

Mandel, M., and D. Carew. 2013. *Regulatory Improvement Commission: A Politically-Viable Approach to U.S. Regulatory Reform.* Washington, DC: Progressive Policy Institute.

McLaughlin, P., and O. Sherouse. 2015. “Industry-Specific Classification of Legal Texts.” *Working Paper.*

Morgenstern, R., W. Pizer, and J. Shih. 2002. “Jobs Versus the Environment: An Industry-Level Perspective.” *Journal of Environmental Economics and Management* 43: 412-436.

Mulligan, C., and A. Shleifer. 2005. “The Extent of the Market and the Supply of Regulation.” *Quarterly Journal of Economics* 120: 1445-1473.

Nicoletti, G., and S. Scarpetta. n.d. “Regulation, productivity and growth: OECD evidence.” *Economic Policy* 18: 9-72.

Pedregosa, F., G. Varoquaux, V. Michel, B. Thirion, O. Grisel, and M. Blondel. 2011. “Scikit-Learn: Machine Learning in Python.” *Journal of Machine Learning Research* 12: 2825-2830.

Peltzman, S. 1975. “The Effects of Automobile Safety Regulation.” *Journal of Political Economy* 83: 677-725.

Peretto, P., and M. Connolly. 2007. “The Manhattan Metaphor.” *Journal of Economic Growth* 12: 329-350.

Peretto, Pietro F. 2007. “Corporate Taxes, Growth and Welfare in a Schumpeterian Economy.” *Journal of Economic Theory* 353-382.

Stigler, G. 1971. “The Theory of Economic Regulation.” *Bell Journal of Economics and Management Science* 2: 3-21.

Walker, W. R. 2013. “The Transitional Costs of Sectoral Reallocation: Evidence From the Clean Air Act and the Workforce.” *Quarterly Journal of Economics.*

Figure 1. Timepaths of both the Total Pages in the Code of Federal Regulations and The Total Count of Restrictions in RegData 2.2, 1977-2012

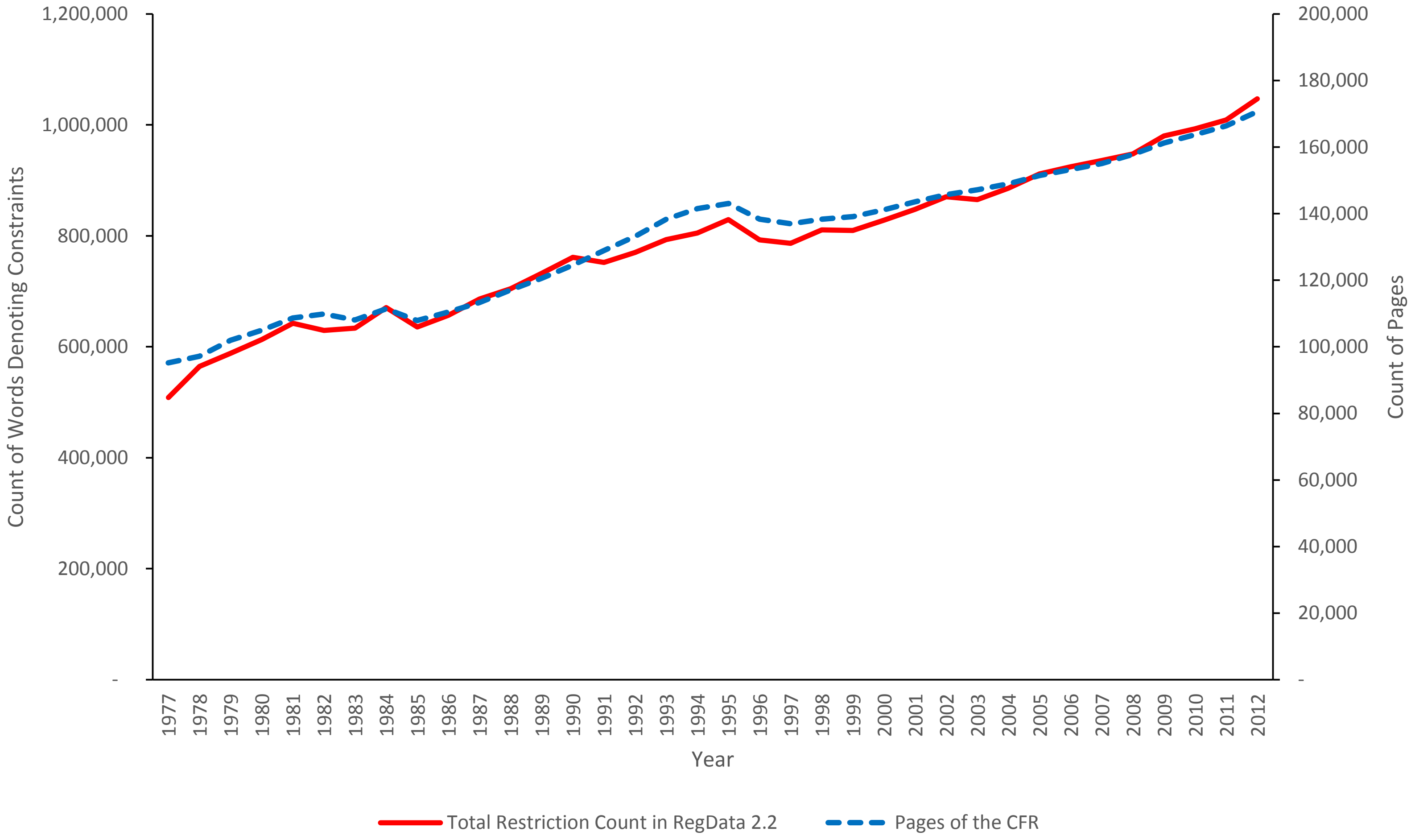

Figure 2. Timepath of Regulatory Constraints (in log scale) by Industry, 1977-2012

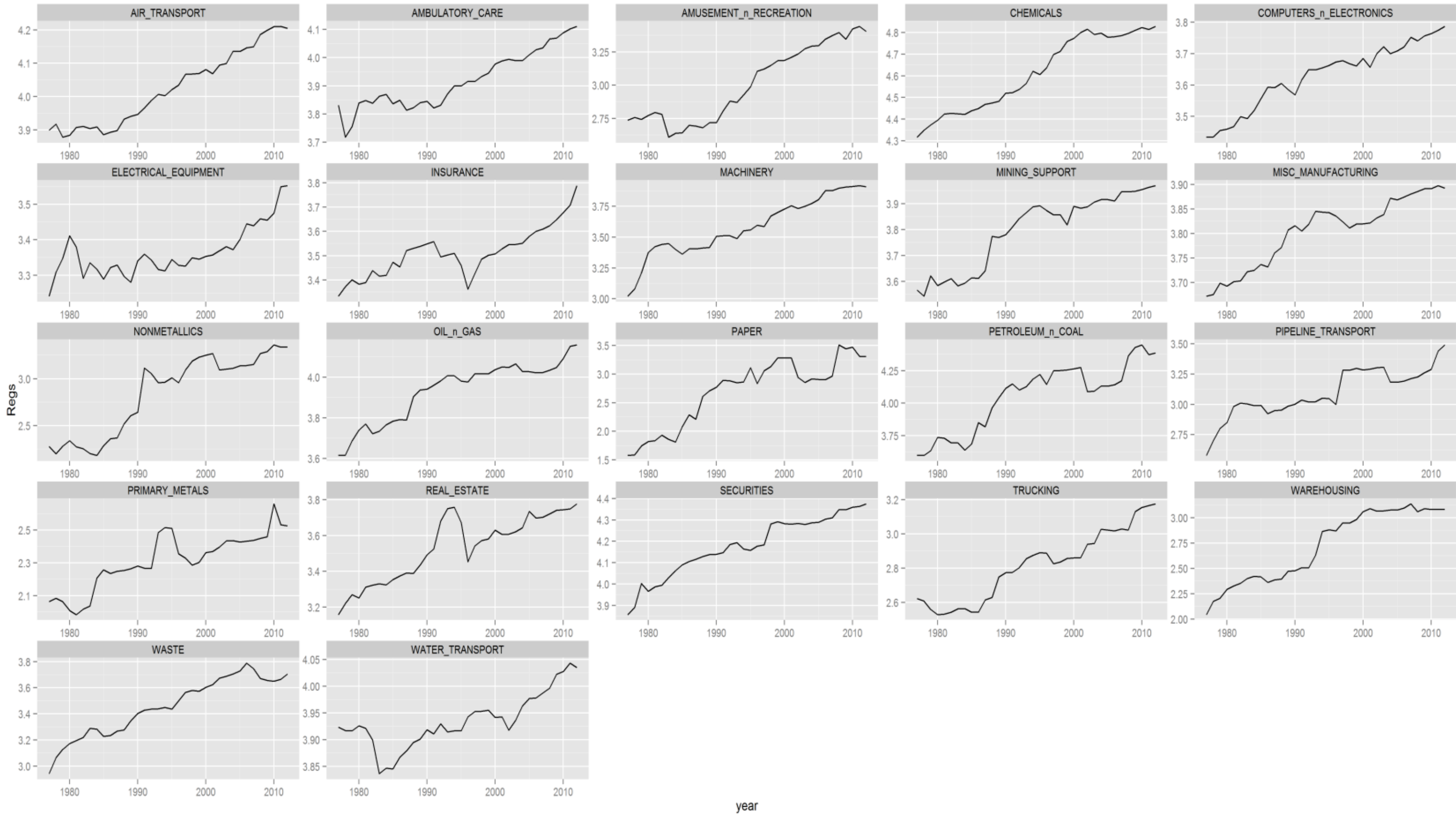

Figure 3. Goodness of Fit for Real Investment in Equipment using Bayesian Estimates (Blue) and OLS (Red) as a Basis for Comparison

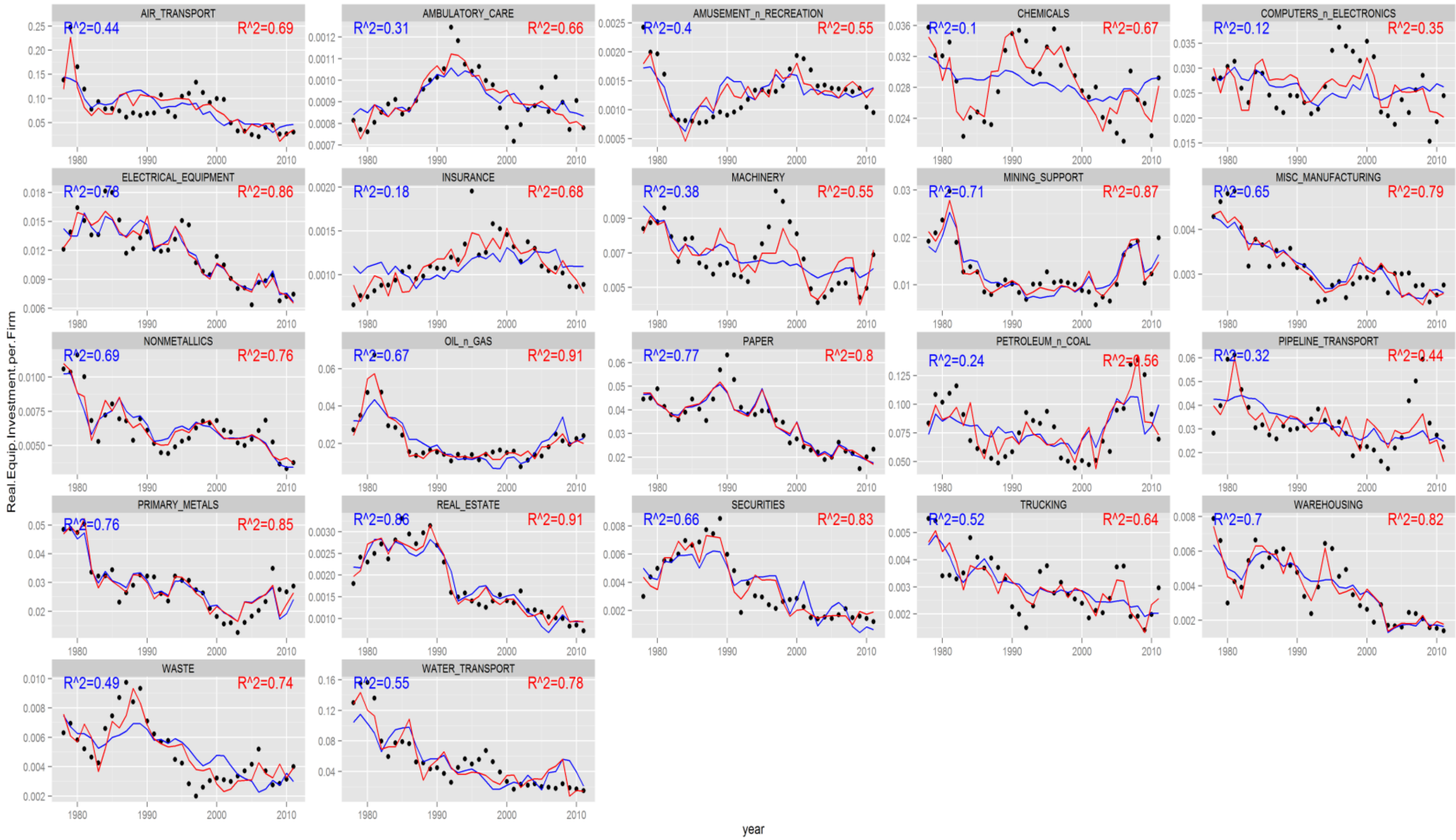

Figure 4. Goodness of Fit for Real Investment in Fixed Capital using Bayesian Estimates (Blue) and OLS (Red) as a Basis for Comparison

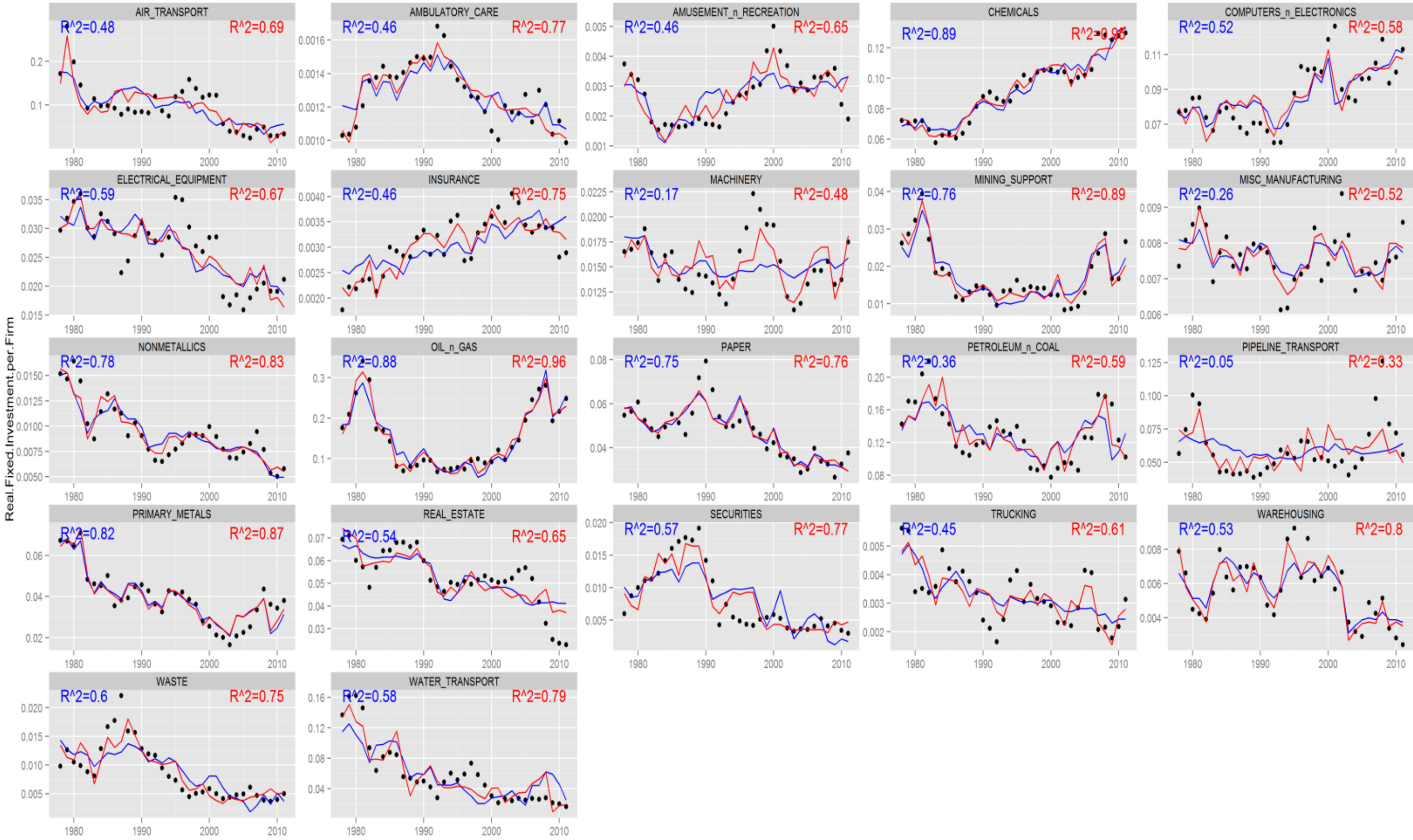

Figure 5. Goodness of Fit for Real Investment in Structures using Bayesian Estimates (Blue) and OLS (Red) as a Basis for Comparison

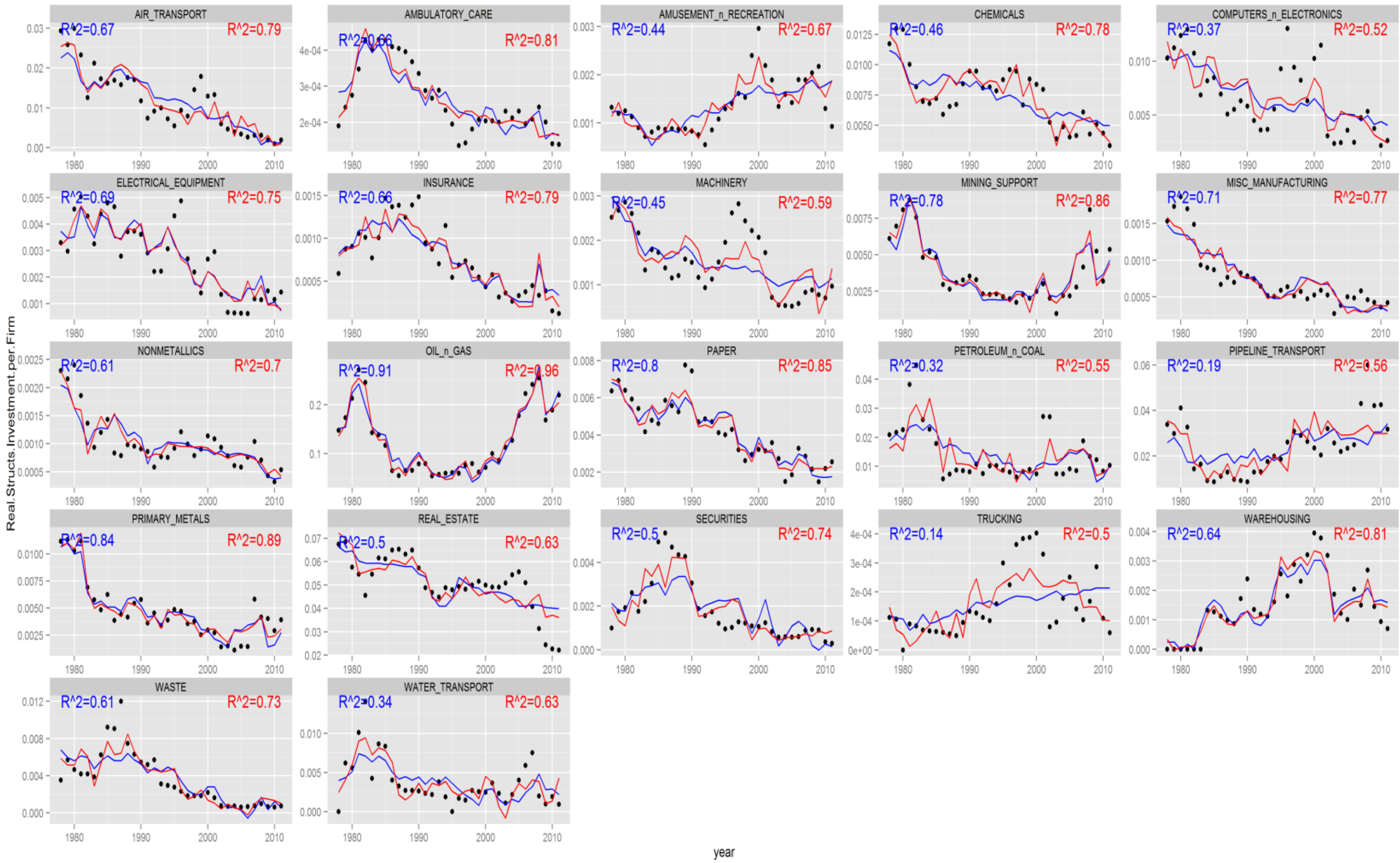

Figure 6. Goodness of Fit for Real Investment in IP using Bayesian Estimates (Blue) and OLS (Red) as a Basis for Comparison

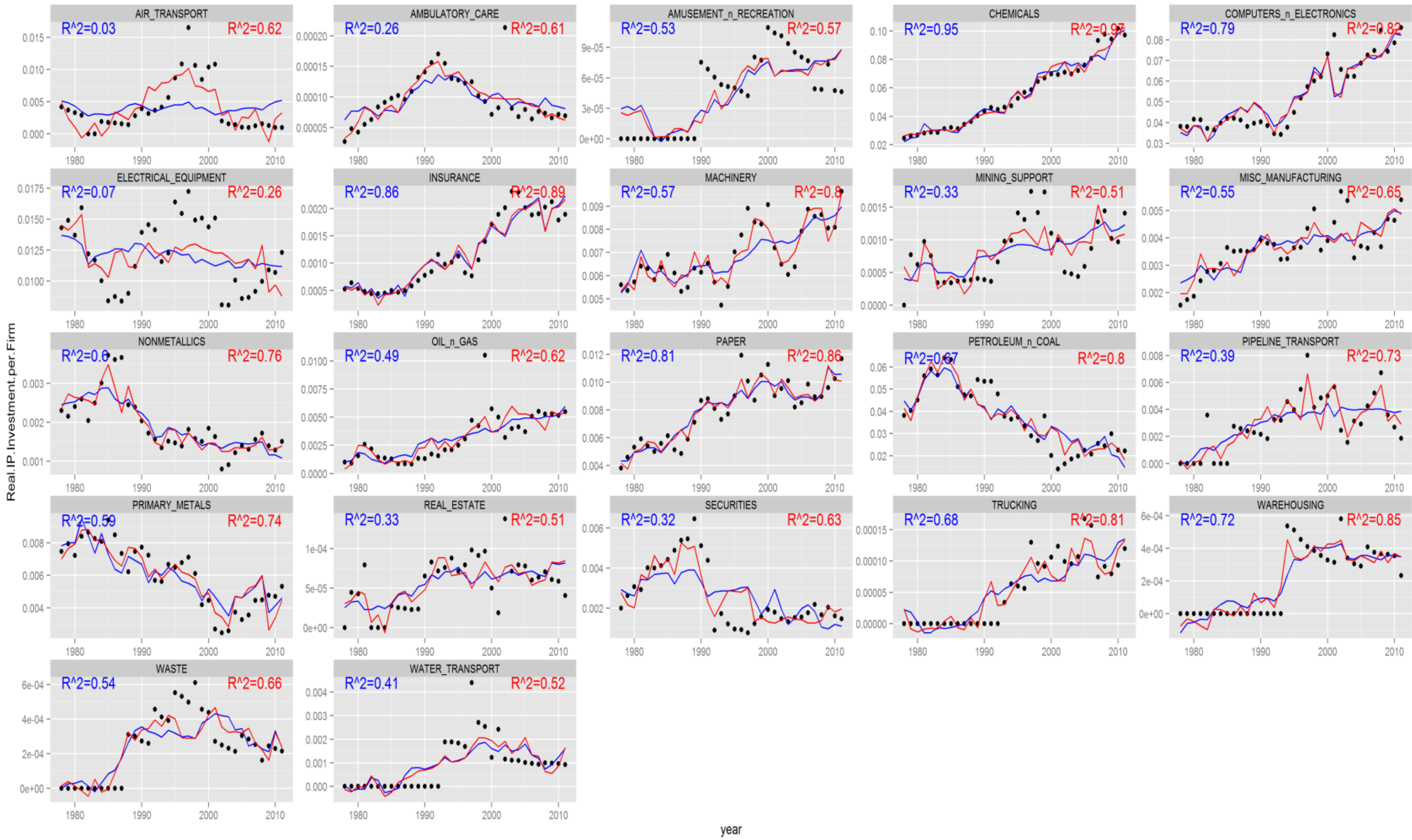

Figure 7. Goodness of Fit for Real Output per Establishment using Bayesian Estimates (Blue) and OLS (Red) as a Basis for Comparison

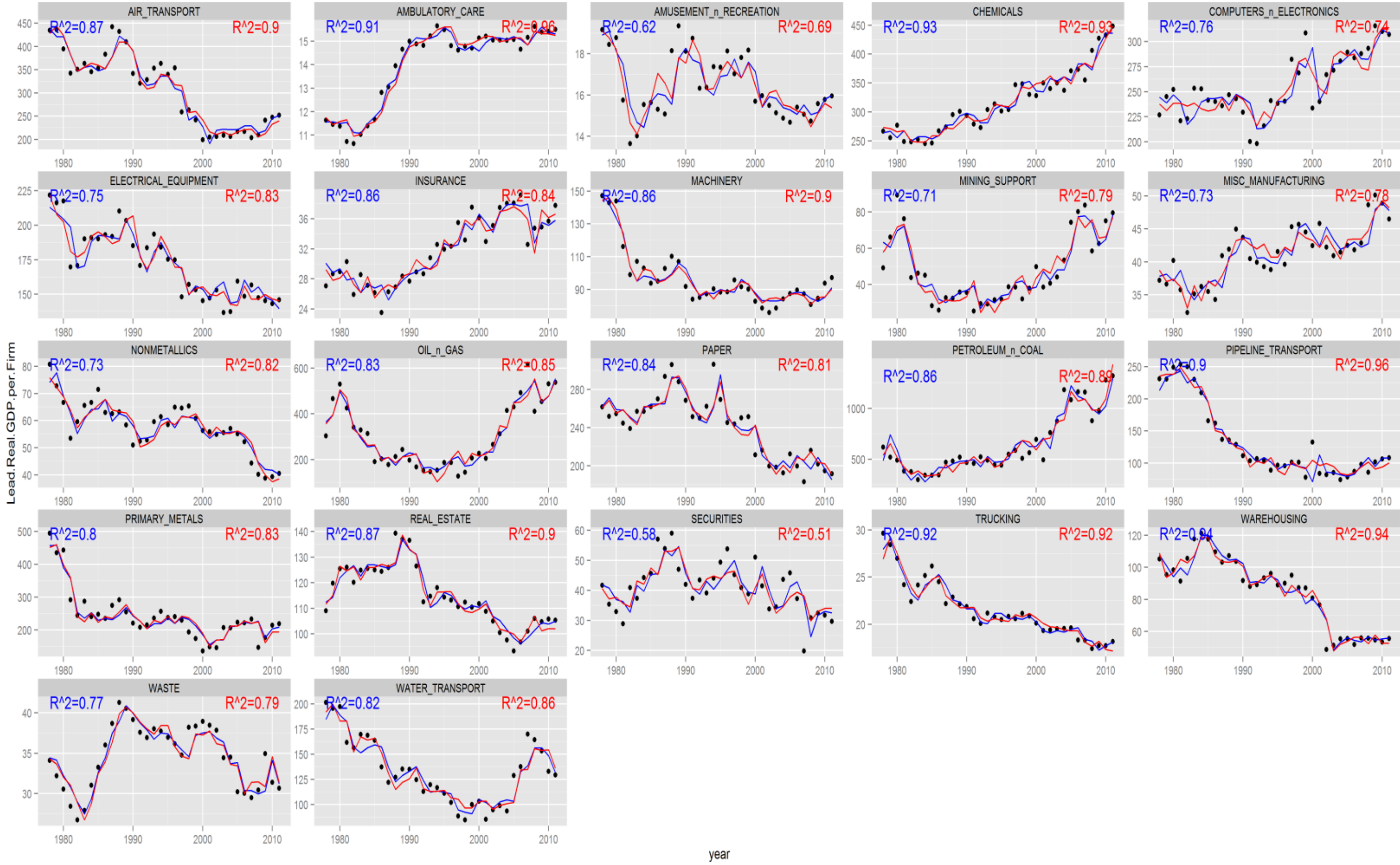

Figure 8. Goodness of Fit for Persons per Establishment using Bayesian Estimation (Blue) and OLS as a Basis for Comparison (Red)

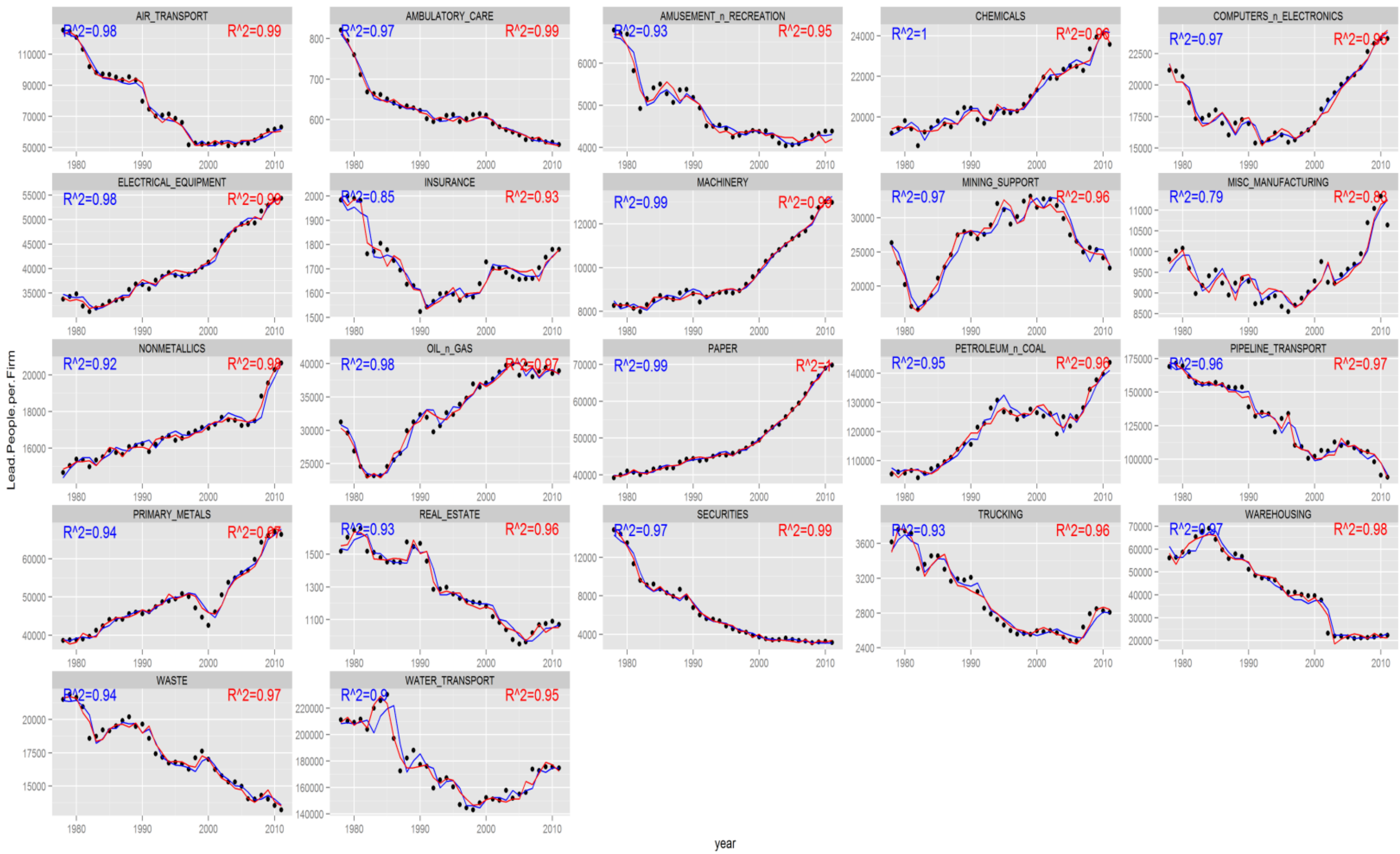

Figure 9. Factual (Black) and Counterfactual (Reds) Timepaths of Regulatory Constraints

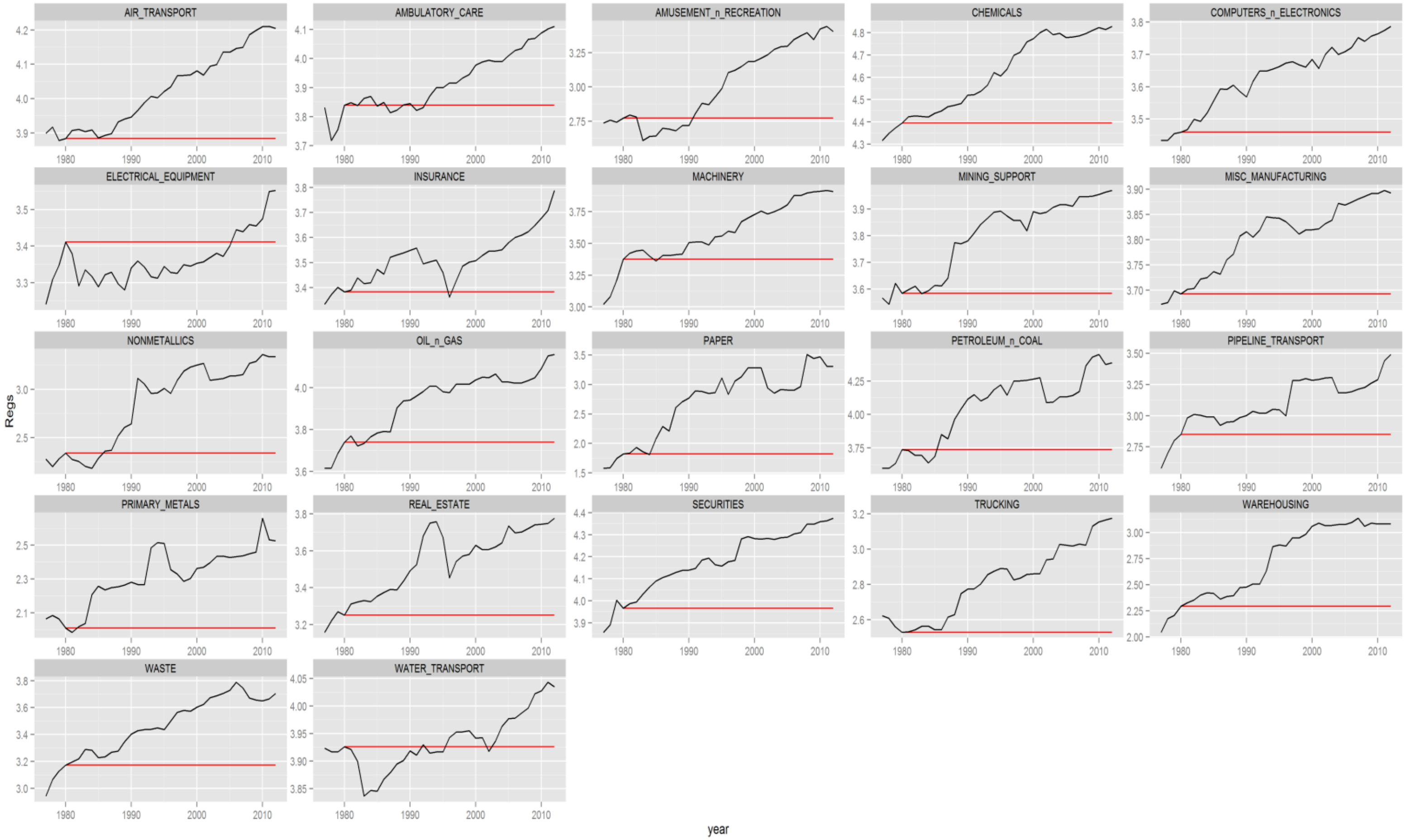

Figure 10. Factual (Black) and Counterfactual Real Investment in Equipment Predicted from Bayesian Estimates (Blue) and OLS (Red)

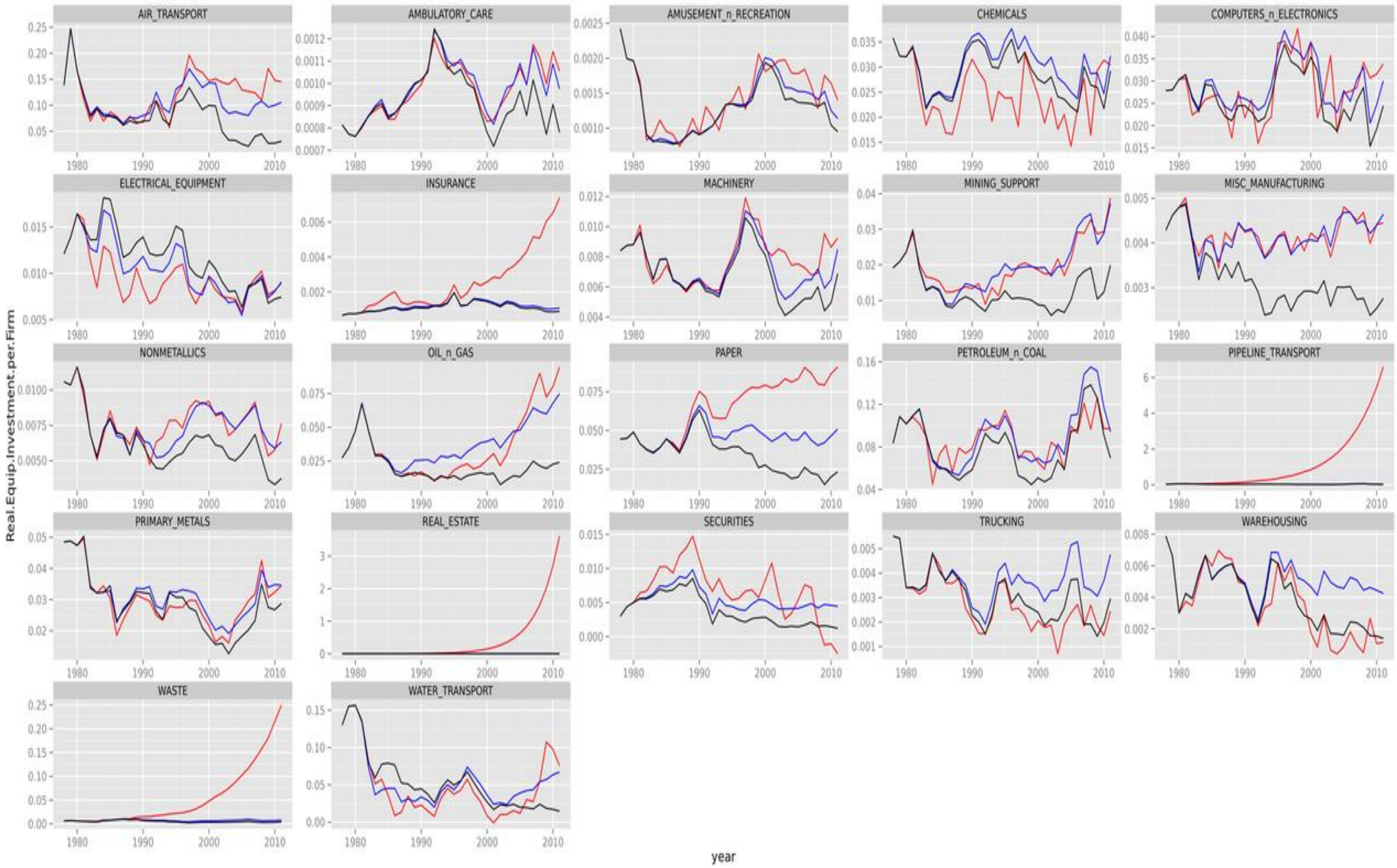

Figure 11. Factual (Black) and Counterfactual Real Investment in Fixed Capital Predicted from Bayesian Estimates (Blue) and OLS (Red)

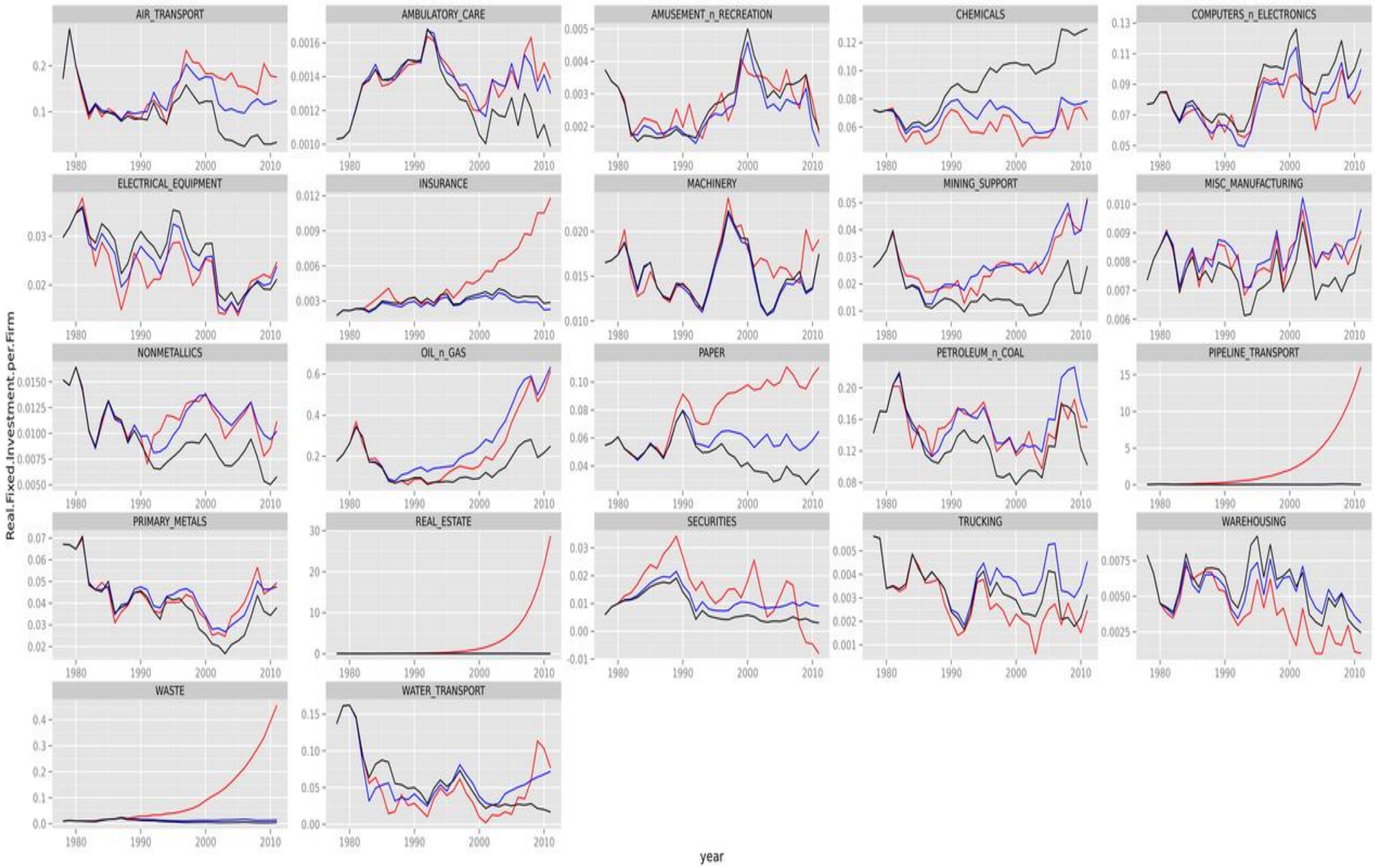

Figure 12. Factual (Black) and Counterfactual Real Investment in Structures Predicted from **Bayesian Estimates (Blue) and OLS (Red)**

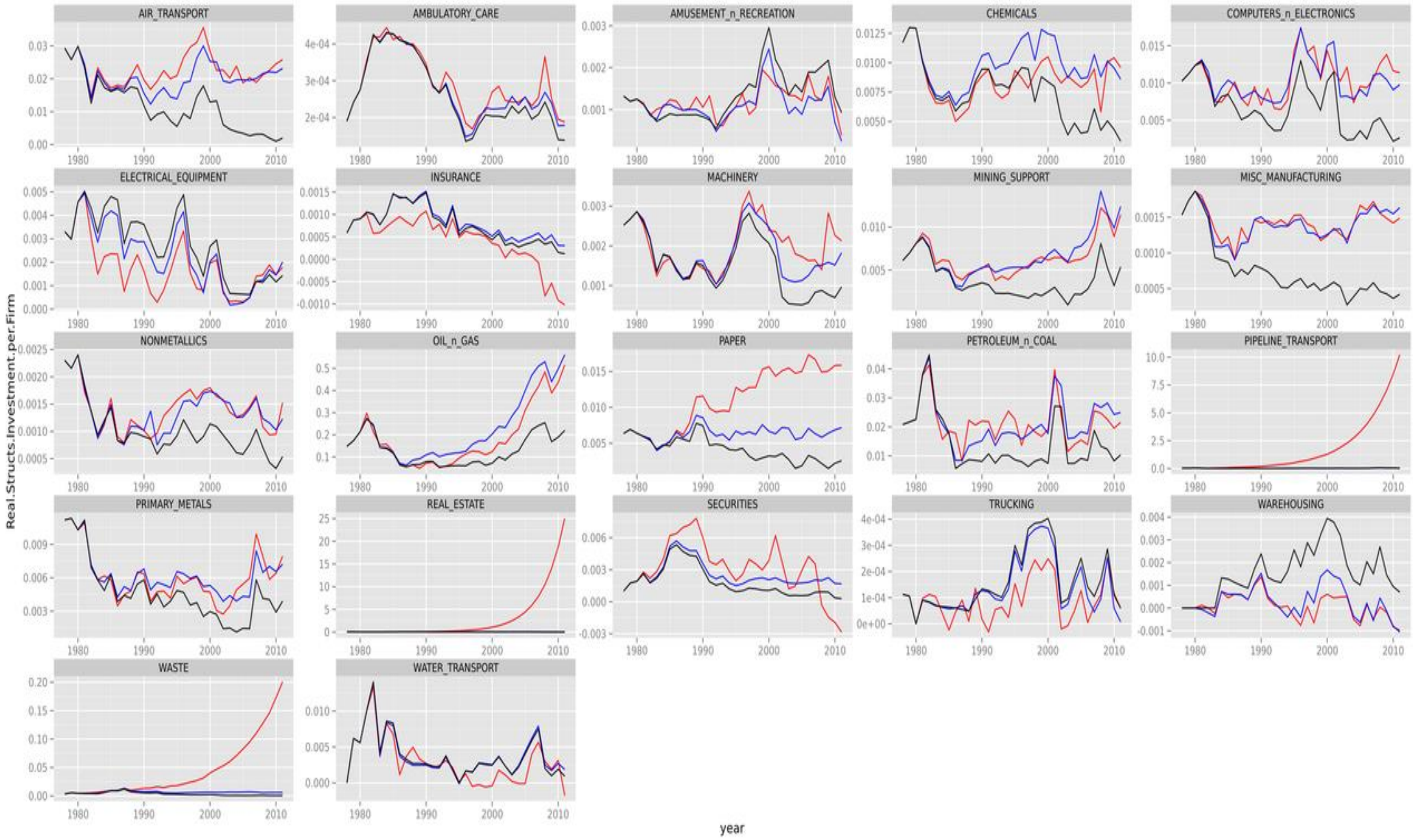

Figure 13. Factual (Black) and Counterfactual Real Investment in IP Predicted from **Bayesian Estimates (Blue) and OLS (Red)**

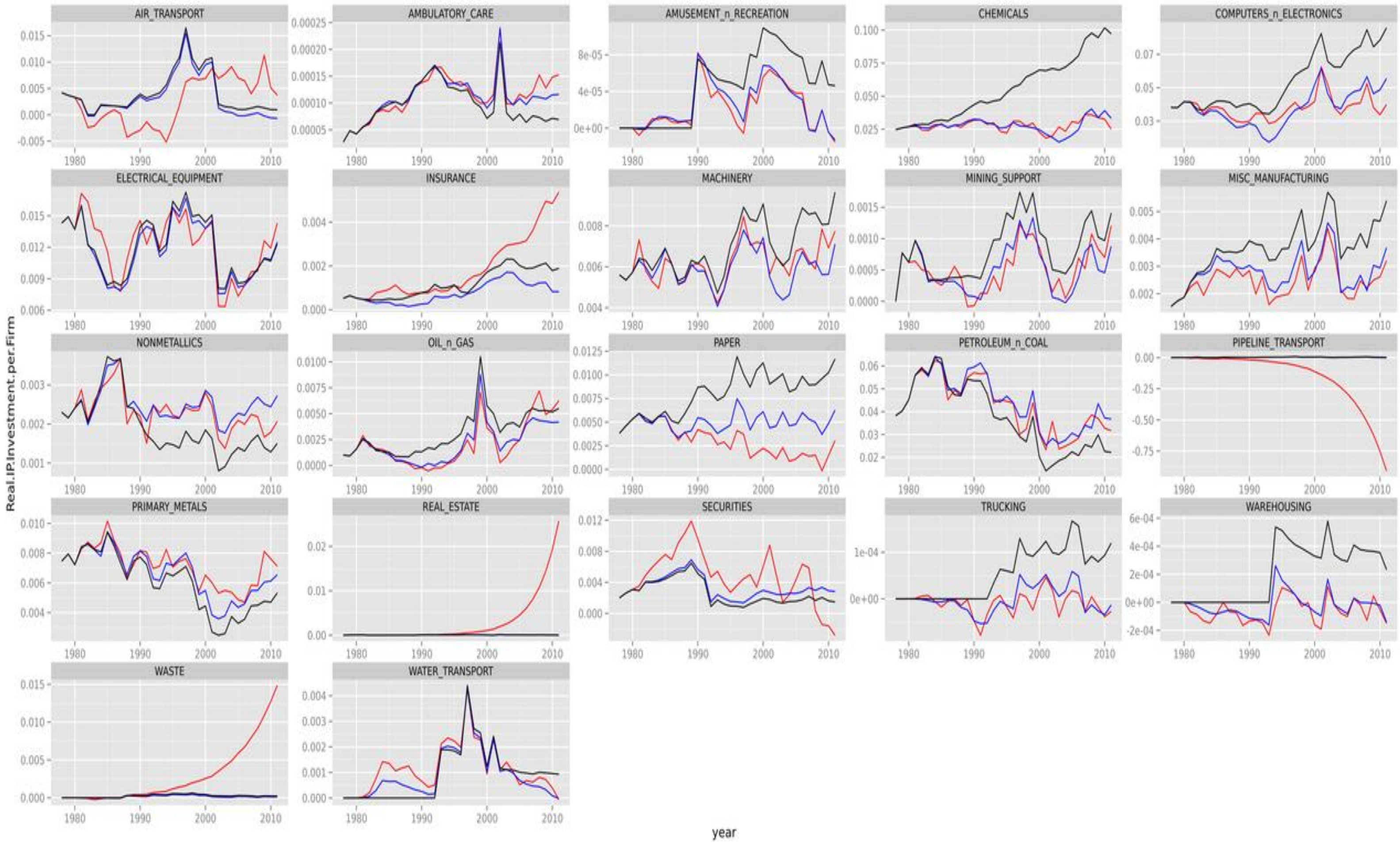

Figure 14. Factual (Black) and Counterfactual Real Output per Establishment Predicted from **Bayesian Estimates (Blue) and OLS (Red)**

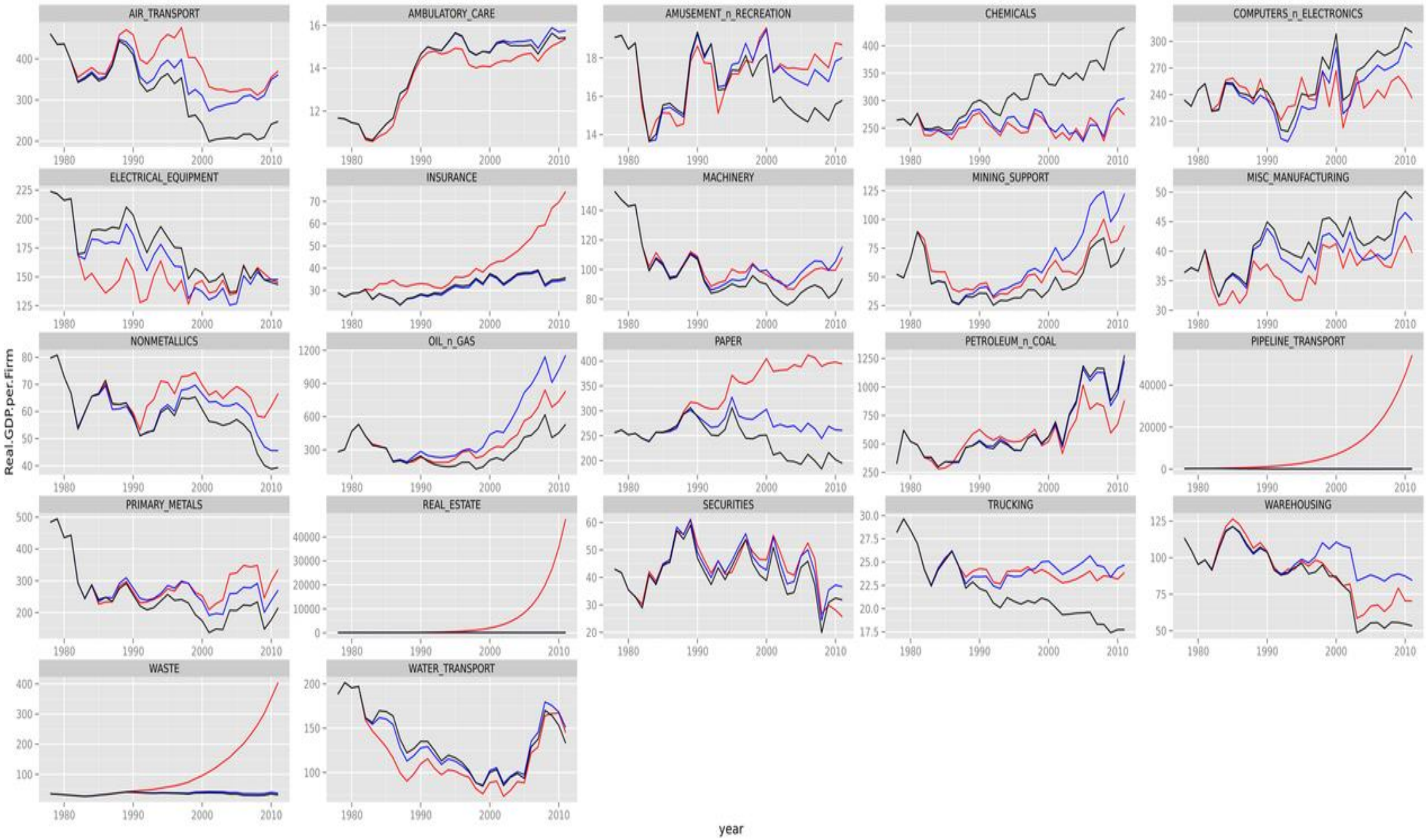

Figure 15. Factual (Black) and Counterfactual People per Establishment Predicted from **Bayesian Estimates (Blue) and OLS (Red)**

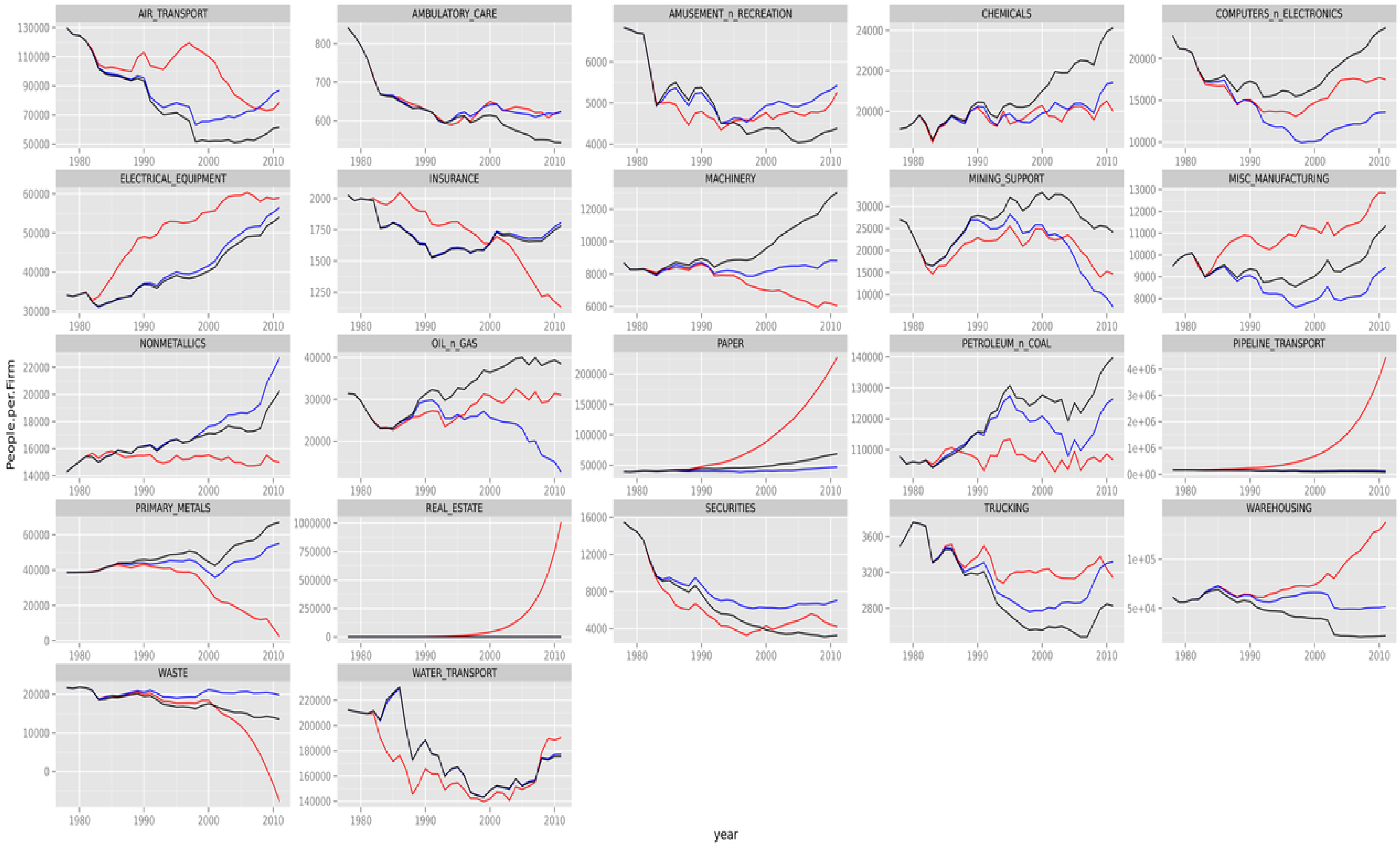

Figure 16. Factual (Black) and Counterfactual Value Added to GDP Peredicted from **Bayesian Estimates (Blue) and OLS (Red)**

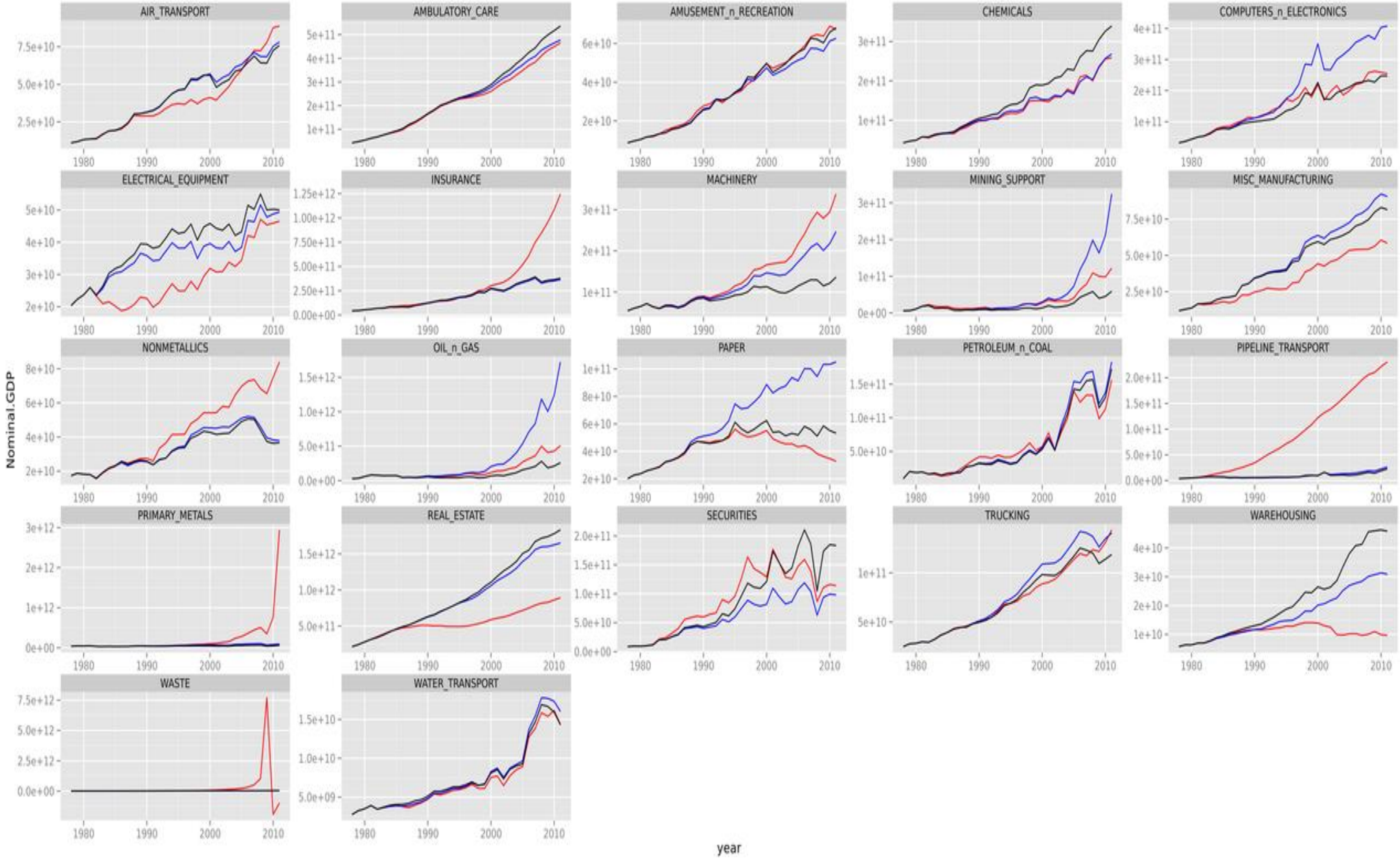

Figure 17. Factual (Black) and Counterfactual [Nominal] GDP Predicted from **Bayesian Estimates (Blue) and OLS (Red)**

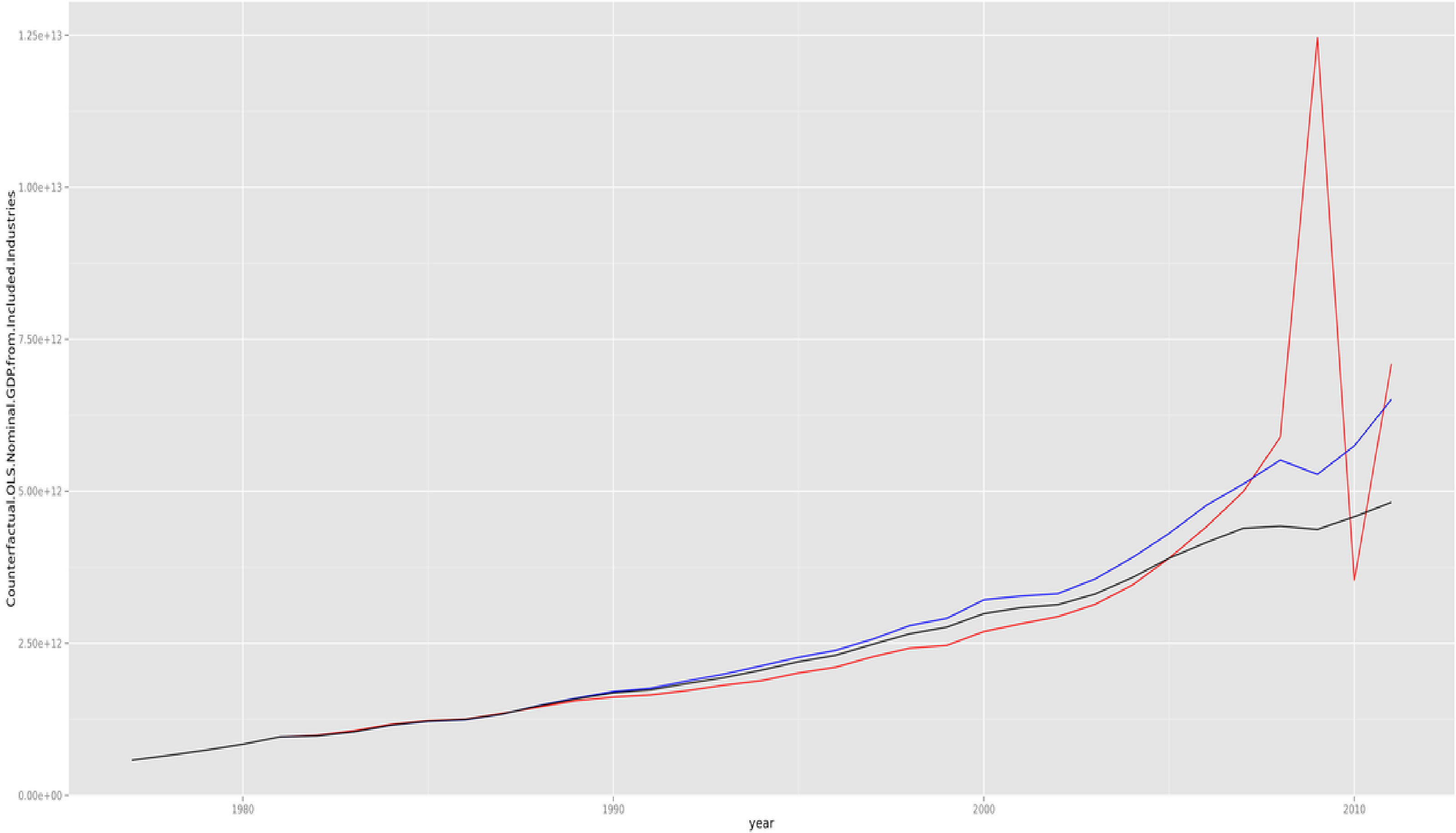

Table 1. Identifiers for 22 Industries Examined in our Study

NAICS Code	Full Description (2012 NAICS Code)	Abbreviated Label
211	OIL AND GAS EXTRACTION	OIL_n_GAS
213	SUPPORT ACTIVITIES FOR MINING	MINING_SUPPORT
322	PAPER PRODUCTS MANUFACTURING	PAPER
324	PETROLEUM AND COAL PRODUCTS MANUFACTURING	PETROLEUM_n_COAL
325	CHEMICAL PRODUCTS MANUFACTURING	CHEMICALS
327	NONMETALLIC MINERAL PRODUCTS MANUFACTURING	NONMETALLICS
331	PRIMARY METALS MANUFACTURING	PRIMARY_METALS
333	MACHINERY MANUFACTURING	MACHINERY
334	COMPUTER AND ELECTRONIC PRODUCTS MANUFACTURING	COMPUTERS_n_ELECTRONICS
335	ELECTRICAL EQUIPMENT, APPLIANCES, & COMPONENTS MANUFACTURING	ELECTRICAL_EQUIPMENT
339	MISCELLANEOUS MANUFACTURING	MISC_MANUFACTURING
481	AIR TRANSPORTATION	AIR_TRANSPORT
483	WATER TRANSPORTATION	WATER_TRANSPORT
484	TRUCK TRANSPORTATION	TRUCKING
486	PIPELINE TRANSPORTATION	PIPELINE_TRANSPORT
493	WAREHOUSING AND STORAGE	WAREHOUSING
523	SECURITIES, COMMODITY CONTRACTS, AND INVESTMENTS	SECURITIES
524	INSURANCE CARRIERS AND RELATED ACTIVITIES	INSURANCE
531	REAL ESTATE	REAL_ESTATE
562	WASTE MANAGEMENT AND REMEDIATION SERVICES	WASTE
621	AMBULATORY HEALTH CARE SERVICES	AMBULATORY CARE
713	AMUSEMENT, GAMBLING, AND RECREATION INDUSTRIES	AMUSEMENT_n_RECREATION

Table 2. Summary Statistics

Variable	N	T	Mean	Std Dev	Min	Max
Real_Investment_in_Equipment_per_Establishment	22	34	0.0203	0.0278	0.0007	0.2471
Real_Investment_in_Fixed_Capital_per_Establishment	22	34	0.0420	0.0502	0.0010	0.3420
Real_Investment_in_Structures_per_Establishment	22	34	0.0125	0.0315	0	0.2714
Real_Investment_in_Intellectual_Property_per_Establishment	22	34	0.0092	0.0178	0	0.1018
Real_Interest_Rate	1	34	0.0258	0.0225	-0.0075	0.0749
Real_Output-per_Establishment	22	34	151.9	162.3	10.6	1279.5
People_per_Establishment	22	34	38968.7	47742.7	543.6	230455.5
Log(Regulatory Constraints)	22	34	3.4928	0.6061	1.5867	4.8233
Population	1	34	266M	29M	220M	314M
GDP	1	34	8.3T	4.4T	2.1T	16.2T

Table 3. Bayesian Hyperparameter Estimates for Real Investment in Equipment

	mean	SE(mean)	Std dev	5th %-tile	50th %-tile	95th %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Intercept	-0.0001	0.0007	0.0222	-0.0352	-0.0002	0.0355	1050	1.0031
hetero_beta_Intercept	0.0195	0.0005	0.0147	0.0017	0.0161	0.0476	1050	0.9979
mean_beta_Int_Rate	-0.724	0.009	0.2909	-1.213	-0.7152	-0.2579	1050	0.996
hetero_beta_Int_Rate	0.1371	0.0028	0.0826	0.0141	0.1321	0.2808	859	0.9972
mean_beta_GDP_per_Firm	0.6589	0.0055	0.1768	0.3927	0.6538	0.9625	1050	0.9986
hetero_beta_GDP_per_Firm	0.1153	0.0023	0.0718	0.0103	0.1108	0.2403	969	0.9963
mean_beta_Regs	-0.2011	0.0042	0.1335	-0.4208	-0.1945	0.0112	1033	0.9986
hetero_beta_Regs	0.1281	0.0032	0.0872	0.0131	0.1156	0.2936	725	1.008
mean_beta_Regs × Int_Rate	0.3609	0.0296	0.943	-1.1587	0.3567	1.852	1012	0.9984
hetero_beta_Regs × Int_Rate	0.1123	0.0024	0.0756	0.0106	0.1037	0.2432	1030	0.9929
mean_beta_Regs × GDP_per_Firm	-0.0933	0.0111	0.3596	-0.6897	-0.0943	0.5144	1050	0.9947
hetero_beta_Regs × GDP_per_Firm	0.116	0.0025	0.0758	0.011	0.1056	0.2482	947	1.0093
mean_beta_Lagged_Regs	-0.1411	0.0044	0.1419	-0.3652	-0.1412	0.0929	1050	0.9971
hetero_beta_Lagged_Regs	0.2342	0.0034	0.0926	0.06	0.2429	0.3836	751	1.0016
mean_beta_Lagged_Regs × Int_Rate	0.4117	0.0304	0.946	-1.0766	0.4038	1.9166	965	1.0006
hetero_beta_Lagged_Regs × Int_Rate	0.1099	0.0022	0.0729	0.0118	0.0981	0.245	1050	1
mean_beta_Lagged_Regs × GDP_per_Firm	-0.0419	0.0113	0.362	-0.6385	-0.0426	0.5702	1020	0.9939
hetero_beta_Lagged_Regs × GDP_per_Firm	0.1233	0.0024	0.0772	0.0136	0.1153	0.2577	1009	1.0071

Samples drawn using NUTS: 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Table 4. Bayesian Hyperparameter Estimates for Real Investment in Fixed Capital

	mean	SE(mean)	Std Dev	5th %-tile	50th %-tile	95th %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Intercept	-0.0007	0.0007	0.0225	-0.0371	-0.0008	0.0359	953	0.9944
hetero_beta_Intercept	0.0187	0.0004	0.0144	0.0015	0.0156	0.0475	1050	0.9978
mean_beta_Int_Rate	-0.9108	0.0093	0.3	-1.4072	-0.9143	-0.4155	1050	1.0037
hetero_beta_Int_Rate	0.1539	0.003	0.0867	0.0178	0.1555	0.2976	854	1.0038
mean_beta_GDP_per_Firm	0.7388	0.0068	0.2214	0.3983	0.7304	1.1136	1050	0.9967
hetero_beta_GDP_per_Firm	0.1241	0.0025	0.0795	0.0151	0.1166	0.264	1050	0.9991
mean_beta_Regs	-0.1915	0.005	0.1633	-0.4625	-0.1971	0.0827	1050	0.9966
hetero_beta_Regs	0.3714	0.0038	0.108	0.1878	0.3743	0.5361	815	1.0047
mean_beta_Regs × Int_Rate	-0.0242	0.032	1.0366	-1.7369	-0.0379	1.6408	1050	0.9997
hetero_beta_Regs × Int_Rate	0.1218	0.0025	0.0789	0.0115	0.1139	0.2642	1001	0.9975
mean_beta_Regs × GDP_per_Firm	0.089	0.0128	0.4155	-0.603	0.0985	0.7494	1050	0.9975
hetero_beta_Regs × GDP_per_Firm	0.1638	0.0027	0.0859	0.0162	0.1669	0.302	1008	1.0002
mean_beta_Lagged_Regs	0.0505	0.0041	0.1336	-0.1669	0.0495	0.2687	1050	0.9984
hetero_beta_Lagged_Regs	0.1635	0.0043	0.1194	0.0123	0.1421	0.3836	780	1.0062
mean_beta_Lagged_Regs × Int_Rate	0.9685	0.0309	1.0022	-0.7506	0.9828	2.6958	1050	1.0001
hetero_beta_Lagged_Regs × Int_Rate	0.1222	0.0027	0.0806	0.0122	0.1112	0.2664	886	1.0096
mean_beta_Lagged_Regs × GDP_per_Firm	-0.2403	0.0117	0.38	-0.8635	-0.2505	0.3671	1050	0.9973
hetero_beta_Lagged_Regs × GDP_per_Firm	0.1207	0.0025	0.0774	0.0138	0.113	0.2597	928	1.0012

Samples drawn using NUTS: 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Table 5. Bayesian Hyperparameter Estimates for Real Investment in Structures

	mean	SE(mean)	Std Dev	5^{th} %-tile	50^{th} %-tile	95^{th} %-tile	$N_{Effective}$	$\hat{R}$
mean_beta_Intercept	0.0001	0.0006	0.0206	-0.0344	0.0004	0.0336	1050	0.9977
hetero_beta_Intercept	0.0189	0.0005	0.0154	0.0012	0.0156	0.0484	1050	1.0049
mean_beta_Int_Rate	-1.0429	0.0096	0.2907	-1.5097	-1.0478	-0.5684	907	1.0001
hetero_beta_Int_Rate	0.1717	0.0033	0.101	0.0235	0.1711	0.3486	922	1.0046
mean_beta_GDP_per_Firm	0.5688	0.0103	0.3268	0.0319	0.5553	1.1126	1006	1.0093
hetero_beta_GDP_per_Firm	0.3335	0.0042	0.128	0.1061	0.3381	0.5474	938	0.9988
mean_beta_Regs	-0.1192	0.0057	0.178	-0.3987	-0.1237	0.1728	989	0.9982
hetero_beta_Regs	0.4588	0.004	0.1215	0.2628	0.4522	0.6652	931	1.0067
mean_beta_Regs × Int_Rate	-0.2044	0.0325	1.0518	-1.9867	-0.121	1.4574	1050	0.9978
hetero_beta_Regs × Int_Rate	0.1359	0.003	0.0899	0.0114	0.1259	0.2935	928	1.0108
mean_beta_Regs × GDP_per_Firm	-0.4815	0.0143	0.4638	-1.2538	-0.4829	0.2946	1050	1.0007
hetero_beta_Regs × GDP_per_Firm	0.2012	0.0046	0.1345	0.0199	0.1844	0.4447	869	0.9997
mean_beta_Lagged_Regs	-0.0813	0.0045	0.1456	-0.3222	-0.0758	0.1539	1050	1.0022
hetero_beta_Lagged_Regs	0.2145	0.0047	0.1435	0.0195	0.2014	0.4715	928	1.0041
mean_beta_Lagged_Regs × Int_Rate	1.2844	0.0321	1.0269	-0.3168	1.2287	3.0626	1024	0.9983
hetero_beta_Lagged_Regs × Int_Rate	0.1336	0.0027	0.0865	0.0133	0.1261	0.2904	1015	1.0061
mean_beta_Lagged_Regs × GDP_per_Firm	0.1652	0.0129	0.4173	-0.4961	0.1579	0.8482	1050	1.0043
hetero_beta_Lagged_Regs × GDP_per_Firm	0.2583	0.0044	0.1394	0.035	0.2549	0.4999	1017	1.0022

Samples drawn using NUTS using 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Table 6. Bayesian Hyperparameter Estimates for Real Investment in Intellectual Property

	mean	SE(mean)	Std Dev	5^{th} %-tile	50^{th} %-tile	95^{th} %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Intercept	0.0008	0.0007	0.0213	-0.0351	0.001	0.036	1050	1.0012
hetero_beta_Intercept	0.0189	0.0004	0.0143	0.0018	0.0161	0.0454	1050	1.0058
mean_beta_Int_Rate	0.1737	0.0091	0.2943	-0.3301	0.1779	0.6295	1040	1.0019
hetero_beta_Int_Rate	0.1326	0.0024	0.0777	0.0164	0.1304	0.27	1017	0.9986
mean_beta_GDP_per_Firm	-0.2612	0.0086	0.2673	-0.7209	-0.2541	0.1638	967	0.9964
hetero_beta_GDP_per_Firm	0.2687	0.004	0.1295	0.0452	0.2739	0.4786	1050	1.0017
mean_beta_Regs	0.0341	0.0047	0.1469	-0.2027	0.0284	0.2788	995	1.0012
hetero_beta_Regs	0.1701	0.0034	0.1072	0.0156	0.1632	0.3547	969	0.9981
mean_beta_Regs × Int_Rate	0.7871	0.0314	1.0175	-0.8527	0.7501	2.4891	1050	1.0046
hetero_beta_Regs × Int_Rate	0.116	0.0023	0.0739	0.0093	0.1125	0.2459	1050	0.9995
mean_beta_Regs × GDP_per_Firm	0.0602	0.0126	0.4095	-0.612	0.0642	0.7562	1050	1.006
hetero_beta_Regs × GDP_per_Firm	0.1769	0.0038	0.1178	0.0161	0.1641	0.3852	960	1.003
mean_beta_Lagged_Regs	0.1538	0.0062	0.1864	-0.1434	0.143	0.4534	895	1.001
hetero_beta_Lagged_Regs	0.4228	0.0035	0.1087	0.2609	0.4194	0.6045	983	1.0049
mean_beta_Lagged_Regs × Int_Rate	-0.9235	0.0311	1.0067	-2.6252	-0.9126	0.6631	1050	1.0064
hetero_beta_Lagged_Regs × Int_Rate	0.1162	0.0023	0.0739	0.0119	0.1067	0.2455	1050	0.9933
mean_beta_Lagged_Regs × GDP_per_Firm	0.4139	0.0139	0.4495	-0.3391	0.4357	1.1505	1050	1.0083
hetero_beta_Lagged_Regs × GDP_per_Firm	0.2512	0.0042	0.1313	0.0358	0.2557	0.4527	964	1.0096

Samples drawn using NUTS using 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Table 7. Bayesian Hyperparameter Estimates for Real Output per Establishment

	mean	SE(mean)	Std dev	5^{th} %-tile	50^{th} %-tile	95^{th} %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Intercept	-0.0004	0.0004	0.014	-0.0239	-0.0001	0.0234	1050	1.0024
hetero_beta_Intercept	0.0122	0.0003	0.0098	0.001	0.0102	0.03	1050	1.0024
mean_beta_Dynamic_Variable	0.7963	0.0043	0.1371	0.5817	0.797	1.0117	1017	1.0033
hetero_beta_Dynamic_Variable	0.0613	0.0013	0.0424	0.0059	0.0563	0.1393	1020	1
mean_beta_Real_Equip_Investment_per_Firm	-0.2462	0.0111	0.3586	-0.8214	-0.2499	0.3531	1050	0.9987
hetero_beta_Real_Equip_Investment_per_Firm	0.0489	0.0012	0.0395	0.0038	0.0406	0.1238	1050	1.0007
mean_beta_Real_IP_Investment_per_Firm	0.17	0.0079	0.2556	-0.2619	0.1734	0.5767	1050	1.0007
hetero_beta_Real_IP_Investment_per_Firm	0.0598	0.0013	0.0435	0.0054	0.0513	0.1411	1050	0.993
mean_beta_Real_Structs_Investment_per_Firm	0.1371	0.0081	0.2636	-0.2993	0.1378	0.5789	1050	0.9985
hetero_beta_Real_Structs_Investment_per_Firm	0.0568	0.0013	0.0415	0.0043	0.0498	0.1356	1050	0.9991
mean_beta_Real_Fixed_Investment_per_Firm	0.3184	0.0113	0.3668	-0.2742	0.2896	0.9304	1050	1.0004
hetero_beta_Real_Fixed_Investment_per_Firm	0.0537	0.0012	0.0399	0.0054	0.0476	0.1312	1050	0.9977
mean_beta_Real_Interest_Rate × Real_GDP_per_Firm	-0.3157	0.0083	0.2695	-0.7696	-0.315	0.1195	1045	1.0003
hetero_beta_Real_Interest_Rate × Real_GDP_per_Firm	0.0786	0.0017	0.0524	0.0076	0.073	0.1752	1001	1.0061
mean_beta_Regs × Lagged_Dynamic_Variable	-0.4173	0.0131	0.4023	-1.0868	-0.4239	0.223	950	1.0019
hetero_beta_Regs × Lagged_Dynamic_Variable	0.0566	0.0013	0.0418	0.0041	0.0501	0.1293	1050	1.0001
mean_beta_Regs × Real_Equip_Investment_per_Firm	0.8133	0.0431	1.3615	-1.4382	0.8256	3.0803	998	0.9973
hetero_beta_Regs × Real_Equip_Investment_per_Firm	0.0465	0.0011	0.0362	0.0026	0.0396	0.1167	1050	0.9948
mean_beta_Regs × Real_IP_Investment_per_Firm	0.3605	0.0224	0.7245	-0.8125	0.3636	1.5227	1050	0.9988
hetero_beta_Regs × Real_IP_Investment_per_Firm	0.0533	0.0012	0.0401	0.0054	0.0454	0.1255	1050	0.997
mean_beta_Regs × Real_Structs_Investment_per_Firm	0.3674	0.031	1.0053	-1.3359	0.3527	2.0762	1050	0.9985
hetero_beta_Regs × Real_Structs_Investment_per_Firm	0.0509	0.0011	0.037	0.0045	0.0442	0.1168	1050	0.9962
mean_beta_Regs × Real_Fixed_Investment_per_Firm	-0.268	0.0547	1.7727	-3.2395	-0.2454	2.5801	1050	0.9966
hetero_beta_Regs × Real_Fixed_Investment_per_Firm	0.049	0.0012	0.0384	0.0044	0.0408	0.12	1050	1.0014

...continued on next page

...continued from previous page: Bayesian Hyperparameter Estimates for Real Output per Establishment								
	mean	SE(mean)	Std Dev	5th %-tile	50th %-tile	95th %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Regs	-0.0691	0.0029	0.093	-0.2226	-0.0708	0.0802	1050	0.9996
hetero_beta_Regs	0.0489	0.0012	0.035	0.0036	0.0433	0.1147	924	1.005
mean_beta_Regs × Real_Interest_Rate × Real_GDP_per_Firm	-0.2646	0.0283	0.9183	-1.7521	-0.2826	1.2873	1050	1.0019
hetero_beta_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.0924	0.0017	0.0562	0.0104	0.0907	0.1888	1050	0.9971
mean_beta_Lagged_Regs	0.0529	0.003	0.095	-0.1019	0.0552	0.2095	1024	1.0027
hetero_beta_Lagged_Regs	0.0508	0.0011	0.0356	0.0047	0.0452	0.1181	1050	0.9947
mean_beta_Lagged_Regs × Lagged_Dynamic_Variable	0.4094	0.0122	0.3894	-0.2327	0.4223	1.0404	1016	0.9963
hetero_beta_Lagged_Regs × Lagged_Dynamic_Variable	0.0593	0.0013	0.0416	0.0056	0.0535	0.133	1050	0.9988
mean_beta_Lagged_Regs × Real_Equip_Investment_per_Firm	-0.615	0.041	1.2977	-2.7923	-0.5983	1.5567	1003	0.998
hetero_beta_Lagged_Regs × Real_Equip_Investment_per_Firm	0.0487	0.0012	0.0379	0.0034	0.0397	0.1235	1050	1.0038
mean_beta_Lagged_Regs × Real_IP_Investment_per_Firm	-0.6094	0.0211	0.6837	-1.6719	-0.6255	0.4803	1050	1.0006
hetero_beta_Lagged_Regs × Real_IP_Investment_per_Firm	0.0539	0.0012	0.0389	0.0041	0.0471	0.1316	987	1.0011
mean_beta_Lagged_Regs × Real_Structs_Investment_per_Firm	-0.5221	0.0298	0.9669	-2.1225	-0.5224	1.0741	1050	0.9995
hetero_beta_Lagged_Regs × Real_Structs_Investment_per_Firm	0.0493	0.0012	0.0382	0.0033	0.0423	0.1212	1044	1.0009
mean_beta_Lagged_Regs × Real_Fixed_Investment_per_Firm	0.0676	0.0529	1.7141	-2.7063	0.0732	3.0241	1050	0.9962
hetero_beta_Lagged_Regs × Real_Fixed_Investment_per_Firm	0.0477	0.0012	0.0366	0.0038	0.0408	0.1136	976	0.9998
mean_beta_Lagged_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.4779	0.0284	0.9201	-1.0352	0.4853	1.9586	1050	0.9988
hetero_beta_Lagged_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.0894	0.0018	0.0572	0.0084	0.0845	0.1849	1020	0.996

Samples drawn using NUTS using 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Table 8. Bayesian Hyperparameter Estimates for Real Persons per Establishment

	mean	SE(mean)	Std Dev	5th %-tile	50th %-tile	95th %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Intercept	0.0004	0.0002	0.0064	-0.0102	0.0002	0.0106	1050	1.0002
hetero_beta_Intercept	0.0058	0.0001	0.0046	0.0005	0.0047	0.0151	1042	1.0023
mean_beta_Dynamic_Variable	0.9916	0.0018	0.0592	0.8926	0.9908	1.0871	1050	1.0009
hetero_beta_Dynamic_Variable	0.0233	0.0006	0.0183	0.0015	0.0193	0.0596	965	1.0026
mean_beta_Real_Equip_Investment_per_Firm	0.1482	0.0068	0.2217	-0.2407	0.1573	0.4922	1050	0.9991
hetero_beta_Real_Equip_Investment_per_Firm	0.0336	0.0007	0.0241	0.0031	0.029	0.0789	1050	0.9998
mean_beta_Real_IP_Investment_per_Firm	-0.031	0.0033	0.1055	-0.2044	-0.0307	0.1336	1050	1.0003
hetero_beta_Real_IP_Investment_per_Firm	0.0278	0.0006	0.0203	0.0022	0.0249	0.066	998	0.9978
mean_beta_Real_Structs_Investment_per_Firm	-0.2762	0.0038	0.1247	-0.4812	-0.2775	-0.0694	1050	0.998
hetero_beta_Real_Structs_Investment_per_Firm	0.0396	0.0008	0.0266	0.0042	0.0363	0.0872	1050	0.9942
mean_beta_Real_Fixed_Investment_per_Firm	0.1537	0.0072	0.2319	-0.2027	0.1427	0.542	1050	1.0016
hetero_beta_Real_Fixed_Investment_per_Firm	0.0394	0.0009	0.0283	0.0031	0.0358	0.0913	1014	1.0037
mean_beta_Real_Interest_Rate × Real_GDP_per_Firm	-0.1249	0.0022	0.0716	-0.2449	-0.1235	-0.0083	1050	1.0081
hetero_beta_Real_Interest_Rate × Real_GDP_per_Firm	0.0228	0.0006	0.0167	0.0019	0.02	0.0532	867	0.9992
mean_beta_Regs × Lagged_Dynamic_Variable	-0.1973	0.0054	0.1765	-0.4898	-0.1943	0.0907	1050	0.9986
hetero_beta_Regs × Lagged_Dynamic_Variable	0.0243	0.0006	0.0184	0.0022	0.0208	0.0594	1050	0.9999
mean_beta_Regs × Real_Equip_Investment_per_Firm	-0.5355	0.0241	0.7816	-1.892	-0.5365	0.7281	1050	0.9998
hetero_beta_Regs × Real_Equip_Investment_per_Firm	0.0274	0.0007	0.0215	0.0021	0.0232	0.0666	1050	1.0027
mean_beta_Regs × Real_IP_Investment_per_Firm	0.096	0.0103	0.3351	-0.438	0.0967	0.6285	1050	0.9986
hetero_beta_Regs × Real_IP_Investment_per_Firm	0.0284	0.0006	0.0209	0.0024	0.0252	0.0676	1050	0.992
mean_beta_Regs × Real_Structs_Investment_per_Firm	0.3545	0.0111	0.361	-0.2548	0.3625	0.9516	1050	0.9952
hetero_beta_Regs × Real_Structs_Investment_per_Firm	0.0341	0.0007	0.024	0.0029	0.0302	0.0794	1050	0.9991
mean_beta_Regs × Real_Fixed_Investment_per_Firm	0.3231	0.0283	0.9169	-1.2225	0.3149	1.8263	1050	0.9963
hetero_beta_Regs × Real_Fixed_Investment_per_Firm	0.0376	0.0009	0.0283	0.0029	0.032	0.0906	1050	0.9959

…continued on next page

…continued from previous page: Bayesian Hyperparameter Estimates for Real Persons per Establishment

	mean	SE(mean)	Std Dev	5^{th} %-tile	50^{th} %-tile	95^{th} %-tile	$N_{Effective}$	$\widehat{R}$
mean_beta_Regs	0.006	0.0018	0.0572	-0.0884	0.0051	0.1018	1050	1.0038
hetero_beta_Regs	0.0654	0.0011	0.0333	0.0115	0.0643	0.1237	969	1.0028
mean_beta_Regs × Real_Interest_Rate × Real_GDP_per_Firm	-0.0735	0.0083	0.2678	-0.4976	-0.0795	0.3835	1050	1.0026
hetero_beta_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.0243	0.0006	0.0178	0.002	0.0207	0.0577	966	1.002
mean_beta_Lagged_Regs	0.0333	0.0017	0.0544	-0.0504	0.0303	0.1288	1050	1.0018
hetero_beta_Lagged_Regs	0.0466	0.0009	0.0281	0.0055	0.0451	0.0955	945	0.9971
mean_beta_Lagged_Regs × Lagged_Dynamic_Variable	0.1102	0.0053	0.1706	-0.17	0.1042	0.404	1050	0.9988
hetero_beta_Lagged_Regs × Lagged_Dynamic_Variable	0.0218	0.0005	0.0166	0.0019	0.0185	0.0522	1050	1.0039
mean_beta_Lagged_Regs × Real_Equip_Investment_per_Firm	0.3339	0.0231	0.7483	-0.8456	0.3237	1.614	1050	1.0014
hetero_beta_Lagged_Regs × Real_Equip_Investment_per_Firm	0.029	0.0007	0.0223	0.0022	0.0249	0.0701	1050	0.9936
mean_beta_Lagged_Regs × Real_IP_Investment_per_Firm	-0.1076	0.0099	0.3201	-0.6206	-0.1098	0.4182	1050	0.9979
hetero_beta_Lagged_Regs × Real_IP_Investment_per_Firm	0.029	0.0007	0.0208	0.0022	0.0251	0.0669	992	1.0026
mean_beta_Lagged_Regs × Real_Structs_Investment_per_Firm	-0.1188	0.0104	0.3369	-0.6682	-0.1385	0.4579	1050	0.996
hetero_beta_Lagged_Regs × Real_Structs_Investment_per_Firm	0.0342	0.0008	0.025	0.0035	0.0286	0.0803	1020	1.0019
mean_beta_Lagged_Regs × Real_Fixed_Investment_per_Firm	-0.3946	0.0274	0.8874	-1.9358	-0.3969	1.0564	1050	0.9964
hetero_beta_Lagged_Regs × Real_Fixed_Investment_per_Firm	0.0373	0.0008	0.0262	0.0039	0.0325	0.0877	1050	1.0017
mean_beta_Lagged_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.1889	0.0085	0.2746	-0.2855	0.2	0.6323	1050	1.0007
hetero_beta_Lagged_Regs × Real_Interest_Rate × Real_GDP_per_Firm	0.0235	0.0006	0.0173	0.0016	0.0205	0.0545	889	0.9998

Samples drawn using NUTS using 15 chains, each with iter=1200; warmup=500; thin=10; post-warmup draws per chain=70, total post-warmup draws=1050

Appendix

The household saves according to the following Euler equation:

$$\frac{c}{c}+\frac{p_Y}{p_Y}= \ 1-\tau_A \ r-\rho \tag{16}$$

The individual's optimal trade-off of labor and leisure results in the following aggregate supply of labor:

$$L = Pl = P - \frac{1+\tau_C \ \lambda p_Y cP}{1-\tau_L \ W} \tag{17}$$

The representative firm for the final goods sector must have a unit cost equal to p_Y for there to be 0 profits:

$$\mathrm{p}_Y = \left(\sum_{j=1}^{J} \psi_j^{\xi} p_{Yj}^{1-\xi}\right)^{\frac{1}{1-\xi}} \tag{18}$$

Its conditional demands for the commodity outputs of industry j, Yj, are implicitly given by:

$$p_{Yj}Y_j = \psi_j^{\xi} p_{Yj}^{1-\xi} p_Y^{\xi-1} p_Y Y \tag{19}$$

Summing over all J industries gives:

$$\sum_{j=1}^{J} p_{Yj}Y_j = p_Y Y \tag{20}$$

To characterize the behavior of the typical intermediate goods firm, consider the following Current Value Hamiltonian:

$$\mathrm{H}_{ij} = \ 1-\tau_\pi \ \Pi_{ij} + q_{ij}\kappa K_j \mathrm{L}_{Zij} \tag{21}$$

where the costate variable, q_{ij}, is the value of the marginal unit of knowledge. With X_{ij} pinned down by market demand given $p_{X_{ij}}$, the control variables are the price of the intermediate good ($p_{X_{ij}}$) and the effort in knowledge accumulation ($L_{Z_{ij}}$). Note that the production and investment decisions are separable in the Hamiltonian. The first-order condition on $p_{X_{ij}}$ yields the intermediate goods firm's pricing strategy as a mark-up (depending on the price elasticity of demand) over marginal cost:

$$p_{Xij} = \left(\frac{\chi}{\chi - 1}\right) [Z_{ij}^{-\zeta} W] \tag{22}$$

Substituting into the firm's pre-tax flow of profits both this pricing strategy and the implicit equation for the conditional factor demand under within-industry symmetry across firms:

$$\Pi_{ij} = \left(\frac{1}{\chi}\right) \left[\frac{p_{Yj} Y_j}{N_j}\right] - W\phi_j - L_{Zij} W \tag{23}$$

Since the Hamiltonian is linear in L_{Zij}, its partial derivative yields three cases: 1) that the after-tax cost exceeds the value of the marginal unit of knowledge. The firm, then, does not invest; 2) that the after-tax cost is less than the value of the marginal unit of knowledge. Since the firm would then demand an infinite amount of labor to invest in knowledge accumulation, this case violates the general equilibrium conditions and is ruled out; 3) the first order conditions for the interior solution are given by equality between marginal revenue and marginal cost of knowledge:

$$\kappa K_j q_{ij} = \ 1 - \tau_\pi \ W \Longleftrightarrow L_{Zij} > 0 \tag{24}$$

This first-order condition is paired with the constraint on the state variable and a terminal condition:

$$\lim_{s \to \infty} e^{-\int_t^s [r \ \nu \ + \delta_j] d\nu} \mathrm{q}_{ij} \ s \ \mathrm{Z}_{ij} \ s \ = 0 \tag{25}$$

The partial derivative with respect to the state, the remaining part of the Hamiltonian's first-order conditions (which gets equated to costate times the discount rate less the time derivative), yields a differential equation in the costate variable that amounts to an arbitrage condition:

$$r + \delta_j = \frac{q_{ij}}{q_{ij}} + \frac{1 - \tau_\pi \dfrac{\partial \Pi_{ij}}{\partial Z_{ij}}}{q_{ij}} \tag{26}$$

that defines the (after-tax) rate of return to investment in knowledge as the ratio between additional profit from the knowledge and its shadow price plus (minus) the appreciation (depreciation) in the value of knowledge. Substituting the condition for an interior solution for investment (in knowledge accumulation) into the arbitrage equation, including in time derivatives, the dividend and corporate tax rates drop out of the equation governing optimal investment in quality improvement because investments in R&D get expensed:

$$r + \delta_j = \frac{\dot{W}}{W} - \frac{\dot{K}_j}{K_j} + \frac{\zeta W K_j^{\zeta - 1} p_{Xij} X_{ij} / p_{Xij}}{\left(W / \kappa K_j\right)} \tag{27}$$

We can further simplify this equation by substituting in the knowledge accumulation technology, the pricing strategy, and conditional demands under within-industry symmetry of identical intermediate goods firms:

$$r_{Zj} = \frac{\kappa}{W}\left[\zeta\left(\frac{\chi - 1}{\chi}\right)\frac{p_{Yj} Y_j}{N_j} - \frac{L_{Zj} W}{N_j}\right] + \frac{W}{W} - \delta_j \tag{28}$$

Entrepreneurs create new entrants in industry j by making an investment that requires their labor:

$$\mathrm{I}_{Nj} = W \mathrm{L}_{Nj} \tag{29}$$

Entrepreneurs target only new product lines because entering an existing product line in Bertrand competition with the existing supplier leads to losses. With the non-rival, non-excludable nature of knowledge, any new entrant starts in a state that is identical to the industry's incumbents. Because the new entrant solves a problem that is identical to an incumbent firm, the present value of profits from entry is V_{ij}. Entrepreneurs compare that present value of profits from introducing a new good to the entry cost. If the entry cost is higher then there is no new entry. If the entry cost is less, then entrepreneurs would demand an infinite quantity of resources which would violate resource constraints. Therefore, any entry indicates that the value of firms equals the entry cost, given in the following equation:

$$V_{ij} = W\theta \left[\frac{p_{Yj}Y_j}{N_j}\right] \Longleftrightarrow N_j > 0 \tag{30}$$

Given this expression for the cost of entry, we can divide the investment in new entry by this entry cost to obtain the differential equation that describes growth of the number of firms as a function of the labor expended on forming new entrants given the scale of the industry and the firm failure hazard rate:

$$\frac{N_j}{N_j} = \frac{\mathrm{L}_{Nj}}{\theta[p_{Yj}Y_j]} - \delta_j \tag{31}$$

The rate of return on an investment in a new entrant to industry j is governed by an equation easily derived from the logs and time derivatives for the value of incumbent firms:

$$r + \delta_j = \frac{V_{ij}}{V_{ij}} + \frac{1 - \tau_\pi \;\; \Pi_{ij}}{V_{ij}} \tag{32}$$

Analogous to the other equation describing the rate of return on an investment in capital accumulation, this equation states that the rate of return to the entrepreneurial activity of creating new firms is the capital gain/loss (i.e. the growth rate of the market price of the firm, V_{ij}) and the dividend-price ratio [i.e. the ratio of after-tax profits, $(1 - \tau_\pi)\Pi_{ij}$, to the market price of the firm, V_{ij}]. Substituting in for V_{ij} and its time derivative using our equation for entry costs, as well as profits, we obtain:

$$r_{Nj} = \frac{1 - \tau_\pi \;\; N_j}{\theta W [p_{Yj}Y_j]} \left(\frac{[p_{Yj}Y_j]}{\chi N_j} - W\phi_j - \frac{L_{Zj}W}{N_j}\right) + \frac{W}{W} + \frac{[p_{Yj}Y_j]}{[p_{Yj}Y_j]} - \frac{N_j}{N_j} - \delta_j \tag{33}$$

We equate the quantity of assets held by individuals to the value of intermediate goods firms and substitute in our expression for the value of firms to obtain:

$$Pa = \sum_{j=1}^{J} N_j V_{ij} = \sum_{j=1}^{J} N_j \theta W \left[\frac{p_{Yj} Y_j}{N_j}\right] = \theta W \sum_{j=1}^{J} p_{Yj} Y_j = \theta W p_Y Y \tag{34}$$

Taking logs and time derivatives:

$$\frac{\dot{a}}{a} = \frac{\dot{Y}}{Y} + \frac{\dot{p}_Y}{p_Y} + \frac{\dot{W}}{W} - \frac{\dot{P}}{P} \tag{35}$$

Dividing both sides of the household's budget constraint by an individual's assets and substituting in the previous two equations yields:

$$\frac{a}{a} = \frac{Y}{Y} + \frac{p_Y}{p_Y} + \frac{W}{W} - \underbrace{\frac{P}{P}}_{\omega} = [r \; 1 - \tau_A \; - \omega] + \frac{1 - \tau_L \; Wl}{\theta W p_Y Y/P} - \frac{1 + \tau_C \; p_Y c}{\theta W p_Y Y/P} \tag{36}$$

To solve for the interest rate in the model, we turn to the Euler equation which describes the after-tax rate of return required by households to equal their subjective rate of time preference plus the growth in the individual's expenditure on consumption:

$$r_A = \left(\frac{1}{1 - \tau_A}\right)\left[\frac{c}{c} + \frac{p_Y}{p_Y} + \rho\right] = \left(\frac{1}{1 - \tau_A}\right)\left[\frac{C}{C} - \frac{P}{P} + \frac{p_Y}{p_Y} + \rho\right] \tag{37}$$

Where C is aggregate consumption. Aggregate consumption can be eliminated using the goods market clearing condition (which takes the familiar form except without any investment term because investment in this model, i.e. $L_{Zj} + I_{Nj}$, is denominated in units of labor instead of final goods):

$$Y = C + [G_I + G_C - M] \tag{38}$$

If the share of final goods (net of net imports) controlled by the government is constant, than then this reduces to a simple solution for aggregation consumption equaling a fraction of can be re-written as:

$$C = \ 1 - g \ Y \tag{39}$$

Where g is the government's share of the final goods. This immediately implies that the growth rate of aggregate consumption equals the growth rate of output. Hence the Euler equation becomes:

$$r_A = \left(\frac{1}{1 - \tau_A}\right)\left[\frac{Y}{Y} + \frac{p_Y}{p_Y} + \rho - \omega\right] \tag{40}$$

In equilibrium, the same interest rate holds throughout the economy: $r = r_A = r_{N_j} = r_{Z_j}$ $\forall$j. Hence, we can substitute this into the households budget constraint (without assets already solved out) to obtain:

$$\frac{W}{W} = [\rho - \omega] + \frac{1 - \tau_L \ L}{\theta p_Y Y} - \frac{1 + \tau_C \ \ 1 - g}{\theta W} \tag{41}$$

Substituting the individual's labor supply into L provides an equation relating the value of final goods output to the wage rate:

$$\frac{W}{W} = [\rho - \omega] + \frac{1 - \tau_L \ P}{\theta p_Y Y} - \frac{1 + \tau_C \ \ 1 - g \ \ 1 + \lambda}{\theta W} \tag{42}$$

Because this is a general equilibrium model where only relative prices matter, we can normalize any single price to a positive constant. The obvious choices are the prices on final goods (p_Y) and time sold as labor (W). The former is more useful for our empirical analysis, while the latter is more useful for solving the model. Therefore, we set the numeraire to the time sold as labor in order to solve the model, after which we can renormalize final goods to be the numeraire for conducting our empirical analysis. The most convenient value with which to set the numeraire's price isn't actually 1 but rather:

$$W^* = \frac{1 + \tau_C \ \ 1 - g \ \ 1 + \lambda}{\theta \ \rho - \omega \ + \ 1 - \tau_L} \tag{43}$$

Substituting this into the previous equation reveals that, in these units, the value of output per capita equals 1:

$$\frac{p_Y Y}{P} = 1 \tag{44}$$

Hence, this measure of the value of output grows at the same rate as the population, which reduces the equilibrium interest rate to a constant:

$$r = \left(\frac{1}{1-\tau_A}\right)\rho \tag{45}$$

Substituting this expression into the rate of return to knowledge accumulation and using the technology for knowledge accumulation, we can solve for the growth rate of knowledge (given positive investment in knowledge accumulation) as a function of the value of the industry's output and the number of firms in the industry:

$$\frac{K_j}{K_j} = \frac{\kappa L_{Zj}}{N_j} = \left[\frac{\kappa\zeta \;\; \chi-1}{\chi W^*}\right]\frac{p_{Yj}Y_j}{N_j} - \frac{\delta_j \;\; 1-\tau_A \;\; +\rho}{1-\tau_A} \tag{46}$$

At this point, we invoke an additional assumption to diagonalize the system of differential equations that describe the model's growth dynamics. This effectively restricts the solution to the set of balanced growth paths because, otherwise, one particular industry could grow to a scale that dwarfs all others. By assuming that ψ=1, which is equivalent to a Cobb-Douglas technology in the aggregation of industry commodity outputs into final goods, the value of industry j's output is just a constant fraction of the value of final goods:

$$p_{Yj}Y_j = \psi_j p_Y Y = \psi_j P \tag{47}$$

Given this simplifying assumption, it is most convenient to work in the number of firms per capita (nj) as the state variable:

$$n_j = \frac{N_j}{P} \tag{48}$$

Using this simplifying assumption and state variable, as well as accounting for the non-negativity constraint in the growth of knowledge, we can write the growth rate of industry j's stock of knowledge as a function of the number of firms per capita in industry j:

$$\frac{K_j}{K_j} = \begin{cases} \left[\frac{\psi_j \kappa \zeta \;\; \chi - 1}{\chi W^*}\right]\left(\frac{1}{n_j}\right) - \left[\frac{\delta_j \;\; 1 - \tau_A \;\; + \rho}{1 - \tau_A}\right] & n_j < n_j \\ 0 & n_j \geq n_j \end{cases} \tag{49}$$

Where n_j is the critical mass of firms per capita, above which firms do not invest in innovation:

$$n_j = \frac{\psi_j \kappa \zeta \;\; \chi - 1 \;\;\; 1 - \tau_A}{\delta_j \;\; 1 - \tau_A \;\; \chi W^* + \rho \chi W^*} \tag{50}$$

Substituting these terms into the rate of return to new entry and solving for the growth rate of n_j:

$$\frac{n_j}{n_j} = \gamma_{j0}(1\{n_j < n_j\}) + \gamma_{j1}(1\{n_j < n_j\}) n_j \tag{51}$$

Where:

$$\begin{aligned} \gamma_{j1}(1\{n_j < n_j\}) &= \frac{1 - \tau_\pi}{\theta \kappa \psi_j}\left[1\{n_j < n_j\}\left[\frac{\rho + \delta_j \;\; 1 - \tau_A}{1 - \tau_A}\right] - \kappa \phi_j\right] \\ \gamma_{j0}(1\{n_j < n_j\}) &= \left(\frac{1 - \tau_\pi}{\theta W^*}\right)\left[\frac{1}{\chi} - 1\{n_j < n_j\}\zeta\left(\frac{\chi - 1}{\chi}\right)\right] - \left[\frac{\rho + \delta_j \;\; 1 - \tau_A}{1 - \tau_A}\right] \end{aligned} \tag{52}$$

This piecewise ordinary differential equation is of the Bernoulli class with a closed-form solution following logistic growth:

$$n_j = \frac{\gamma_{j0}(n_j) n_j \;\; 0}{[\gamma_{j1}(n_j) n_j \;\; 0 \;\; + \gamma_{j0}(n_j)] e^{-\gamma_{j0} t} - \gamma_{j1}(n_j) n_j \;\; 0} \tag{53}$$

Which ultimately converges to the following steady-state number of firms per capita:

$$n_j^* = \frac{-\gamma_{j0}(n_j)}{\gamma_{j1}(n_j)} = \frac{\kappa\psi_j \left[\frac{1}{\chi} - 1\{n_j < n_j\}\zeta\left(\frac{\chi - 1}{\chi}\right)\right] - \left[\frac{\rho + \ 1 - \tau_A \ \delta_j}{1 - \tau_A}\right]}{W^* \left[1\{n_j < n_j\}\left[\frac{\rho + \delta_j \ 1 - \tau_A}{1 - \tau_A}\right] - \kappa\phi_j\right]} \tag{54}$$

Assuming $n_j^* < n_j$ as a regularity conditions on the parameters, which is consistent with us observing firms still investing in productivity enhancements when their industry is undergoing consolidation (i.e. reducing the number of firms per capita), the steady-state growth rate in knowledge accumulation is given by the following expression:

$$\frac{K_j}{K_j}^{*} = \left[\frac{\psi_j \kappa \zeta \ \chi - 1}{\chi W^*}\right]\left(\frac{1}{n_j^*}\right) - \left[\frac{\delta_j \ 1 - \tau_A \ + \rho}{1 - \tau_A}\right] \tag{55}$$

This solution utilizes an industry's number of firms per capita as the state variable in that industry's dynamics. Via a simple transformation, we can also recast the state variable as a scalar multiple of the per-firm revenue:

$$\left(\frac{1}{\psi_j}\right) \times \frac{p_{Yj} Y_j}{N_j} = \left(\frac{1}{\psi_j}\right) \times \frac{\psi_j p_Y Y}{N_j} = \frac{1}{N_j} \times \frac{p_Y Y}{P/P} = \frac{1}{N_j/P} \times \underbrace{\frac{p_Y Y}{P}}_{1} = \frac{1}{n_j} \tag{56}$$

Although this transformed state variable may be less intuitive for some, it further simplifies the dynamics from logistic growth to the exponential growth exhibited by a piecewise linear differential equation:

$$\left[\frac{p_{Yj} Y_j}{N_j}\right] = -\psi_j \times \gamma_{j1}(1\{n_j < n_j\}) - \gamma_{j0}(1\{n_j < n_j\})\left[\frac{p_{Yj} Y_j}{N_j}\right] \tag{57}$$

Regardless of our choice of state variable, the resulting solution is unchanged.