Technical Paper Series
Congressional Budget Office
Washington, DC

A PUTTY-CLAY MODEL OF BUSINESS FIXED INVESTMENT

Mark Lasky

Congressional Budget Office
Washington, DC
E-mail: mark.lasky@cbo.gov

September 2003

2003-9

Abstract

In this paper, I specify and estimate a model of investment using ex post fixed proportions, or putty-clay capital, for five categories of business fixed investment. The specification of expectations causes investment to depend on growth in output less growth in total factor productivity (TFP) and on labor productivity times growth in labor hours at full employment. Thus, demographics can play an important role in investment. TFP at full employment is estimated using a Kalman filter to determine the level that would satisfy an Okun's Law relationship exactly. Labor and aggregate capital are combined using a Cobb-Douglas production function, but the function aggregating different types of capital allows for elasticities of substitution greater than in a Cobb-Douglas function. The model explains past investment fairly well without the use of constants or lagged dependent variables. The collapse of investment between 2000 and 2003 can be attributed to a slowdown in growth of output relative to growth of productivity, a higher cost of funds caused by lower stock prices, and a slowdown in growth of labor hours at full employment.

1 Introduction

Weak business investment in new plant and equipment has played a major role in the economic slump of the past three years. In absolute terms, the decline in real business fixed investment between the third quarter of 2000 and the first quarter of 2003 was easily the largest ever. In relative terms, that decline subtracted more from GDP growth—1.6 percentage points—than any decline in business fixed investment since World War II except the steep drop during 1982-1983. The downturn of investment from 2000 to 2003 followed a boom in investment during the 1990s.

What factors have caused businesses to curtail investment by such a large magnitude since 2000? What factors caused the preceding boom in investment? What factors might trigger a recovery of investment over the next few years? These questions can be answered using an economic model of investment.

A basic assumption of this paper is that capital is putty-clay. That is, once a piece of equipment has been manufactured or a structure has been built, the ratio of capital to labor embodied in that asset remains the same as long as it is in use. Once a truck is built, for example, it requires exactly one driver at any one time it is being operated. Once a PC is built, only one person operates it at any one time. Over time, a given capital asset may be used by different workers, or even different firms, but the number of workers required or able to operate it at any one time remains fixed. The putty-clay assumption, also known as ex post fixed proportions, has a long history in the economics literature, going back to Johansen (1959). In the short run, that assumption implies that a one percent reduction in the cost of each type of capital boosts investment by one percent but allows a one percent rise in output to boost investment by more than one percent.

I assume that each worker produces output according to a Cobb-Douglas function of labor and a composite of the amount of each type of capital used by that worker. However, that composite is not itself a Cobb-Douglas function of those different types of capital. Instead, the elasticity of substitution between different types of capital is greater (more negative) than -1. That feature allows the share of GDP going to investment in information technology to trend upward over time even though capital's share of GDP is roughly stable over time.

The combination of time to build and of businesses' expectation that the gap between actual output and output at full employment will narrow over time means that both growth in actual output and growth in output at full employment affect investment. However, higher productivity reduces the amount of capital needed to produce a given amount of output, so investment depends on growth in output less growth in cyclically-adjusted total factor productivity, rather than on growth in output alone. Under realistic assumptions, a one percent rise in productivity leads to a roughly one percent rise in investment but produces no change in unemployment. Growth in labor hours at full employment affects growth in output at full employment, and so affects investment.

The model specified in this paper is not completely consistent with the putty-clay assumption. I assume that replacement investment equals the value of capital depreciated, but in a true putty-clay model, replacement machines will have greater capacity than the depreciated machines they replace. In addition, the cost of capital is not fully consistent with the putty-clay assumption. While those assumptions greatly simplify the estimation, I hope to relax them in the future.

The model, estimated for five different asset classes covering most of business fixed investment, does a reasonably good job of explaining investment in those assets over the past 30 years. In particular, the model can explain most of the downturn in investment between 2000 and 2003 as a result of three factors: a downturn in output relative to cyclically-adjusted total factor productivity; a rise in the real cost of funds; and below-trend growth in full-employment labor hours that left less scope for a cyclical rebound than in previous recoveries.

Section 2 describes the production function in detail, including its background, specification, and some simplifying assumptions. Section 3 specifies a model of investment using that production function. Businesses choose the size of new machines (new capital per worker) that maximizes profits. Businesses assume that part of the existing gap between actual output and output at full employment will close during the time to build. Section 3 concludes with a discussion of the effects of demand shocks and productivity shocks on investment and a comparison of the putty-clay investment model with the standard neoclassical investment model.

Section 4 discusses the data used to estimate the model. Those data include estimated data for labor hours at full employment, the capital stock at full employment, total factor productivity at full employment, and fitted cost shares for different groups of assets. Section 5 discusses the empirical results for the investment equations and how they help explain the past behavior of investment.

2 The Production Function

The results of this paper ultimately stem from assumptions about how businesses produce output from labor and capital. The most important of those assumptions is that capital is putty-clay. Even so, the output of a given combination of labor and capital can change over time with technical progress. In addition, I assume that output is produced by a Cobb-Douglas function of labor and aggregate capital. However, aggregate capital is not a Cobb-Douglas function of individual types of capital, so the elasticity of substitution between different types of capital generally exceeds (is more negative than) -1.

To derive a tractable expression for the expected output of a new structure, piece of equipment, or software, I must make some simplifying assumptions. First, I assume that the same capital aggregating function applies to each worker, meaning that each worker uses a similar mix of capital. Second, I assume that all workers have the same labor quality. Finally, I assume that businesses reallocate capital among workers frequently and at no cost. While none of these assumptions hold in the real world, they are likely to have little impact on the results, while making those results much easier to obtain.

2.1 Basics of Production: Machines, Output, and Capacity

The mechanics of production are different in a world with putty-clay capital than in the standard neoclassical world. In the standard neoclassical production function, output is a function of total workers, or total labor hours, and of the total capital stock. That means that any investment affects the marginal productivity of all labor and of all existing capital. For example, consider the effects of a purchase of a billion dollars of new computers in an economy with 100 million workers. Those additional computers boost the marginal productivity of all workers; implicitly, each worker uses an extra 10 dollars of computer. At the same time, due to decreasing marginal returns, the marginal productivity of the pre-existing stock of computers falls slightly.

In a putty-clay world, capital cannot be reshaped and lumped together, but is made up of many machines, buildings, and software programs used by different workers. The purchase of a billion dollars of new computers only affects the productivity of the workers using those new computers. It leaves the productivity of workers using other computers, or no computers, unaffected.[1] The new computers do not reduce the productivity of existing computers.

This paper breaks each type of capital into units called "machines." If there are M different types of capital, then each employed worker uses exactly M machines, one of each different type. Also, each machine is used by only one worker. Thus, the total number of machines used is M times the total number of people working. If an economy were in a no-growth steady state, the size of each machine of type m would equal the total stock of type-m capital, in dollars, divided by the total number of workers at full employment. In the real world, many workers do not use every type of machine, and many machines, e.g., an airport or an airplane, are operated by more than one worker at a time. However, the assumption of one machine of each type for each worker greatly simplifies the model without introducing important distortions to the results. (See the section on simplifying assumptions below.)

[1] This does not mean that the living standards of those other workers are unaffected; the additional computers reduce the relative price of goods and services produced using those computers. For example, barbers benefit when automakers upgrade their computers, because the cost of a car falls relative to the price of a haircut.

In this paper, the “output” of a machine refers to its gross output, namely the entire output of the worker using that machine. Thus, the output of the computer used by a worker is the same as the output of the office used by that worker, which is the same as the output of the furniture used by that worker. The fact that the whole of a given worker’s output can be attributed to each machine used means that the sum of the outputs of all the machines that depreciate in a given year can actually exceed total output if the number of machines depreciating exceeds the total number of workers. The same is true for the sum of the outputs of replacement machines and is even more likely to be true for the sum of the outputs of new machines, which include replacement machines as well as machines that expand capacity. One must remember the potential for double-counting whenever “adding up” the output produced by different machines.

Capacity is the output that a set of machines—one of each type—produces when used by a worker, ignoring cyclical effects on productivity. Because productivity is procyclical, the output of a set of machines can exceed capacity when demand is rising rapidly, and fall short of capacity when demand is rising slowly or falling. If a set of machines is not used by a worker, its output is zero.

In general, the economy is capable of producing more output than it does. For example, capacity utilization in manufacturing has never risen above 92 percent. That could be because there are always several businesses with unused plant and equipment, or because businesses that could employ existing capital for a greater number of hours per day, for example by adding another shift, do not.

Based on that empirical observation, this paper assumes that businesses always have spare capacity that could be utilized if they increased total hours worked. However, I do not model the reason firms are willing to purchase capacity that will likely be idle.[2] Instead, I define capacity at the level of the business, and thus of the economy, to exclude the fraction that, on average, is spare capacity.

2.2 Background for the Production Function

This paper assumes that the production function has certain properties: capital is putty-clay; technical progress is Hicks-neutral, and is the same for all capital; and output is Cobb-Douglas in labor and a non-Cobb-Douglas aggregate of capital. Before presenting the production function in mathematical detail, it is useful to step back and review the reasons for assuming those properties.

[2] One possibility, suggested by Gilchrist and Williams (1998), is that businesses do not know ex ante how productive a given asset will be, and ex post leave their least productive assets idle.

2.2.1 Putty-Clay

Before capital is put in place, a business can choose from a wide array of different ratios of labor to capital—a property known as ex ante variable proportions. For example, a trucking company can choose from a variety of different truck sizes. Since only one driver operates a truck at once, a more expensive truck implicitly means a higher ratio of capital to labor. Similarly, a business buying new personal computers, peripheral equipment, and software can choose from a wide variety of capacities, at different prices. Assuming that only one person can operate a PC at once, equipping workers with more-expensive information technology implicitly boosts the level of capital per worker.

However, once that capital is put in place, the ratio of capital to labor embodied in that asset remains the same as long as it is in use. That assumption of ex post fixed proportions differs from the older and more common neoclassical assumption of ex post variable proportions (e.g., Solow (1956)). Ex post variable proportions is also known as putty-putty, because businesses are assumed to freely re-mold existing capital in order to vary its labor requirement as desired. Thus, as the cost of capital falls relative to the cost of labor, businesses mold new investment together with existing capital in order to hold constant the ratio of a worker's compensation per hour to the cost of the capital being operated by that worker. Every time the cost of capital falls one percent relative to the hourly cost of labor, businesses increase the size of each existing machine by one percent. With ex post fixed proportions, changes in the cost of capital affect only the labor intensity of new capital. Businesses cannot change the labor requirements of existing capital.

The empirical behavior of investment supports the assumption of ex post fixed proportions (clay) over the assumption of ex post variable proportions (putty). Researchers have found that the short-run response of investment to changes in output is much larger than the short-run response to changes in the real cost of capital (see the survey by Chirinko (1993)). This is consistent with clay capital but not with putty capital. Assume that production is Cobb-Douglas between capital and labor, consistent with the observation that capital's share of income is fairly stable over long periods of time.[3] Then, if capital were putty, each one percent fall in the real cost of capital would lead to a one percent rise in the capital stock, the same short-run result as from a one percent rise in real output. However, since new investment in any year is only a fraction of the capital stock, a one percent rise in the capital stock means a much larger percentage increase in investment. If capital is clay, a one percent fall in the real cost of capital only affects the labor intensity of new capital, and thus leads to a one percent rise in investment, a much smaller change in investment than from

[3] Stability of capital's share of income (vK/PY, where v is the rental cost of capital, K is the capital stock, P is the price of output, and Y is output) over time, coupled with the first order conditions ($\frac{\partial(PY)}{\partial K} = v$) implies that $\frac{\partial(PY)}{\partial K} = \alpha\frac{PY}{K}$, where α is capital's share of income. This in turn implies that dlnY=αdlnK. Adding the assumption of constant returns to scale produces the Cobb-Douglas production function $Y = AK^{\alpha}L^{1-\alpha}$.

a one percent rise in output.

2.2.2 Technical Progress

Although the labor requirement of any unit of capital remains constant over its lifetime, the output of that capital can change. This paper assumes Hicks-neutral technical progress and assumes that such technical progress is the same for all capital (in percentage terms). Thus, a one percent rise in technology boosts each worker's output by one percent, no matter the age or combination of equipment and structures that worker is using. However, it leaves the labor requirement of each existing unit of capital unchanged. The output of a given worker thus depends only on the overall level of technology and on the real quantity of each type of capital that worker is using, and not on the vintage of that capital.[4]

Alternatively, one could instead assume that productivity growth occurs only as new vintages of capital are introduced, and that once capital is put in place, the amount of output it can produce is fixed. However, at least some evidence argues against this assumption. One example is Lundberg's "Horndal effect," discussed in Arrow (1962). The Horndal iron works in Sweden had no new investment for a period of 15 years, yet output per manhour rose on average close to 2 percent per year.

A second counterexample to a pure vintaging model for technical change is the behavior of labor productivity in the U.S. nonfarm business sector during the early 1990s. In 1992, nonfarm business productivity grew 3.7 percent, after growing no faster than 1.3 percent during any of the previous five years. According to a vintaging model, this could only happen if new capital available in 1992 were significantly more productive than existing capital. However, productivity grew just 0.5 percent in 1993, despite more rapid growth in output, higher investment, and much more rapid growth in hours. (Nonfarm business hours actually fell in 1992, making a cyclical explanation for that year's strong productivity growth difficult.) A pure vintaging model would imply that in 1993, equipment makers forgot how to make the types of capital they made in 1992. Changes in the productivity of existing capital can better explain the data.

The model also assumes that demand has an important impact on productivity. That is, given labor, capital and technology, a higher level of demand, or more rapid growth in actual output than in output at full employment, will boost labor productivity. The assumption that demand or demand growth affect productivity has a long history in the literature (e.g., Kuh (1965)).

Except for fluctuations in productivity due to demand, I assume constant returns to scale at the aggregate level. Thus, for a given technology, a one percent increase in

[4] The vintage only matters in that the size of the machine depends on its cost at the time it was built.

labor hours and a one percent increase in the capital aggregate lead to a one percent increase in output.

2.2.3 Capital in the Production Function

This paper assumes that labor hours and aggregate capital services are combined to produce output according to a Cobb-Douglas production function. Thus, a one percent drop in the aggregate cost of capital relative to the price of output leads to a one percent rise in the desired flow of aggregate capital services relative to output. The assumption of a Cobb-Douglas production function in empirical work on business fixed investment goes back at least to Jorgenson (1963). The relative stability of labor's share of total income since World War II lends support to this assumption (see Figure 1). However, the model presented in this paper could accommodate alternative specifications of the production function.

In addition, I assume that there are many types of capital, each with its own service life and tax treatment.[5] A brief look at BEA's estimated depreciation rates or at the tax code confirms that economic depreciation, asset lives, and depreciation rules vary across different types of capital.

However, this paper does not combine capital services across those different assets using the same Cobb-Douglas methodology used to combine labor and aggregate capital services. When capital services are aggregated using a Cobb-Douglas function, a one percent drop in the cost of one type of capital relative to costs of other types of capital leads to a one percent rise in the desired flow of services from that type of capital relative to services from other types of capital.

Instead, this paper assumes that changes in the cost of one type of capital lead to decreases in demand for substitute forms of capital.[6] The elasticities of substitution between different types of capital exceed 1.0 in absolute value. Thus, for example, a one percent fall in the cost of computers relative to the cost of other capital will boost demand for computer services relative to the services of other capital by more than one percent. The general upward trend in the shares of computers and software in nominal business fixed investment is consistent with this assumption.[7]

These assumptions about how capital enters the production function mean that com-

[5] For the sake of tractability, some of the categories of investment that I estimate are aggregates of several types of capital. However, one could estimate investment in a broader array of capital types within the structure of the model.

[6] Kiley (2001) presents a similar capital aggregate in a putty-putty model of investment.

[7] Alternatively, Tevlin and Whelan (2000) explain the upward trend in the share of computers and software in total nominal investment by assuming that the elasticity of substitution between computers and labor exceeds (is more negative than) -1.0, while the elasticity of substitution between other investment and labor is less than -1.0. For some period of time, that assumption may be consistent with labor having a stable share of total income. Eventually, however, if real computer prices continue to fall, labor's share of total income must fall as well.

puters and software substitute for other forms of capital, and not for labor. Any model of investment that assumes computers (or computers and software) have an elasticity of substitution greater than -1 while other types of investment have an elasticity of substitution equal to -1 implicitly assumes that computers substitute for labor.

This paper's assumptions about capital's role in production are consistent with postwar data on income shares. Another way of explaining the postwar data, not explored in this paper, is that computers or software do substitute for labor, but that a shift in output toward industries with higher-than-average labor shares of income has masked this trend.

2.3 The Basic Production Function

For worker n at time t, the production function is

$$y_{nt} = A_t\ U_t\ \left[G_t\left(k_{1nt}, ..., k_{Mnt}\right)\right]^{\alpha},$$

where y_{nt} is output of worker n at time t, A is economy-wide technology, U is the effect of economy-wide intensity of usage, or effort, on productivity, G() is the function aggregating different types of capital, k_{mnt} is the size of the machine of type m (in constant dollars) used by worker n at time t, and α is capital's coefficient in the Cobb-Douglas production function. There are M types of capital. Production of worker n is zero if he is not employed.

Total factor productivity (TFP) is the product of economy-wide technology A and economy-wide effort U. Technology is what TFP would be if there were no fluctuations in labor usage from "full employment." Effort depends primarily on changes in the ratio of actual output to output at full employment, but also on the level of that ratio. Empirically, most movements in TFP stem from movements in technology A.

The production function exhibits constant returns to scale. If x workers use the same amounts of each type of capital, then total output is x times the output of an individual worker, and total capital used is x times the amount of capital used by each worker.

This paper assumes that the capital aggregating function G also exhibits constant returns to scale. Except in the case of a Leontief function, this means that

$$\sum_{m=1}^{M} s_{mnt} = 1, \qquad \text{where } s_{mnt} = \frac{\partial \ln G_t\left(k_{1nt}, ..., k_{Mnt}\right)}{\partial \ln k_{mnt}}.$$

Note that both changes in the function G and price-induced changes in the relative k may produce changes in the capital shares s. For example, both the introduction of the personal computer and relative declines in the price of computing power appear to have boosted computers' share of businesses' total capital budget.

Constant returns to scale in the function G is not necessary for constant returns to scale in the overall production function. The production function guarantees constant returns to scale no matter what G is. Whatever the function G, a doubling of the amount of labor and capital leaves capital per worker k_{mnt}, and thus output per worker y, unchanged, causing a doubling of total output. However, if the function G does not exhibit constant returns to scale, it is unlikely that labor's share of income will equal the parameter α.

2.4 Simplifying Assumptions

> All theory depends on assumptions which are not quite true. That is what makes it theory. The art of successful theorizing is to make the inevitable simplifying assumptions in such a way that the final results are not very sensitive.
>
> Robert Solow (1956, p. 65)

This paper makes some important simplifying assumptions about how output is produced. Briefly, I assume that each worker faces the same production function, that each worker has the same labor quality, and that capital is reallocated among workers freely and often. While none of these assumptions are quite true, I do not think that they affect the final results in a fundamental way. However, they do allow the expected production of a new machine to be expressed in a convenient form.

2.4.1 Same Production Function for All Workers

This paper assumes that, at any one time, each worker is using the same production function. This assumption has two implications. First, the prices of each type of good and service produced are the same, or, equivalently, there is only one type of good produced. Second, every worker uses a similar, although not identical, mix of different types of plant and equipment.

Obviously, none of this is true in the real world. However, the assumption of one price is not as distorting as it may seem. Over time, relative prices between goods and services move inversely to the relative productivity of the industries or processes producing them. For example, the relative price of computers has gradually fallen because the productivity of the processes producing its components, especially semiconductors, has risen over time. Thus, while price (p) and technology (A) will likely be different for different goods and services, the product of the two (pA) will move in similar ways over time. Since it is this product that matters for investment, the assumption of one price and level of technology for all goods and services should not have an important effect on the results.

It is more difficult to say whether the assumption of a common mix of capital has an important effect, but it seems unlikely. In reality, there is no worker who uses every type of capital in doing her job. Still, non-linearities are probably not large enough that investment would be much different if one worker drove a truck and another operated a drill press than if each used both machines at different times. As long as the labor requirement of a unit of capital remains the same over its useful life, investment should be broadly consistent with the model in this paper.

Similarly, in the real world, the percentage of workers using certain types of capital, like computers, can change considerably over time. Implicitly, that means that the fraction of workers using a production function containing that type of capital changes over time. However, if such changes occur gradually, a shift in the fraction of workers using a certain type of capital can be closely approximated by a rise in the share of such capital in each worker's capital aggregate.

More problematic is the implicit assumption that every industry has the same capital intensity. That assumption does not matter if the composition of output remains the same over time. However, if demand shifts toward or away from industries with higher-than-average capital intensity, investment would be affected. The relative stability of labor's share of income over time (discussed above) could be consistent either with a constant composition of output or with a shift away from capital-intensive industries combined with a decline in labor's share of income in most industries.

2.4.2 Same Labor Quality for All Workers

This paper assumes that every worker is of the same quality, both at any one time and across time periods. For the most part, this assumption probably makes little difference. Variations in average labor quality over time have the same effect as variations in technology, and so are subsumed in the technology variable A. In practice, not separating out known variations in labor quality before using a Kalman filter to parse TFP into technology and intensity of use may impart some bias to those estimates.

Similarly, variations in labor quality between workers at any one time could be handled by replacing actual labor hours with effective labor hours, with each worker's effective hours being equivalent to that worker's quality relative to the average. Problems arise only to the extent that a worker's relative quality changes over time. Given that such changes are frequently gradual (with experience), the results should not be much affected.

Implicitly, the assumption that every worker is of the same quality also eliminates the possibility that, as a piece of capital ages and becomes less productive than new capital of the same type, it is passed down from high-quality workers to lower-quality workers. Such exchanges of capital would inject an element of ex post variable proportions into the model. Given the lack of empirical support for ex post variable proportions, the assumption that such exchanges do not occur seems justified.[8]

2.4.3 Capital Reallocation is Frequent and Costless

The model assumes that businesses reallocate capital among workers freely and often. Obviously, that assumption is not literally true, since a given worker will often use the same piece of equipment or work in the same building for a considerable length of time. However, if one does not make that assumption, then the size of each new machine k_{mnt} depends on the sizes of existing machines of other types used by worker n. Even more problematic, if any of those existing non type-m machines requires replacement during the service life of the new machine, the size of its replacement depends on both the size of the new type-m machine and the size of the type-m machine that will replace it in the future. Even for an economy with only one worker, mathematically solving for the optimal size of a new machine is enormously complex.

The assumption of frequent reallocation of capital means that a business cannot tell which new machine will be combined with which combination of existing machines. Thus, the size of new machines of a given type at time t is uniform, and depends on the overall distribution of existing machines of different types. In addition, the size of any replacement non type-m machine that a new type-m machine is later combined with does not depend on the size of the new type-m machine, but on the entire distribution of type-m machines existing at the time of replacement.

The possibility of endogenous replacement means that the assumption of frequent reallocation has little impact on the results, because the difference in productivity between a combination of old capital and a combination of new capital cannot be greater than the cost of scrapping the old capital and buying new capital. This guarantees that non-linearities are not large. So, while the output of a worker using a combination of machines each 10 percent as large as the average and the output of a worker using a combination of machines each 190 percent as large as the average would fall well short of the output of two workers using a combination of average machines, the possibility of endogenous replacement rules out such outcomes.

2.5 Expected Output

The simplifying assumptions discussed above allow us to rewrite the production function for a worker into expected output associated with a machine of type m installed at time t. Since all labor is assumed to have the same quality, a machine's output

[8] If one added labor quality to the model, the possibility of gradually passing capital down the ranks of labor quality during its service lifetime could be eliminated by assuming that each piece of equipment remains in the same tier of labor quality throughout its useful life, and that the relative position of those tiers also remains constant over time.

is the same no matter which particular worker uses that machine. In addition, the reallocation assumption means that the expected output of a machine is not tied to a particular bundle of other types of machines, but rather to the expected overall distribution of other types of machines at the time of production. Thus, expected output associated with a particular machine depends only on the size of that machine and on economy-wide variables: expected technology A (including average labor quality), expected effort U, and the expected size of machines of other types. In consequence, all machines of type m installed at time t are the same size, $k_{m,t}$.

Let $y_{m,t+i,i}$ denote the output at time t+i of a worker using capital of type m aged i periods at time t+i, i.e., installed at time t, and let the prescript t denote expectations at time t. Then the expectation at time t of the output to be produced at time t+i by a worker using capital of type m installed at time t, ${}_ty_{m,t+i,i}$, is the expected value of $A_{t+i}\ U_{t+i}\ \left[G_{m,t+i}\left(k_{m,t}\right)\right]^{\alpha}$. As an approximation, I assume

$$ {}_ty_{m,t+i,i} = {}_tA_{t+i}\ {}_tU_{t+i}\ \left[{}_tG_{m,t+i}\left(k_{m,t}\right)\right]^{\alpha}. $$

The expected sizes of machines of types other than m at time t+i are subsumed in the function $G_{m,t+i}$. Partly for that reason, the function G depends on time and is specific to m.

3 Investment

The total number of new machines equals the number of replacement machines plus the net change in the number of machines. Define "expansion machines" as the net change in the number of machines of each type (in a frictionless world, these are all equal), and "expansion capacity" as the amount by which the expansion machines change total capacity. Depending on the total desired change in capacity, the number of expansion machines may be positive or negative. By the assumption of frequent reallocation, most expansion machines are used in combination with existing machines of other types, so expansion capacity (if positive) will likely differ from the amount of output that new machines would produce if they were used in combination with each other instead of with existing machines.

During any period of time, businesses want to invest so that expansion capacity plus the growth of the productive capacity of existing capital plus the difference in output between replacement machines and the machines they replace equals the desired change in capacity during that period. Rearranging that verbal equation, expansion capacity equals the desired change in capacity less the increase in capacity from replacing depreciated capital with new capital less gains in the productive capacity of workers using existing capital:

$$\begin{aligned}\text{Expansion Capacity} = {} & \text{Desired Change in Capacity} \\ & - \text{Increase in Capacity From Replacing Depreciated Capital} \\ & - \text{Change in Capacity of Non-Depreciated Capital.}\end{aligned} \tag{1}$$

Several steps are required to go from the verbal equation (1) to an expression for aggregate investment. First, one must determine the conditions under which equation (1), governing decisions about investment made by each business, also governs aggregate investment. The second step is to solve equation (1) for expansion capacity in a frictionless world where there is no lag between the time that new capacity is needed and the time it is installed. The third step is to determine the optimal size of new machines. In the next step, I combine information on expansion capacity, capacity of replacement machines, and machine size into an expression for aggregate investment in each type of capital in a frictionless world. The fifth step is to add simple dynamics, in which there is a fixed "time to build" between the time that new capacity is needed and the time that it is installed. In the final step, I enrich the dynamics to include multiple times to build and differing time lags for the effects of output and the cost of capital.

This section concludes with two subsections discussing the properties of the investment equations. The first of those examines the effects of demand shocks and productivity shocks on investment. The second highlights the differences between the investment model of this paper and the standard neoclassical investment model.

3.1 Aggregating Investment Across Businesses

To use the firm-level equation (1) as an explanation for aggregate investment, one must be able to add up each term of the equation across firms. In most cases, the ability to add up is obvious: total expansion capacity is the sum of expansion capacity for each firm; the change in the total capacity of non-depreciated capital is the sum of the change in the capacity of each firm's non-depreciated capital; and the total increase in capacity from replacing depreciated capital is the sum of each business's increase in capacity from replacing depreciated capital. But the desired change in capacity obtained from aggregate data need not equal the sum of the changes in capacity desired by each firm.

In order to create a total desired change in capacity that can be considered the sum of each firm's desired change in capacity, it must be true that the sum of the market shares that each firm in the economy expects is one, i.e., that expectations are collectively consistent. This is not unreasonable for firms in industries with customer markets in the style of Okun (1981), in which "the current level of demand experienced by a firm [depends] positively upon its volume of sales in the past."

It seems unlikely for industries producing commodities, like mining and farming. Consequently, this paper does not examine investment done by those industries.

3.2 Capacity of New Capital in a Frictionless World

In a frictionless world with perfect information, the assumption that every worker uses every type of capital means that the number of expansion machines of each type is the same, whether positive or negative. Let $N_{e,t}$ be the number of expansion machines of each type put in place at time t, expressed at an annual rate. At full employment, $N_{e,t}$ is also the number of workers using expansion machines of each type at time t. To make the math easier, time is continuous in this subsection and the next two, but is lumped together into discrete periods beginning in the subsection examining time to build.

Let y_t be the average output during period t of workers using only capital existing before time t, excluding the effect of deviations in effort U from average effort.[9] If new machines were the same size as existing machines, expansion capacity would just be $N_{e,t}\ y_t$. However, if new machines differ in size from existing machines, their output will differ from that of existing machines. The output of a worker using a new machine of type m, $y_{m,t,0}$, equals y_t plus the increment to output from using a type-m machine of different size than those existing prior to time t, $y_{m,t,0}$-y_t. Summing up the output of expansion machines and eliminating the double-counting caused by attributing the entire output of each new machine to that machine, expansion capacity can be expressed as

$$N_{e,t}\ y_t \left[1 + \sum_{m=1}^{M} \left(\frac{y_{m,t,0}}{y_t} - 1\right)\right],$$

at an annual rate.

Let $y^h_{m,t,i}$ be the hypothetical output of a worker using a machine of type m aged i years at time t had it not depreciated at time t. That output includes the effects of any improvements in technology at time t. In a putty-clay world, the output of a replacement machine ($y_{m,t,0}$, the same as the output of an expansion machine) will generally be larger than the output of the hypothetical output of a worker using a depreciated machine; capital per worker generally grows over time, so replacement machines will be more productive than those they replace. Personal computers are an obvious example. New PCs are generally much more powerful than the old PCs they replace.

Let $R_{m,t,i}$ be the number of machines of type m and age i depreciating at time t, at an annual rate. The total number of replacement machines of type m, $R_{m,t}$,

[9] Deviations in effort U from 1.0 are generally short-lived, so firms assume that future effort equals 1.0 when planning capacity, i.e., ${}_tU_{t+i} = 1$.

equals $\int_{i=0}^{L_m} R_{m,t,i} di$, the total number of type-m machines depreciating at time t, all at annual rates. At time t, the annual rate of increase in capacity obtained by replacing depreciated machines of all types with new machines is thus:

$$\sum_{m=1}^{M} R_{m,t}\ y_{m,t,0} - \sum_{m=1}^{M} \int_{i=0}^{L_m} R_{m,t,i}\ y_{m,t,i}^{h}\ di.$$

Let $\dot{A}_t$ be annualized growth of technology at time t. Then the rate of increase of capacity due to technology growth at time t is $\left(\dot{A}_t/A_t\right) N_t y_t$, where N_t is the total number of machines of each type at time t. In discrete time, N_t becomes N_{t-1}, the number of workers using machines already existing at time t.

Finally, let $\dot{Y}E_t$ be the desired annualized change in output at time t. In a frictionless world, with no time to build, the verbal equation for output from new investment (1) can be expressed mathematically as:

$$\begin{aligned} N_{e,t}\ y_t \left[1 + \sum_{m=1}^{M} \left(\frac{y_{m,t,0}}{y_t} - 1\right)\right] &= \dot{Y}E_t - \sum_{m=1}^{M} R_{m,t}\ y_{m,t,0} \\ &\quad + \sum_{m=1}^{M} \int_{i=0}^{L_m} R_{m,t,i}\ y_{m,t,i}^{h}\ di - \left(\dot{A}_t/A_t\right) N_t\ y_t. \end{aligned} \tag{2}$$

The output of new machines of type M can be found by adding the output of type-m replacement machines to both sides of equation (2) and rearranging, to yield:

$$\begin{aligned} \left(N_{e,t} + R_{M,t}\right)\ y_{M,t,0} &= \dot{Y}E_t + R_{M,t}\ y_t - \left(\dot{A}_t/A_t\right) N_t\ y_t \\ &\quad N_{e,t} \sum_{m=1}^{M-1} \left(y_{m,t,0} - y_t\right) + R_{M,t}\left(y_{M,t,0} - y_t\right) \\ &\quad - \sum_{m=1}^{M} R_{m,t}\ y_{m,t,0} + \sum_{m=1}^{M} \int_{i=0}^{L_m} R_{m,t,i}\ y_{m,t,i}^{h}\ di. \end{aligned} \tag{3}$$

The first line of equation (3) states that the capacity of workers using new machines equals the desired change in capacity plus (roughly) the capacity of depreciated machines less the change in capacity from using improved technology with existing machines. The second and third lines adjust that identity for the difference between the output of new machines and existing machines.

3.3 Optimal Size of New Machines

To determine investment, one must replace output per worker $y_{M,t,0}$ in equation (3) with the size of new machines of type M. Businesses choose the size of machine that maximizes expected profits. The optimal size of new machines depends on output per worker and on the cost of capital.

The present discounted value (PDV) of after-tax profits associated with a new machine of type m equals the PDV of after-tax revenues of the worker using that machine, less the PDV of the after-tax cost of that worker, less the PDV of the after-tax cost of the non type-m machines also used by that worker, less the purchase price of the machine, plus any investment tax credit, plus the PDV of the tax value of depreciation allowances. The reallocation assumption means that the size of a new machine of type m does not affect the expected size of machines of non type-m machines used with it. In addition, the after-tax cost of a unit of labor is independent of the size of any particular new machine of type m. Consequently, in choosing the size of a new machine of type m, the after-tax costs of both labor and non type-m machines can be ignored.

Mathematically, the PDV of after-tax profits associated with a new machine of type m purchased and put into service at time t, excluding associated after-tax costs of labor and non type-m machines, is

$$\pi^*_{m,t} = \int_0^{L_m} \left[(1-u_{t+i})\, p_{t+i}\; y_{m,t+i,i}\; F_{t,i}\right]\, di \; - q_{m,t}\; k_{m,t}\,(1-C_{m,t}) + q_{m,t}\; k_{m,t}\; Z^u_{m,t},$$

where L_m is the service lifetime of capital of type m, u is the tax rate on corporate income, p is the price of output y, F is the discount factor $F_{t,i} = e^{-\int_0^i r_{t+j} dj}$ for nominal rate of return r_{t+j} at time t, $q_{m,t}$ is the purchase price of new capital of type m, $C_{m,t}$ is the rate of investment tax credit for capital of type m, and $Z^u_{m,t}$ is the PDV of taxes saved through depreciation allowances for each dollar of new capital of type m.

The PDV of taxes saved through depreciation allowances per dollar of new machine purchased, Z^u, is

$$Z^u_{m,t} \equiv \int_0^{TL_{m,t}} \left[u_{t+i}\; D_{m,t,i}\,(1-B_{m,t}\; C_{m,t})\, F_{t,i}\right]\; di,$$

where, for type of capital m, TL is the tax lifetime, $D_{m,t,i}$ is the share of depreciation allowances taken i years after time of purchase t, and B is the share of the investment tax credit that is deducted from the allowable base for depreciation.

Businesses choose the size of a new machine in order to maximize the expected after-tax profits associated with the use of that machine. Each business takes the prices of output and machines as given.[10] Setting the derivative of expected profits to zero,

$$\begin{aligned}\frac{\partial_t \pi^*_{m,t}}{\partial k_{m,t}} &= \int_0^{L_m} \left[(1 - {}_t u_{t+i}) \;\; {}_t p_{t+i} \frac{\partial_t y_{m,t+i,i}}{\partial k_{m,t}} \;\; {}_t F_{t,i} \right] \, di \; - q_{m,t} \left(1 - C_{m,t}\right) + q_{m,t} \;\; {}_t Z^u_{m,t} \\ &= 0.\end{aligned}$$

That is, at the optimal size of a new machine, the extra after-tax revenues generated from a larger machine would exactly offset the increase in the after-tax cost of that machine.

In order to determine optimal machine size, one must solve the first order condition for $k_{m,t}$. A helpful first step is to replace the derivative $\partial y / \partial k$. Define the elasticity of the capital aggregator G with respect to a change in the size of i-year-old type-m machines at time t+i as

$$s_{m,t+i,i} \equiv \frac{\partial \ln G_{m,t+i}\left(k_{m,t}\right)}{\partial \ln k_{m,t}}.$$

Then we can rewrite the first order conditions as

$$\int_0^{L_m} \left[(1 - {}_t u_{t+i}) \;\; {}_t p_{t+i} \; \alpha \;\; {}_t s_{m,t+i,i} \frac{{}_t y_{m,t+i,i}}{k_{m,t}} \;\; {}_t F_{t,i} \right] \, di \; = q_{m,t} \left(1 - C_{m,t}\right) - q_{m,t} \;\; {}_t Z^u_{m,t},$$

or

$$q_{m,t} \;\; k_{m,t} \left[1 - C_{m,t} - {}_t Z^u_{m,t}\right] = \alpha \int_0^{L_m} \left[(1 - {}_t u_{t+i}) \;\; {}_t p_{t+i} \;\; {}_t s_{m,t+i,i} \;\; {}_t y_{m,t+i,i} \;\; {}_t F_{t,i} \right] \, di.$$

If s_m is constant over time, then businesses choose new machines of a size that equates the expected after-tax cost of a new machine with a share αs_m of the PDV of the expected after-tax output to be produced by the worker using that machine.

In order to calculate the PDVs of expected after-tax output and the expected tax value of depreciation allowances, one must make assumptions about businesses' expectations of taxes, interest rates, prices, growth of output per machine, and elasticity s_m of the capital aggregator G. This paper assumes that businesses expect the tax treatment of profits and depreciation to remain unchanged from current law at the time of investment. In addition, businesses treat the rate of return, the rate of increase for overall prices p, and the rate of growth of real output per worker using a machine of type m purchased at time t as constant over time.[11] Finally, businesses

[10] That is, the elasticity of price p with respect to output per capita y is assumed to be zero. Assuming that this elasticity is a constant negative number would merely mean multiplying the price of output by a constant term in all the equations.

[11] In the empirical section, I assume that expectations for the nominal rate of return and the rate of inflation depend on the service life of the asset. For assets with long service lives, the nominal rate of return and the expected rate of inflation are weighted averages of current values and historical averages.

assume that the elasticity s_m of a new machine remains constant over its lifetime.[12] With these assumptions, the equation for optimal machine size becomes

$$q_{m,t}\ k_{m,t}\left[1 - C_{m,t} - {}_tZ^u_{m,t}\right] = \alpha\ s_{m,t,0}\,(1-u_t)\,p_t\ y_{m,t,0}\int_0^{L_m} \exp\left({}_t\dot{p} + {}_t\dot{y}_m - {}_tr\right)\ di,$$

where a dot over a variable denotes an instantaneous rate of growth.

Evaluating the integral and rearranging terms, the optimal machine size is

$$k_{m,t} = \alpha\ s_{m,t,0}\frac{p_t}{v_{m,t}}y_{m,t,0}, \tag{4}$$

$$\text{where}^{13}\ v_{m,t} \equiv q_{m,t}\frac{{}_tr - {}_t\dot{p} - {}_t\dot{y}_m}{1 - \exp\left[-\left({}_tr - {}_t\dot{p} - {}_t\dot{y}_m\right)L_m\right]}\frac{\left[1 - C_{m,t} - u_t\ {}_tZ_{m,t}\right]}{(1-u_t)} \tag{5}$$

$$\text{and}\ {}_tZ_{m,t} \equiv \int_0^{TL_{m,t}} \left[D_{m,t,i}\,(1 - B_{m,t}\ C_{m,t})\exp\left(-{}_tri\right)\right]\ di.$$

The variable Z is the PDV of depreciation allowances per dollar of new machine of type m. The optimal size of a new machine k of type m is proportional to the elasticity of output with respect to capital of type m at time of purchase, $\alpha s_{m,t,0}$, to the ratio of the price of output p to the cost of capital v, and to output per worker y. Machine size is thus inversely proportional to the real price of new capital (q_m/p). Optimal machine size is inversely correlated with the expected real cost of funds (r-$\dot{p}$) and the corporate tax rate, and is positively correlated with expected growth in the output of labor using the new machine ($\dot{y}_m$), its service life (L_m), the investment tax credit (C_m), and the value of depreciation allowances (Z_m).

The cost of type-m capital $v_{m,t}$ can influence the optimal size of a new type-m machine $k_{m,t}$ through two channels: direct and indirect. The cost of capital appears directly in the equation for the optimal size of new machines. All else equal in equation (4), a one percent drop in the cost of capital would lead to a one percent increase in the size of new machines.

In addition, the cost of capital can affect new machine size indirectly through the elasticity term s_m. Suppose, for example, that the capital aggregating function G exhibits a greater degree of substitution than a Cobb-Douglas function, as appears to be the case empirically. Then, as the relative price of type-m capital falls, the elasticity s_m rises, and businesses put a greater share of their capital budget into

[12] That is, ${}_ts_{m,t+i,i} = s_{m,t,0}$. In fact, $s_{m,t+i,i}$ will generally be slightly declining in i, so that the right-hand side of the equation is slightly smaller than a share $\alpha s_{m,t,0}$ the PDV of the expected after-tax output to be produced using the machine. While the elasticity s of an existing machine of type m remains the same during its lifetime, the elasticity of new machines of type m may change. That is, $s_{m,t+i,0}$ will not necessarily equal $s_{m,t,0}$.

[13] An expression similar to the first ratio on the right-hand side of the equation can be found in both Solow (1962) and Phelps (1963).

type-m capital. So, as the relative price of computers and software has fallen over time, investment in computers and software has accounted for a larger and larger share of total investment.

3.4 Investment in a Frictionless World

In this paper, I make an important assumption that greatly simplifies estimation of the investment equation. I assume that $y_{m,t,i} = y_t$ for all i, i.e., that output per machine for depreciated machines, for new machines, and for existing machines is the same. This simplification is not necessary for estimation, and I intend to relax it in future work. However, the calculation for the variable y^h is quite complex, and is not included in this paper. With this important simplification, equation (3) becomes

$$(N_{e,t} + R_{m,t})\, y_t = \dot{Y}E_t + R_{m,t}\ y_t - \left(\dot{A}_t/A_t\right) N_t\ y_t. \tag{6}$$

The likely impact of the simplifying assumption that $y_{m,t,i} = y_t$ for all i is that equation (6) understates the importance of replacement capital, and thus overstates the importance of expansion capacity. That result can be seen by comparing equation (6) with equation (3). While the output of replacement capital equals the output of depreciated capital in (6), the output of replacement capital exceeds the output of depreciated capital in (3). Thus, by understating the output produced by replacement capital, equation (6) overstates the need for expansion capacity.

Investment in capital of type m at time t ($I_{m,t}$) is found by multiplying the number of machines of type m by the size of new type-m machines $k_{m,t}$:

$$I_{m,t} = (N_{e,t} + R_{m,t})\, k_{m,t}. \tag{7}$$

After substituting for $k_{m,t}$ from equation (4), employing the simplifying assumption that $y_{m,t,0}$ equals y_t, and substituting for $(N_{e,t} + R_{m,t})\, y_t$ using equation (6), equation (7) becomes

$$I_{m,t} = \alpha\ s_{m,t,0} \frac{p_t}{v_{m,t}} \left[\dot{Y}E_t + R_{m,t}\ y_t - \left(\dot{A}_t/A_t\right) N_t\ y_t\right]. \tag{8}$$

(Investment, growth of desired output and technology, and the number of replacement machines R are all expressed at annual rates.) Investment in machines of type m is proportional to the elasticity of output with respect to capital of type m ($\alpha s_{m,t,0}$), to the ratio of output price to the cost of capital, and to a function of desired growth of capacity, replacement demand, and growth in technology.

Based on rough calculations, the simple putty-clay cost of capital defined in equation (5) may imply a higher level of investment than is observed in the data. So instead, I use the standard neoclassical cost of capital under ex post variable proportions v', with the exception that the capital gains term q̇ is replaced by the growth rate of output prices ṗ. Thus,

$$v'_{m,t} = q_{m,t} \left({}_t r - {}_t \dot{p} + db_m / L_m \right) \frac{\left[1 - C_{m,t} - u_t \; {}_t Z_{m,t} \right]}{(1 - u_t)}, \tag{9}$$

where db is a declining-balance parameter that may differ from 1.0 (Fraumeni, 1997).

Reconciling the putty-clay cost of capital with actual investment levels may require raising the marginal cost of capital without raising the marginal output of capital. One way to raise marginal cost without raising marginal output is to add maintenance and repair costs that are proportional to machine size to the model.

3.5 Investment With Simple Time to Build

Thus far, this section has assumed that investment occurs instantaneously with the changes in demand that lead to that investment. In reality, there are important lags between the time that output rises and the time that new capital is finally put in place. First, it takes some time for businesses to recognize the need for more capital, and to plan for how to meet that need. Second, it takes some time after such plans are finalized to put new capital in place. The presence of unfilled orders for nondefense capital goods suggests that equipment manufacturers do not fill all orders immediately. For structures, considerable time elapses between ground-breaking and final completion. Finally, some investment will lag further behind changes in output if capital is "lumpy," i.e., if businesses do not adjust their capital stock until some threshold level of adjustment is required.[14] Similarly, lags will be added to the investment process if integrating a large number of new factories into production during a short period of time is more costly than integrating those same factories into production one by one over a longer period.

Some stylized facts suggest that these lags, referred to collectively as the "time to build," are important. First, percentage fluctuations in the capital stock are smaller than percentage fluctuations in output. If lags did not exist, percentage movements would be identical. Nonetheless, there is a correlation between growth in output and growth in capital, suggesting that the lags or adjustment costs are not so important that businesses do not vary purchases of capital with market conditions. Finally, the correlation between lagged changes in output and current changes in the capital stock (investment) suggests that time to build is important.

Assume that the interval between the beginning of the planning stage and the final delivery of capital—the time to build—is T periods. Then expectations at time t-T and the change in expectations between times t-T-1 and t-T determine investment

[14] For evidence that capital is lumpy, see Cooper, Haltiwanger, and Power (1999) and Doms and Dunne (1998).

at time t. Those expectations include both the expectations determining machine size and the expectations determining the number of new machines needed.

By definition, $I_{m,t} = {}_{t-T}I_{m,t}$, where the prescript t-T denotes expectations at time t-T. Also, assume firms know their future replacement requirements, so that $R_{m,t} = {}_{t-T}R_{m,t}$. Then, in discrete time, equation (8) becomes

$$I_{m,t} = \alpha \; {}_{t-T}s_{m,t,0} \; \frac{{}_{t-T}p_t}{{}_{t-T}v'_{m,t}} \times \left[\begin{array}{c} \left({}_{t-T}YE_t - \; {}_{t-T-1}YE_{t-1}\right) + R_{m,t} \cdot \; {}_{t-T}y_t \\ - \left(\frac{{}_{t-T}A_t - \; {}_{t-T-1}A_{t-1}}{{}_{t-T}A_t}\right) \; {}_{t-T}N_{t-1} \cdot \; {}_{t-T}y_t \end{array} \right]. \tag{10}$$

Since new capital to be installed at time t-1 reflects expectations at time t-T-1, new capital to be installed at time t must incorporate changes in desired future capacity and expected future technology between periods t-T-1 and t-T. So, for example, the term for the change in desired capacity ${}_{t-T}YE_t - \; {}_{t-T-1}YE_{t-1}$ can be split into two pieces, ${}_{t-T}YE_{t-1} - \; {}_{t-T-1}YE_{t-1}$ and ${}_{t-T}YE_t - \; {}_{t-T}YE_{t-1}$. The first piece reflects the revision to capacity desired for period t-1 based on information obtained between periods t-T-1 and t-T. Since orders for capacity in t-1 have already been placed in period t-T-1, that change in desired capacity cannot be achieved in period t-1. Instead, the firm adjusts orders in period t-T to achieve that revision to desired capacity, which then occurs in period t. The second piece is the growth in desired capacity between periods t-1 and t based on information available in period t-T. That includes such factors as expected growth in demand between the two periods.

Assume that the expected gap between actual output Y and output at full employment $\bar{\mathrm{Y}}$ is a linear function of the current gap between actual output and output at full employment, i.e.,

$${}_{t-T}Y_t - {}_{t-T}\bar{Y}_t = \theta_T \left(Y_{t-T} - \bar{Y}_{t-T}\right).$$

In addition, assume that output at full employment T periods hence is expected to be a multiple η_T of current output at full employment, so ${}_{t-T}\bar{Y}_t = \eta_T \; \bar{Y}_{t-T}$. (Implicitly, firms assume that labor hours at full-employment are a random walk. Otherwise, η would depend on the forecast for growth in labor hours.) Absent additional shocks, deviations of actual output from output at full employment tend to disappear over time, so the larger is T, the smaller will be θ_T.[15]

Each business would like to have enough capital at time t to be able to produce its share of total output. Adding firms together, and assuming that expectations of total market share sum to one across firms, total desired capacity equals expected

[15] The coefficient θ will be further reduced by any learning period for workers using new equipment. In such a case, the coefficient depends not only on the portion of the current gap between actual output and output at full employment expected to persist at the time of installation but also on the portion of the current gap expected to remain at various points on the learning curve.

output. Thus ${}_{t-T}YE_t = {}_{t-T}Y_t$, and

$$ {}_{t-T}YE_t = \theta_T Y_{t-T} + (\eta_T - \theta_T)\,\bar{Y}_{t-T}. $$

Taking first differences,

$$ {}_{t-T}YE_t - {}_{t-T-1}YE_{t-1} = \theta_T\,(Y_{t-T} - Y_{t-T-1}) + (\eta_T - \theta_T)\,\left(\bar{Y}_{t-T} - \bar{Y}_{t-T-1}\right). $$

In order to obtain expectations for technology, I assume that businesses expect technology to be a random walk with drift, or ${}_{t-T}A_t = \gamma_T\ A_{t-T}$. Then

$$ \frac{{}_{t-T}A_t - {}_{t-T-1}A_{t-1}}{{}_{t-T}A_t} = \frac{\gamma_T\ A_{t-T} - \gamma_T\ A_{t-T-1}}{\gamma_T\ A_{t-T}} = \frac{A_{t-T} - A_{t-T-1}}{A_{t-T}}. $$

Substituting for the change-in-expectations terms in equation (10), we have

$$ I_{m,t} = \alpha\ {}_{t-T}s_{m,t,0}\frac{{}_{t-T}p_t}{{}_{t-T}v'_{m,t}}\times \left[\begin{array}{l} \theta_T\,(Y_{t-T} - Y_{t-T-1}) + (\eta_T - \theta_T)\,\left(\bar{Y}_{t-T} - \bar{Y}_{t-T-1}\right) \\ +R_{m,t}\cdot\ {}_{t-T}y_t - \left(\frac{A_{t-T}-A_{t-T-1}}{A_{t-T}}\right)\ {}_{t-T}N_{t-1}\cdot\ {}_{t-T}y_t \end{array}\right] \quad (11) $$

All else equal, investment in type-m capital is proportional to type-m capital's share in the total cost of new capital s and to the ratio of output prices p to the cost of type-m capital v'. In addition, investment depends positively on growth in actual output Y and output at full employment $\bar{\text{Y}}$ in excess of growth in technology A. Growth in output at full employment depends on growth in full-employment labor hours multiplied by output per hour. The impact of longer service life is ambiguous, because longer service lives reduce replacement demand but also lower the cost of capital, and so increase machine size. Since output at full employment varies one-for-one with technology, a portion $(\eta - \theta)$ of the final term is cancelled out by the $\bar{Y}_{t-T} - \bar{Y}_{t-T-1}$ term.

3.6 Investment with Variable Time to Build

In reality, the data imply that investment at any given time t is influenced by businesses' expectations at several points in time, not just those at one particular time t-T. One can imagine many possible reasons for this. Businesses may be able to alter the size of a project during its planning stage, or even after it is under way, for example by cancelling orders. In addition, investment tends to be lumpy. That is, businesses often do not invest in new capital until the amount of new capital desired has reached some threshold level. Investment can thus depend on changes in demand that occurred before the planning period for the investment even began. A large change in output during one period may lead to higher investment during several periods if businesses find it difficult to incorporate a large amount of new capital at one time. One would also expect the investment of businesses at different stages of production to respond at different speeds to changes in final demand. Finally, for structures, investment projects generally take time to complete.

The possibilities of lumpy investment, adjustment costs, or a planning period after an increase in demand has occurred all suggest that lags on the cost of capital variables will be shorter on average than lags on changes in output. In all cases, at least some of the change in output leading to new investment occurs before investment plans are finalized. Since machine size need not be determined until those plans are near completion, the longest lag for the change in output (the number of machines) will exceed the longest lag for the cost of capital (machine size).

I assume that the determinants of machine size—share of capital costs s, cost of capital v', and price of output p—each have the same lag structure in the investment equation. Similarly, I assume that the determinants of the number of machines—change in actual output Y, change in output at full employment $\bar{Y}$, and change in technology A—each have the same lag structure in the investment equation. As simplifications, I assume ${}_{t-T}y_t = y_{t-1}$ and ${}_{t-T}N_{t-1} = N_{t-1}$. With variable time to build, equation (11) becomes:

$$I_{m,t} = \alpha \sum_{i=0}^{T_m} \beta_{m,i}\ s_{m,t-i,0} \frac{p_{t-i}}{v'_{m,t-i}} \times \sum_{i=0}^{\tau_m} \zeta_{m,i}\ X_{m,t-i}, \tag{12}$$

where

$$\begin{aligned} X_{m,t-i} &= \theta_m \left(Y_{t-i} - Y_{t-i-1}\right) + \left(\eta_m - \theta_m\right)\left(\bar{Y}_{t-i} - \bar{Y}_{t-i-1}\right) \\ &\quad + R_{m,t}\ \ y_{t-1} - \left(\frac{A_{t-i} - A_{t-i-1}}{A_{t-i}}\right) N_{t-1}\ \ y_{t-1}, \end{aligned} \tag{13}$$

$\tau_m \geq T_m$, and $\sum_i \beta_{m,i} = \sum_i \zeta_{m,i} = 1$ for each type of machine m.[16]

The empirical portion of this paper estimates equation (12), using v' from equation (9). Possible sources of error in equation (12) are: errors in the specification of expectations; errors in the specification of lags; the failure to account for factors causing businesses to shift investment between time periods, such as anticipated changes in tax law; errors resulting from simplifying assumptions; and errors in the model itself.

3.7 The Effects of Demand and Productivity Shocks on Investment

Changes in both productivity and demand can affect investment. Those effects can be summarized as follows. First, a positive, permanent shock to productivity leads

[16] Equation (13) implicitly assumes that θ and η are the same at every lag length. Realistically, η is larger at further lags, and θ may be smaller at more distant lags. The assumption about η makes little difference, but the assumption about θ could be important.

to a proportionate rise in investment. However, such an increase in investment does not reduce unemployment, and thus does not cause business cycle fluctuations in unemployment. Second, a positive shock to demand boosts investment more than proportionately. Third, a positive shock to productivity accompanied by an equal-sized negative shock to demand leads to lower investment.

Simplifying equation (11) makes the effects of demand and productivity on investment more clear. First, assume there is only one type of capital. Second, assume no time to build (T=0). This not only simplifies that equation, but also allows us to replace η with 1. Finally, assume that the number of workers at full employment equals the number of machines N at time t-1. Then combining equation (11) and the production function, investment is approximately

$$I_t = \alpha \frac{p_t}{v_t'} \times \left[\begin{array}{c} \theta\ \Delta Y_t + (1-\theta)\left(\frac{\Delta \bar{N}_t}{\bar{N}_{t-1}} + \frac{\Delta A_t}{A_{t-1}} + \alpha \frac{\Delta \bar{k}_t}{\bar{k}_{t-1}}\right) Y_{t-1} \\ +R_t\ y_t - \frac{\Delta A_t}{A_t} \bar{N}_{t-1}\ y_t \end{array} \right], \tag{14}$$

where Δ denotes the change in a variable from the prior period, $\bar{\mathrm{N}}$ is the number of workers at full employment, and $\bar{\mathrm{k}}$ is (a geometric average of) capital per worker at full employment. (ΔY is the discrete time analogue of $\dot{\mathrm{Y}}$.)

Using a few approximations, equation (14) can be rewritten as

$$I_t = \alpha \frac{p_t}{v_t'} \bar{Y}_t \times \left[\theta \frac{\Delta Y_t}{Y_{t-1}} - \theta \frac{\Delta A_t}{A_{t-1}} + (1-\theta)\left(\frac{\Delta \bar{N}_t}{\bar{N}_{t-1}} + \alpha \frac{\Delta \bar{k}_t}{\bar{k}_{t-1}} \right) + \frac{R_t}{\bar{N}_{t-1}} \right]. \tag{15}$$

Ignoring coefficients, the first term in brackets is the growth rate of output, the second term is the growth rate of technology, the third term is the growth rate of inputs to production at full employment, and the final term is the depreciation rate.

3.7.1 Productivity Shocks

Equation (15) shows the many channels through which increases in TFP can affect investment. First, higher TFP causes a proportional rise in output at full employment $\bar{\mathrm{Y}}$ and thus in investment, as businesses employ more capital per worker. Second, TFP could affect the cost of capital relative to the output price, v/p. For example, if higher productivity leads to lower inflation, the monetary authority could reduce real interest rates, cutting the real cost of capital and boosting investment. Third, faster TFP growth leads to more rapid growth in output ΔY, boosting investment. Finally, improved TFP increases the output of existing machines, reducing the need for investment. This is reflected in the second term in brackets, and offsets the third channel. The coefficient on this term is only θ, instead of 1, because businesses expect higher TFP to boost future output by $(1-\theta)\,\Delta A_t / A_t$.

How large is the impact of higher TFP growth ΔA on output growth ΔY? Consumption, currently 70 percent of U.S. GDP, should rise proportionately with changes in

TFP. Permanent income rises proportionately with TFP, assuming that TFP is a random walk apart from transitory business cycle effects, and total real wealth also rises nearly proportionately with TFP.[17] In equilibrium, the other components of GDP change so that overall non-investment GDP rises roughly proportionately to the change in TFP.[18]

Inspection of equation (15) reveals that the equilibrium solution for ΔY, given a one percent rise in TFP and a one percent rise in non-investment GDP, requires a one percent rise in investment. Overall GDP rises one percent, so $\Delta Y_t/Y_{t-1}$ and $\Delta A_t/A_{t-1}$ each rise by 0.01, leaving the expression in brackets unchanged. However, output per worker y rises one percent, resulting in a one percent rise in investment I. Adding the missing lag terms back into equation (15) would not fundamentally change this result.

The increased output of new machines in the case of the productivity shock eventually leads to further increases in output above the no-shock base case. Output is higher both because TFP is higher and because the average machine size increases. If capital were putty-putty, the full increase in average machine size would be immediate, because businesses would adjust the size of existing machines in order to get the capital-output ratio back up to base case levels. With putty-clay capital, however, the overall capital-labor ratio increases to its new higher equilibrium level only gradually.

Note that while higher productivity boosts output, it has no impact on the unemployment rate. Both output and productivity are one percent higher than in the base case. Employment is thus unchanged from the base case. In order to use productivity shocks to explain business cycle fluctuations in unemployment, those shocks must either be known to be temporary and induce intertemporal trade-offs between labor and leisure, as in Kydland and Prescott's (1982) real business cycle model, or have some impact on output beyond those impacts mentioned in this section.

3.7.2 Demand Shocks

Since a rise in TFP will boost both ΔY and ΔA in equation (15), a demand shock is defined as a change in output ΔY that occurs independently of the other terms in equation (15). For typical coefficient values, a one percent increase in demand will lead to a short-run increase in investment of more than one percent. This is just the acceleration principle, first expounded by Clark (1917). For equipment excluding computers, communications, mining, and farming equipment, the term in brackets

[17] The value of nonresidential capital rises proportionately, but the value of residential capital is unchanged.

[18] Tax revenues rise roughly proportionately to TFP. If these extra revenues are spent, the government component of GDP rises proportionately to TFP. If instead tax rates are reduced, consumption rises by a similar amount, in addition to the increase due to greater permanent income. The dollar adjusts to keep the trade deficit, which is a small percentage of total GDP, roughly unchanged.

has averaged about 0.082 over the past 30 years. However, the contemporaneous value of θ is 0.111. (Total θ for this category of investment is 0.31, but the effects are spread out over three years.) Thus, an extra 1.0 percent of demand results in an extra 1.35 percent (=0.111/0.082) investment in equipment excluding computers, mining, farming, and communications equipment.

3.7.3 Opposing Productivity and Demand Shocks

Now suppose that the shocks discussed in the previous two sections are combined, so that productivity rises by one percent at the same time that the ratio of demand to supply (output at full employment) falls by one percent, leaving ΔY unchanged. The sum of the two effects will be a decline in investment. Investment in equipment excluding computers, mining, farming, and communications equipment would fall 0.35 percent. In equation (15), the term in brackets, indicating the number of new machines, would decline 1.35 percent, while the size of new machines, captured by output at full employment $\bar{\mathrm{Y}}$, would rise 1.0 percent.

Another way to say this is that an increase in productivity that fails to be reflected in output growth will lead to a reduction of investment. For this reason, the surge in productivity during late 2001 and early 2002 failed to produce an increase in investment. The slow recovery from the 1990-1991 recession also featured a surge in productivity in 1992 that was unmet by stronger growth in output.

3.8 Differences with the Standard Neoclassical Model

In the standard neoclassical model with more than one type of capital, gross investment equals depreciation plus a polynomial distributed lag of the change in the desired stock of capital K*:

$$I_{m,t} = \sum_{i=0}^{T_m} \zeta_i \left(K^*_{m,t-i} - K^*_{m,t-i-1}\right) + Dep_{m,t},$$

where $\mathrm{Dep}_{m,t}$ is the real value of capital of type m depreciating in period t. If depreciation is geometric at rate δ_m, then $\mathrm{Dep}_{m,t}$=$\delta_m K_{m,t-1}$. The desired stock of capital is defined as

$$K^*_{m,t} = \alpha\ s_m\ \frac{p_t\ Y_t}{v_{m,t}}.$$

Combining the two equations,

$$I_{m,t} = \alpha\ s_m\ \sum_{i=0}^{T_m} \zeta_i \left(\frac{p_{t-i}\ Y_{t-i}}{v_{m,t-i}} - \frac{p_{t-i}\ Y_{t-i}}{v_{m,t-i}}\right) + Dep_{m,t}.$$

A crucial difference between putty-clay and putty-putty models is that in a putty-clay model, investment depends on $(p/v)\Delta Y$, while in a putty-putty model, investment depends on $\Delta(pY/v)$. That is, investment is much more sensitive to changes in the ratio of the price of output to the cost of capital p/v in a putty-putty model than in a putty-clay model. When capital is putty-clay, a change in the cost of capital only affects the capital/labor ratio of new capital. With a unit elasticity of substitution between capital and labor, a one percent fall in the cost of capital leads to a one percent rise in investment. However, when capital is putty-putty, firms can change the capital/labor ratio in both new and existing capital. With a unit elasticity of substitution between capital and labor, a one percent fall in the cost of capital leads to a one percent rise in the desired capital stock, which initially produces a rise in investment of much more than one percent.

Technology shocks also have a different short-run effect on investment in this paper than in the standard neoclassical model. In the putty-clay model, a proportionate rise in output and TFP leads firms to boost the capital/labor ratio of new capital but not of existing capital. Thus, investment rises proportionately to output and TFP. Over time, the higher rate of investment leads to further increases in output, generating further increases in investment. Eventually, both the capital stock and output increase more than proportionately to the rise in TFP.

In a putty-putty model, that long-run process happens much more quickly. The rise in the real wage resulting from higher technology leads firms to boost the capital/labor ratio of all capital. The desired stock of capital rises proportionately to output, so investment rises more than proportionately to output and TFP. (This assumes that the price of capital goods moves proportionately to the price of all output, so that p/v remains constant.) As in the putty-clay model, the higher rate of investment leads to further increases in output, generating further increases in investment. However, since putty-putty firms can change the capital/labor ratio of existing capital as well as new capital, the economy converges to its new long run equilibrium much more rapidly.

The different short-run effects of a rise in TFP lead to different drivers for the output accelerator in the two models. In a putty-clay model, the accelerator term depends on growth in output less growth in technology. However, in the standard model, the accelerator term depends on growth in output only.

The putty-clay assumption underlying the investment model in this paper produces two other differences with the standard neoclassical model of investment. First, the cost of capital in a putty-clay model depends on the rate of general price inflation rather than on the rate of price inflation for capital goods. The marginal product of an extra dollar of putty-clay capital rises over time with the price of the output that the worker using it produces. However, the marginal product of an extra dollar of putty-putty capital is always equal to the rental cost of new capital, and thus depends on the price of new capital. If the price of new capital will fall in the future, then the marginal product of capital will also fall.

Second, in a putty-putty model, real replacement demand equals the real value of currently-depreciating capital when it was purchased. However, in a putty-clay model, if the cost of labor has risen relative to the cost of capital, replacement machines will be of greater quality than those they replace. Thus, replacement demand typically exceeds depreciation.

Two other important differences between the model in this paper and the standard neoclassical model of investment stem from factors other than the difference between putty-putty and putty-clay. First, this paper assumes that the elasticity of substitution between individual types of capital may differ from unity. Many authors have made this assumption in a putty-putty model, Tevlin and Whelan (2000) being a recent example. However, those models generally assume that the elasticity of substitution between overall capital and labor also differs from unity. Kiley (2001) is a rare exception.

Second, I assume that firms base their capacity decision on both output and output at full employment, rather than on output alone. Because output at full employment depends on labor hours at full employment, that assumption means that growth in labor hours at full employment can affect investment. Similar to a growth model, faster growth in the workforce leads businesses to invest more. In the standard neoclassical model, businesses do not boost investment until those extra workers are producing output.

4 Data Used in the Estimation

Several sources of data were employed in the estimation. This section describes that data and those sources. In addition, this section discusses how data were constructed for labor hours at full employment, technology, output at full employment, and fitted cost shares of the various types of capital.

4.1 Sector Coverage

Ideally, the output and labor concepts match the investment series to be explained. For business fixed investment excluding that by the mining and farming sectors, output should be output of the nonfarm business sector, excluding housing and mining, plus output of households and institutions.[19] However, I have not subtracted mining output from the total. Similarly, labor hours should be total private labor hours

[19] The gross domestic product of households and institutions consists entirely of labor compensation. The imputed income from the capital used by such labor is included in nonfarm business output. Thus, the total output produced by the labor of households and institutions is split between the nonfarm business sector and households and institutions. This is the relevant output concept for investment in capital to be used by labor from households and institutions.

less labor hours in the mining and farming sectors. As in the case of output, I have excluded hours of the agricultural sector, but not of mining.

Because mining is much more capital-intensive than the rest of the economy, subtracting out that sector has a larger impact on investment than on the output and labor measures. Over the 1970-2001 period, investment in mining exploration, shafts and wells and in mining and oilfield machinery averaged 4.9 percent of business fixed investment outside farming. However, over that same period, GDP in the mining sector averaged 2.8 percent of private nonfarm GDP, while labor compensation in the mining sector averaged 1.5 percent of nonfarm private labor compensation. The higher capital intensity of mining is another reason to treat investment in that sector separately from other investment.

4.2 Labor Input

Labor hours at full employment are obtained by adjusting actual labor hours for cyclical movements. First, I explain as much of the movement in labor hours as possible using population variables, time trends and dummy variables, the ratio of output in nonfarm business plus households and institutions to total GDP excluding housing, and the difference between the unemployment rate and the estimate of the unemployment rate at full employment made by the Congressional Budget Office (CBO). Labor hours at full employment are then found by removing the effect of deviations in the unemployment rate from its full employment level.

Actual labor hours for nonfarm business, households, and non-profit institutions are obtained from the Bureau of Labor Statistics. The dependent variable in the regression is the logarithm of the ratio of that labor hours variable to the population aged 16 and over. (Using logarithms of labor hours means that a given change in the unemployment rate always has the same percentage effect on the participation rate.) Because participation rates for persons aged 16 to 24 are lower than for persons aged 25 and over, the logarithm ratio of the population aged 25 to 64 to the population aged 16 to 64 is added as an independent variable. Separate time trends for 1948-1979, 1980-1989, and 1990 on were used to capture trends in participation rates and hours per week. A dummy variable captures an apparent break in the participation rate in the second quarter of 1962.

The logarithm of the ratio of GDP for nonfarm business, households, and institutions to the ratio of non-housing GDP is added as an independent variable, because the population variables in the equation apply to overall labor hours, but we are only interested in labor hours in nonfarm business, households, and institutions. Finally, the equation includes the difference between the civilian unemployment rate and CBO's estimate of the civilian unemployment rate at full employment.

As expected, the unemployment has a strongly negative coefficient: all else equal, a higher unemployment rate causes labor hours worked to fall (Table 1). That coeffi-

Table 1: Impact of Unemployment on Labor Hours

Variable	Coefficient	T-statistic
Constant	-6.868	-278.6
$\log\left(\frac{\text{population aged 25 to 64}}{\text{population aged 16 to 64}}\right)$	0.175	3.9
Dummy beginning 1962:q2	-0.022	-6.8
Time	-0.0002	-0.9
Time trend beginning 1980:q1	0.0048	10.5
Time trend beginning 1990:q1	-0.0031	-6.4
$\log\left(\frac{\text{GDP in nonfarm business, households and institutions}}{\text{GDP excluding housing sector GDP}}\right)$	0.747	9.0
Unemployment rate: actual minus full employment	-0.0208	-39.4
R^2	0.963	
Durbin-Watson	0.44	

The dependent variable is the logarithm of hours worked in nonfarm business, households, and institutions divided by the population aged 16 and over.
Time equals 0 in 1900:q1 and rises by 0.25 each quarter. The sample is quarterly, from 1948:q2 to 2003:q1.

cient is nearly identical under different specifications for the population variable. For example, if the dependent variable uses population aged 16 to 64 instead of the population aged 16 and over, the coefficient on the unemployment rate remains -0.0208. Labor hours at full employment then equal actual labor hours divided by the exponential of the actual unemployment rate minus the full-employment unemployment rate, multiplied by its coefficient.

4.3 Capital Stock

In a putty-clay model, one can only construct an aggregate capital stock under special conditions, which are unlikely to be satisfied in the real world. Output at full employment of labor should be constructed using information on the size of existing machines of different vintages. As a simplification, however, I calculate output at full employment of labor using an aggregate capital stock.

The capital stock must be consistent with labor hours. Modelers typically assume that, in a putty-putty world, utilization rates of labor and capital are the same. Thus, the capital stock utilized at full employment equals the existing capital stock multiplied by its average historical utilization rate. In a putty-clay world, however, utilization rates for capital and labor do not move proportionately over the business cycle. The capital stock utilized at full employment need not equal the existing capital stock multiplied by its historical utilization rate.

Utilization rates for capital and labor move differently because the total capital stock reacts to demand shocks, whereas labor hours at full employment do not. Consider a series of positive shocks to demand that boost output above output at full employment. Define the utilization rate for labor as actual labor hours divided by full-employment labor hours and the utilization rate for capital as utilized machines divided by the total number of machines. In the very short run, the number of machines and full-employment labor hours are both fixed, so utilization rates for capital and labor rise by the same percentage amount. As time passes, businesses will invest in more plant and equipment than they otherwise would have, thereby trimming the rise in the utilization rate of capital. However, labor hours at full employment are unaffected, so there will be no offset to the rise in the utilization rate of labor.

The implication is that, in a putty-clay world, the capital stock multiplied by its average historical utilization rate is likely to be a poor proxy for capital utilized at full employment of labor. Late in a cyclical expansion, when investment is strong, the capital stock grows more rapidly than capital utilized at full employment of labor. TFP growth estimated using a putty-putty measure of capital will understate true TFP growth. Conversely, late in a recession or early in a recovery, when investment is weak, the capital stock grows more slowly than capital utilized at full employment of labor. TFP growth estimated using a putty-putty measure of capital will overstate true TFP growth.

To construct capital utilized at full employment, I start with the estimate of capital services CBO uses in constructing potential output in the nonfarm business sector. I then adjust that measure for changes in the stock due to the business cycle. For this purpose, I assume that businesses boost their capital stock by 0.40 percent for each 1 percent rise in the ratio of actual to full-employment labor hours, and that this adjustment occurs gradually over 8 quarters.[20] Thus, capital utilized at full employment is

$$\bar{K}_t = \frac{K_t}{0.60 + 0.40 * \left(\sum_{i=0}^{7} \frac{N_{t-i}}{\bar{N}_{t-i}}\right)/8},$$

where $\bar{K}$ is capital utilized at full employment, K is CBO's estimate of capital services, N is actual labor hours, and $\bar{N}$ is labor hours at full employment.[21]

With ex post fixed proportions, the ratio of capital utilized to capital utilized at full employment equals the ratio of labor hours utilized (worked) to labor hours at full employment.[22] Thus, utilized capital KU is

$$KU_t = \bar{K}_t \frac{N_t}{\bar{N}_t} = \frac{K_t}{0.60 + 0.40 * \left(\sum_{i=0}^{7} \frac{N_{t-i}}{\bar{N}_{t-i}}\right)/8} \frac{N_t}{\bar{N}_t}.$$

[20] That coefficient and lag length are loosely consistent both with the coefficients and lag lengths in the investment equations and with simple regressions of the growth rate of the total capital stock on labor hours and on the ratio of labor hours to full-employment labor hours.

[21] Technically, the right-hand side should be multiplied by the average historical utilization rate. For convenience, I have omitted that rate.

[22] This statement assumes that the mix of different vintages of capital used does not depend on the utilization rate of labor.

Utilized capital and actual labor hours are the capital and labor inputs in production.

Note that, in the short run, the amount of capital utilized is proportional to the number of labor hours rather than to the number of workers. Since capital per worker is fixed in the short run, a one percent increase in the number of workers and a one percent increase in hours per worker each increase machine-hours (sets of machines used times hours of use) by one percent. In the longer run, the historical downward trend in hours per worker has reduced capital per worker relative to capital per labor hour. My estimation procedure implicitly captures that trend as a reduction in TFP, but it is really a change in hours per machine.

4.4 Technology

In preparing its economic forecast, the Congressional Budget Office calculates estimates of potential output and total factor productivity.[23] Those estimates are designed to capture long-run trends in technology. However, short-run changes in technology are better indicators of the short-run movements in investment that an investment equation should capture. Thus, investment equations require an estimate of technology that focuses more closely on short-term changes.

As discussed above, total factor productivity is composed of TFP at full employment, or underlying technology A, and intensity of effort U. This paper assumes that intensity of effort is a function of current and lagged values of $Y/\bar{Y}$, the ratio of actual output to full-employment output. Initially, businesses meet a positive shock to demand by boosting both labor hours and worker effort. Over time, a greater share of the increase in output is met through additional labor hours, while effort reverts back to pre-shock levels. Consequently, the cyclical component of productivity depends positively on the contemporaneous ratio of actual to full-employment output, but negatively on lags of that ratio. Mathematically,

$$U_t = \left(\frac{Y_t}{\bar{Y}_t}\right)^{c_1} \left(\frac{Y_{t-1}}{\bar{Y}_{t-1}}\right)^{c_2} \left(\frac{Y_{t-2}}{\bar{Y}_{t-2}}\right)^{c_3}, \tag{16}$$

where c_1, c_2, and c_3 are coefficients, with $c_1>0$, $c_2<0$, and $c_3<0$.

Because neither the underlying technology nor the intensity of effort are observable, this paper employs a Kalman filter technique to estimate these. I assume that underlying technology is a random walk with exponential drift. In a sense, this procedure calculates the value of potential output that makes Okun's Law–the relationship between the unemployment rate and current and lagged ratios of actual GDP to potential GDP–fit in every period.

In order to estimate the Kalman filter, one must calculate TFP, the observable product of the unobservable underlying technology and intensity of effort. In this paper,

[23] For a discussion of the methodology employed, see Congressional Budget Office (2001).

I make the putty-putty assumption that output can be expressed as a function of aggregate labor and aggregate capital, i.e.,

$$Y_t = TFP_t \ KU_t^{\alpha} \ N_t^{1-\alpha}, \tag{17}$$

where KU and N are aggregate utilized capital and labor and TFP=A×U. Under constant returns to scale, the coefficients on capital and labor equal the shares of capital and labor in total income. This share is difficult to calculate in the noncorporate sector, because proprietors' income contains income of both capital and labor. In the corporate sector, labor compensation has accounted for an average 65 percent of corporate GDP since World War II. However, corporate GDP includes indirect business taxes. Some of these, mainly property taxes, are levied on capital, but most, like sales taxes, are levied on the combined output of capital and labor. If one excludes indirect taxes other than property taxes from corporate GDP, then labor's share of income jumps to 70 percent. Thus, I assume that α=0.30 and 1-α=0.70.

At full employment, equation (17) becomes

$$\bar{Y}_t = A_t \ \bar{K}_t^{\alpha} \ \bar{N}_t^{1-\alpha}. \tag{18}$$

Dividing equation (17) by equation (18), and remembering that with ex post fixed proportions the ratios of utilized capital to capital at full employment and utilized labor to labor at full employment are equal, yields

$$\frac{Y_t}{\bar{Y}_t} = U_t \left(\frac{KU_t}{\bar{K}_t}\right)^{\alpha} \left(\frac{N_t}{\bar{N}_t}\right)^{1-\alpha} = U_t \left(\frac{N_t}{\bar{N}_t}\right). \tag{19}$$

Substituting equation (19) into equation (16),

$$U_t = \left(U_t \frac{N_t}{\bar{N}_t}\right)^{c_1} \left(U_{t-1} \frac{N_{t-1}}{\bar{N}_{t-1}}\right)^{c_2} \left(U_{t-2} \frac{N_{t-2}}{\bar{N}_{t-2}}\right)^{c_3}.$$

But U=TFP/A, so

$$\frac{TFP_t}{A_t} = \left(\frac{TFP_t}{A_t}\frac{N_t}{\bar{N}_t}\right)^{c_1} \left(\frac{TFP_{t-1}}{A_{t-1}}\frac{N_{t-1}}{\bar{N}_{t-1}}\right)^{c_2} \left(\frac{TFP_{t-2}}{A_{t-2}}\frac{N_{t-2}}{\bar{N}_{t-2}}\right)^{c_3}. \tag{20}$$

The state variable (SV) is the logarithm of technology less its time trend

$$SV_t = \log(A_t) - c_4 t, \tag{21}$$

where c_4 is a coefficient to be estimated. Taking logarithms of equation (20), substituting for $\log(A_{t-i})$ using equation (21), and rearranging terms produces the signal equation

$$\begin{aligned} \log(TFP_t) \quad = \quad & SV_t + c_4 t + \frac{c_1}{1-c_1} \log\left(\frac{N_t}{\bar{N}_t}\right) \\ & + \frac{c_2}{1-c_1}\left[\log\left(TFP_{t-1}\frac{N_{t-1}}{\bar{N}_{t-1}}\right) - SV_{t-1} - c_4(t-1)\right] \\ & + \frac{c_3}{1-c_1}\left[\log\left(TFP_{t-2}\frac{N_{t-2}}{\bar{N}_{t-2}}\right) - SV_{t-2} - c_4(t-2)\right]. \end{aligned} \tag{22}$$

Table 2: Kalman Filter Estimate of Total Factor Productivity

	Unconstrained		Constrained	
Variable	Estimate	z-statistic	Estimate	z-statistic
c_1	0.319	8.2	0.292	8.0
c_2	-0.164	4.8	-0.183	6.0
c_3	-0.082	2.8	-0.109	
c_4	0.0134	6.0	0.0134	6.1
Standard Error	0.0079		0.0079	
Log Likelihood	696.4		695.8	

Time equals 0 in 1900:q1 and rises by 0.25 each quarter. The sample is quarterly, from 1950:q2 to 2003:q1.

Thus, the state variable SV is chosen to produce levels of effort U and output at full employment $\bar{Y}$ that make equation (16), describing the cyclical behavior of effort, fit exactly.

The signal equation (22) was estimated both with and without the constraint that $c_1+c_2+c_3=0$. Data are quarterly, so t rises by 0.25 per quarter, and t-x is replaced by t-0.25*x. Table 2 shows results of the estimation. If $c_1+c_2+c_3=0$, then fluctuations in labor utilization have no permanent effect on effort. That constraint is not rejected at the 90 percent level. Without the constraint, about 23 percent of the contemporaneous rise in effort resulting from a rise in demand remains after two quarters.

The time series for technology A can then be found by creating a smoothed estimate of the state variable SV and reversing equation (21), using the estimated value of c_4. Output at full employment is calculated by solving equation (18).

By construction, the ratio of actual GDP to GDP at full employment for nonfarm business, households and institutions closely mirrors the behavior of the actual unemployment rate minus the unemployment rate at full employment (see Figure 2). The fact that the scale for GDP is roughly twice that of the scale for the unemployment rate indicates that a one percent rise in GDP relative to its full-employment level reduces the unemployment rate relative to full employment by about a half a percentage point.

The general pattern of the ratio of actual GDP to GDP at full employment in the nonfarm business, household and institutional sectors is very similar to the ratio of actual GDP to CBO's measure of potential GDP for the same sectors (see Figure 3). However, GDP at full employment is much more volatile than CBO's estimate of potential GDP (see Figure 4). To explain quarterly movements in investment, one must have a measure of TFP that captures that volatility.[24]

[24] Even after adjusting for the business cycle, there still seems to be some cyclicality left in TFP at full employment. It may be due to simultaneity, with TFP growth influencing the ratio of actual

4.5 Data for the Cost of Capital

A time series for the cost of capital must be constructed for each of the five categories of investment examined in this paper: computers; software; communications equipment; other private nonresidential equipment, excluding mining and farming; and private nonresidential structures, excluding mining and farming. In addition to data for the price deflators, which is constructed simultaneously with real investment, those calculations require data for the cost of funds, expected inflation, service lives, and the tax treatment of capital.

4.5.1 Cost of Funds and Expected Inflation

Conceptually, the nominal cost of funds ${}_tr$ is the rate of expected payments by business to capital (including retained earnings), net of taxes, plus the expected rate of revaluation of assets due to inflation. The real cost of funds is the nominal cost of funds less the expected aggregate rate of inflation. In a putty-clay world with no TFP growth, assets are revalued according to the change in the value of what they produce, i.e., the aggregate rate of inflation. Thus, the expected revaluation of assets due to inflation and the expected rate of aggregate inflation are the same, and the real cost of funds, or ${}_tr - {}_t\dot{p}$ in equation (9), is just expected net payments to capital.[25] Taking account of TFP growth would have an ambiguous effect, because it means the value of output from a given piece of capital rises faster than inflation, but also means that the cost of the associated labor input rises faster than inflation.

For corporations, the real cost of funds is a weighted average of the after-tax yield on debt and the after-tax yield on equity, excluding the revaluation of assets for inflation. I assume that the cost of funds for all other legal forms of organization is the same as the cost of funds for corporations. The weights on debt and equity are constructed from data in the Flow of Funds accounts of the Federal Reserve Board. Debt is the value of credit market instruments of nonfarm nonfinancial business, while equity is the market value of equities of nonfarm nonfinancial business. I use two-quarter moving averages because the flow of funds data are for the end of each quarter.

The after-tax yield on debt is calculated as Moody's corporate bond rate times one minus the federal corporate income tax rate times one minus the state corporate tax rate.[26] The after-tax yield on equity, excluding capital gains from inflation, is

GDP to GDP at full employment, or to an inadequate estimate of the impact of the business cycle.

[25] The equality of the real rate of return with expected payments to capital net of taxes is consistent with the Bureau of Labor Statistics's methodology for creating the cost of capital, as described in Appendix C of U.S. Department of Labor (1983). Similarly, Harper, Berndt, and Wood (1989) show that using an "internal own rate of return model" defined using expected payments to capital net of taxes "yields the same rental prices ... as would the nominal internal rate of return model provided average capital gains were employed." The putty-clay model meets this condition.

[26] This formula is used, instead of one minus the sum of the two corporate tax rates, because taxes paid to the states are deductible at the federal level. The Moody's corporate bond rate series is

calculated using the dividend-discount model, and thus equals the dividend yield for the nonfarm nonfinancial business sector divided by the average ratio of dividends to after-tax profits in that sector during 1952 to 2001, 0.556.

Expected inflation, modelled using adaptive expectations, is calculated as a polynomial distributed lag of rates of growth in the price index for GDP. That differs from the capital gains term in a putty-putty model, which depends on the change in the price of existing assets, rather than on the change in price of the output they produce. Using a different measure of expected inflation would have little effect on the results: in a putty-clay model, expected inflation only affects the present discounted value of depreciation allowances.

At least for debt, the real cost of funds as calculated above applies only to the period to maturity of securities currently being issued. Beyond this, businesses could reasonably assume that the real cost of funds will return to its historical average. Consequently, the real cost of funds for communications equipment and for other private nonresidential equipment, excluding mining and farming, is an average of the real cost of funds as calculated above and its historical average. For private nonresidential structures, excluding mining and farming, the weights are 0.10 on the real cost of funds as calculated above and 0.90 on its historical average. Putting a much larger weight on the current real cost of funds significantly worsens the fit of the equation for investment in structures.

4.5.2 Service Lives

Service lives serve two purposes in the model presented in this paper. First, the service life of new capital contributes to a firm's calculation of the cost of that capital. Second, the service life of existing capital determines how much of the existing stock of capital must be replaced during a particular period, and thus determines the replacement portion of investment. Service lives of the investment categories may change over time, either because service life of an individual type of asset changes or because the mix of investment shifts toward assets with longer or shorter lives. Consequently, the service life used in the cost of capital may not be the same as the service life used to determine replacement demand.

Service lives for each asset are taken from the U.S. Department of Labor, Bureau of Labor Statistics (2001). In addition, U.S. Department of Labor, Bureau of Labor Statistics (1983) makes an adjustment to account for the decline in the efficiency of existing capital as it ages. That adjustment is equivalent to setting parameter db_m in equation (9) to a value less than 1. However, I found that such an adjustment worsens the fit of four of the five investment equations, so I left it out.

For aggregates of assets with different service lives, those lives are weighted together

extended before 1970 by averaging the rates on Aaa-rated bonds and Baa-rated bonds.

using investment shares, rather than shares of the capital stock. That choice is fairly obvious for the service life used in calculating the cost of capital, since that cost is ultimately used in an investment equation, not in an equation for the desired capital stock. That choice is also the right one for replacement demand, since shares of replacement demand are likely to be more closely correlated with shares of investment than with shares of the capital stock.

To avoid possible simultaneity problems and to adjust for the volatility of investment shares, weights are calculated over an eight-year period, rather than for the period contemporaneous with investment. Replacement demand is determined by depreciation of capital installed in the past, so I use simple eight-year moving averages of investment shares. (Eight years is one of the shortest service lives in the BLS data set.) For the cost of capital, which depends on service lives of capital currently being installed, I use eight-year centered moving averages. To calculate those averages for the final four years of history, I assume that investment shares in the future are the same as in the last quarter of available history.

4.5.3 Tax Treatment of Capital

The raw data on the tax treatment of depreciation—tax lifetimes, methods of depreciation used, and declining balance parameters—and on investment credit rates is drawn from Gravelle (1994). The federal tax rate on corporate income is the statutory rate. The state and local tax rate on corporate income is obtained from the National Income and Product Accounts by dividing state and local corporate tax receipts by corporate profits less profits earned by the Federal Reserve less net rest-of-world profits.

For aggregate categories of investment, rates of investment tax credit and tax lifetimes of assets with different lifetimes are both weighted together using the same eight-year centered moving averages used to weight together service lives. Methods of depreciation used and basis adjustment for the investment tax credit do not need to be weighted, since they are the same across all assets within each category of investment that I estimate.

Methods of depreciation include straight-line, accelerated (using declining balance), sum of digits, and expensing. I assume that nonprofit institutions and tax-exempt cooperatives use expensing. Using expensing for those entities when calculating the cost of capital is mathematically equivalent to setting the corporate tax rate for them to zero.

4.6 Fitted Cost Shares

In order to estimate equation (12), one needs estimates of the cost shares $s_{m,t-i,0}$ of new capital. Using a time series of the actual cost shares would introduce a simultaneity bias, since an error in any single investment equation would also result in a roughly proportionate error in the cost share for that asset. However, changes in the relative costs of different types of capital or shifts in technology may result in changes in the cost shares that should be captured in the investment equation. The solution is to use fitted cost shares in the investment equations. Doing so eliminates the potential for simultaneity while at the same time allowing changes in relative costs of capital to affect investment in different types of capital.

To obtain a raw estimate of the cost share of new capital of type m, I divide the total cost of new capital of type m by a rough estimate of the total output of workers that will use that new capital. The total output of workers using that new capital is the sum of expansion demand and replacement demand for type-m capital. Expansion capacity is modeled as a weighted average of annualized growth in actual output and output at full employment less TFP during the prior three years. The weights on growth in actual output and output at full employment are round estimates of the coefficients on growth in actual output and output at full employment in the investment equations. The output of replacement machines is modeled as the depreciation rate times output at full employment. The raw estimates of cost shares of new capital are normalized to sum to one.

Kiley (2001) estimates cost shares using a translog cost function. However, for the investment categories examined in this paper, an even broader array of functional forms is necessary. In particular, visual inspection of the cost share for software suggests that a logit would work better for that share (see Figure 5).

I estimate the cost shares for both computers and software as logit functions. The upper bounds are set roughly equal to 150 percent of the historical maximum cost share. For computers, the relative cost variable is a weighted average of the logarithms of ratios of the cost of capital for computers to the costs of non-computer non-software capital. The weights are historical averages of cost shares for those non-computer non-software types of capital. In addition, the cost share for computers moved noticeably higher in the late 1970s and early 1980s, roughly coincident with the introduction of the personal computer. Consequently, I included a dummy variable (dummy_1) equal to zero before 1979, rising linearly to 1.0 by the end of 1982, and equal to 1.0 thereafter. Estimated coefficients for the fitted cost shares are shown in Table 3.

The relative cost variable for software is similar to that for computers, but uses the cost of capital for software instead of the cost of capital for computers. I also included a linear time trend. The time trend captures most of the rise in the share.

The cost share for structures is modeled as a logit function of the remaining cost

Table 3: Estimation Results for Fitted Cost Shares

Computers
$\log\left(\frac{s_1}{0.10-s_1}\right) = -0.102*\left[0.461*\log\left(\frac{v_1}{v_5}\right)+0.468*\log\left(\frac{v_1}{v_3}\right)+0.071*\log\left(\frac{v_1}{v_4}\right)\right]$
(-5.5)
$+1.17*dummy_1 - 0.97$ $\quad R^2 = 0.918, \quad D.W. = 0.24$
(13.7) (-9.3)
Software
$\log\left(\frac{s_2}{0.17-s_2}\right) = -0.429*\left[0.461*\log\left(\frac{v_2}{v_5}\right)+0.468*\log\left(\frac{v_2}{v_3}\right)+0.071*\log\left(\frac{v_2}{v_4}\right)\right]$
(-2.6)
$+0.081*time - 7.47$ $\quad R^2 = 0.985, \quad D.W. = 0.17$
(14.0) (-10.3)
Structures
$\log\left(\frac{\frac{s_3}{s_3+s_4+s_5}-0.30}{0.63-\frac{s_3}{s_3+s_4+s_5}}\right) = -0.235*\left[0.490*\log\left(\frac{v_3}{v_1}\right)+0.510*\log\left(\frac{v_3}{v_2}\right)\right] -0.51$
(-9.0) (-7.4)
$R^2 = 0.362, \quad D.W. = 0.11$
Telecommunications Equipment
$\frac{s_4}{s_4+s_5} = -0.006*\log\left(\frac{v_5}{v_4}\right)+0.034*dummy_4+0.114$
(-1.3) (10.8) (35.2)
$R^2 = 0.750, \quad D.W. = 0.31$
T-statistics are in parentheses below estimated coefficients.
Suffixes: 1 denotes computers; 2 denotes software; 3 denotes nonresidential structures; 4 denotes communications equipment; 5 denotes other equipment.
Time equals 0 in 1900:q1 and rises by 0.25 each quarter. Dummy variables are defined in the text. The sample is quarterly, from 1967:q1 to 2003:q1.

shares, with the range of the logit set equal to twice the range of structures' share of remaining costs. The relative cost variable is a weighted sum of the logarithms of the ratios of cost of capital for computers and software to that for structures. Again, the weights are historical averages of cost shares. The negative coefficient on this variable suggests that businesses have substituted software and computers for structures.

The cost share for telecommunications equipment is modeled as a share of remaining capital costs. That cost share, like the cost share of computers, rose sharply in the late 1970s and early 1980s. Hence, I add a dummy variable ($dummy_4$) equal to 0 before 1980 and equal to 1 thereafter. The logarithm of the ratio of the cost share of telecommunications equipment to that for "other" equipment (equipment excluding computers, telecommunications, mining and farming) has a small effect. The cost share for "other" equipment is then calculated as a residual.

5 Estimates of Investment Equations

This paper estimates investment equations for five categories of capital: computers; software; communications equipment; other private nonresidential equipment, excluding mining and farming; and private nonresidential structures, excluding mining and farming. Adding an autoregressive correction has little impact on the estimates for most categories of capital but significantly improves the theoretical properties of the software equation. Excluding the last ten years from the estimation period has little effect on the fitted values of overall investment during that period, although the fit for individual equations suffers. Tests fail to show an independent role for productivity growth beyond that in the model.

5.1 Equation to Be Estimated

For each type of capital m, the investment equation to be estimated is

$$I_{m,t} = \alpha \sum_{i=0}^{T_m} \beta_{m,i} \; s_{m,t-i,0} \frac{p_{t-i}}{v'_{m,t-i}} \times \sum_{i=0}^{\tau_m} \zeta_{m,i} \; X_{m,t-i} + \epsilon_{m,t}, \tag{23}$$

where

$$\begin{aligned} X_{m,t-i} &= 4 \cdot \theta_m \left(Y_{t-i} - Y_{t-i-1}\right) + 4 \cdot (1-\theta_m) \left(\bar{Y}_{t-i} - \bar{Y}_{t-i-1}\right) \\ &\quad + \frac{db_m}{L_m} \; \bar{Y}_{t-i} - 4 \cdot \left(\frac{A_{t-i} - A_{t-i-1}}{A_{t-i}}\right) \left(1 - \frac{db_m}{4 \cdot L_m}\right) \bar{Y}_{t-i}, \end{aligned} \tag{24}$$

$\tau_m \geq T_m$, $\sum_i \beta_{m,i} = \sum_i \zeta_{m,i} = 1$ for each type of machine m, and ϵ is an error term. Consistent with theory, equation (23) is estimated without a constant term.

Equation (24), capturing the non-price drivers of investment, is equation (13) with some modifications. First, because I use annualized quarterly data, I annualize changes in actual output, potential output, and underlying technology by multiplying by 4. Second, I assume that the parameter η equals 1. Third, the depreciation rate times output at full employment replaces the number of replacement machines times output per full employment worker. Finally, output at full employment replaces the number of full employment workers times output per worker.

The error term ϵ_m in equation (23) is likely to be correlated with nominal output at full employment $p\bar{Y}$ divided by the rental cost of type-m capital. As a correction for heteroskedasticity, I multiply each side of equation (23) by the lagged ratio of the price of m to nominal output at full employment. Thus, the dependent variable is very nearly nominal investment as a share of nominal output at full employment. The cost shares s are the fitted shares as described in the previous section. I assume db_m=1 for all m.

The parameters to be estimated are α, the $\beta_{m,i}$, the $\zeta_{m,i}$, and the θ_m. I do not impose the restriction that α is the same across all types of capital, but it nearly is. Typically, the lag structure in investment equations is specified as a polynomial distributed lag.[27] However, the necessity of estimating θ simultaneously with the β and ζ complicates the identification problem.

Instead, for computers, software, and telecommunications equipment, I impose a moving average on the output and cost of capital terms in the investment equation, setting $\beta_{m,i} = \beta_{m,1}$ and $\zeta_{m,i} = \zeta_{m,1}$ for all i and choosing the T_m and τ_m that produce the smallest standard error for the equation. For the other two categories of investment, I use two separate moving averages of the output term X (two different values of ζ) but a single moving average of the cost term (a single β). Although one could employ a separate θ_m for the two moving averages of output, this produces odd results, so I do not.[28]

5.2 Estimation Results: Ordinary Least Squares

For the most part, the results from estimating equation (23) for the five different types of capital are consistent with theory (see Table 4). Lags on output are longer than those on the cost of capital. The estimated coefficient on capital's share of income, 0.257 to 0.269, is somewhat lower than the 0.30 share I assume in calculating TFP, but there are good reasons for that. First, I excluded mining investment when calculating cost shares, but did not exclude mining from the output measure used in the equations. Second, capital's share of income is higher in the mining sector than in the overall corporate sector, so capital's share of income in the corporate sector excluding mining is lower than that in the overall corporate sector.

The consistency of estimates of the coefficient θ with theory is more mixed. At the time an investment is being planned, the coefficient θ reflects the share of the contemporaneous difference between actual output and output at full employment that businesses think will still exist when the investment is put in place. Thus, types of capital with short planning cycles, i.e., those with the shortest lags on output, should also have the highest θ. And so, computers, with the shortest lag on output, also has the highest estimated coefficient for θ, 0.628. However, the asset with the next-shortest time-to-build, software, actually has a negative θ. Structures, the asset with the longest time-to-build, has the highest estimated θ of the remaining asset types.

[27] See, for example, Bischoff (1971), Hall and Jorgenson (1971), or Laurence H. Meyer & Associates, Ltd. (1995).

[28] For other equipment, the estimate of θ for recent growth in demand is smaller than the estimate of θ for more distant lags of growth in demand.

Table 4: Least Squares Results for Investment Equations

	Computers	Software	Telecom. Equipment	Other Equipment	Structures
Lags on Cost of Capital	0 to 1	1 to 4	0 to 1	0 to 6	0 to 4
Lags on Output	1 to 5	2 to 7	1 to 10	0 to 7	1 to 12
$\sum \zeta_{m,i}$	1	1	1	0.808	0.613
(t-statistic)				(23.2)	(16.0)
Lags on Output				8 to 12	13 to 24
$\sum \zeta_{m,i}$				0.192	0.387
α	0.259	0.263	0.259	0.257	0.269
(t-statistic)	(124.3)	(206.2)	(134.0)	(250.6)	(116.3)
θ_m	0.628	-0.177	0.232	0.227	0.281
(t-statistic)	(8.1)	(-2.1)	(6.0)	(9.6)	(7.2)
R^2	0.944	0.990	0.769	0.797	0.714
D.W.	0.34	0.11	0.28	0.38	0.11
Sample Period	1967:q4-2003:q1	1968:q2-2003:q1	1967:q2-2003:q1	1969:q4-2003:q1	1970:q3-2003:q1

"Other equipment" is equipment excluding computers, communication equipment, farm tractors, agricultural machinery except tractors, and mining and oilfield machinery. "Structures" is private nonresidential structures excluding farms and mining exploration, shafts, and wells.
Regressions were run with ordinary least squares on quarterly data.

5.3 Estimation Results: Autocorrelation Correction

The very low Durbin-Watson statistics in the regressions using ordinary least squares (OLS) suggests autocorrelation of the residuals. To test the sensitivity of the OLS estimates to this autocorrelation, I applied an AR(1) correction to the equations. In the case of software, the estimated AR(1) process is nearly nonstationary (i.e., the estimated autoregressive coefficient is close to 1.0), so I used an AR(2). In all cases except software, I retained the same lag structure. For software, I added the first lag of output.[29]

Results are shown in Table 5. Except for structures, for which the estimated AR(1) process is nearly nonstationary, estimates of α are similar to those for OLS. The autocorrelation correction also raises the coefficient θ for software to a value more consistent with estimates of θ for other categories of investment.

[29] While adding the first lag of output to the software equation raises the R^2 only slightly, it removes the need for a moving average adjustment, and it makes the estimated α the same as in the OLS equation.

Table 5: Investment Equations with Autoregressive Correction

	Computers	Software	Telecom. Equipment	Other Equipment	Structures
Lags on Cost of Capital	0 to 1	1 to 4	0 to 1	0 to 6	0 to 4
Lags on Output	1 to 5	1 to 7	1 to 10	0 to 7	1 to 12
$\sum \zeta_{m,i}$	1	1	1	0.727	0.678
(t-statistic)				(22.9)	(11.5)
Lags on Output				8 to 12	13 to 24
$\sum \zeta_{m,i}$				0.273	0.322
α	0.258	0.262	0.257	0.255	0.133
(t-statistic)	(34.4)	(43.4)	(38.3)	(86.0)	(4.6)
θ_m	0.427	0.290	0.236	0.306	0.452
(t-statistic)	(4.4)	(3.6)	(3.9)	(8.0)	(3.9)
AR(1) coefficient	0.85	*	0.86	0.83	0.99
R^2	0.983	0.999	0.940	0.936	0.972
Sample Period	1968:q1-2003:q1	1968:q3-2003:q1	1967:q3-2003:q1	1970:q1-2003:q1	1970:q4-2003:q1

"Other equipment" and "Structures" are defined in the notes to Table 4.

Regressions were run with ordinary least squares on quarterly data.

* For software, the AR(1) coefficient is 1.45, and the AR(2) coefficient is -0.50.

5.4 Test of Productivity Effect

Section 3 shows that a positive shock to productivity unaccompanied by higher output leads businesses to reduce investment. To test this, I altered equation (24), the equation for the number of new machines, to

$$\begin{aligned} X_{m,t-i} &= 4 \cdot \theta_m \left(Y_{t-i} - Y_{t-i-1}\right) + 4 \cdot (1-\theta_m)\left(\bar{Y}_{t-i} - \bar{Y}_{t-i-1}\right) \\ &\quad + \frac{db_m}{L_m}\,\bar{Y}_{t-i} - 4 \cdot (1-\psi_m)\left(\frac{A_{t-i} - A_{t-i-1}}{A_{t-i}}\right)\left(1 - \frac{db_m}{4 \cdot L_m}\right)\bar{Y}_{t-i} \\ &\quad - 4 \cdot \psi_m \Delta\tilde{A}\left(1 - \frac{db_m}{4 \cdot L_m}\right)\bar{Y}_{t-i}, \end{aligned}$$

where $\Delta\tilde{A}$ is the sample average of $\left(\frac{A_{t-i} - A_{t-i-1}}{A_{t-i}}\right)$.

Under the null hypothesis that the accelerator effect depends on differential growth between actual output and output at full employment calculated with actual technology and that growth in TFP does not aid investment through some channel outside the model of this paper, the parameter ψ is equal to zero for all types of machines m. One alternative hypothesis would be that the accelerator effect depends on differential growth between actual output and output at full employment calculated with trend TFP. Under that hypothesis, $\psi_m = \theta_m$. Another alternative hypothesis would be that not only does the accelerator effect depend on trend TFP, but also that growth in TFP aids investment directly through some channel outside the model. In that case, $\psi_m > \theta_m$.

Estimates of the parameter ψ are actually negative in most cases, supporting the null hypothesis that the accelerator effect should be calculated using actual technology and that productivity growth aids investment only through channels already captured in the model. Table 6 reports results with autoregressive corrections in cases where the roots were not close to unity, and OLS results otherwise.[30] The coefficient ψ is positive (but not statistically significant) only for other equipment, and is significantly negative for software.

5.5 Explaining Past Investment

The estimated equations do a reasonably good job of explaining private nonresidential fixed investment, excluding mining and farming, over the past 30 years (Figure 6). For equipment and software, I use the equations fitted with autocorrelation corrections. Collectively, those equations fit about as well as the equations estimated with ordinary least squares, but have better theoretical properties. For structures, the equation fitted with an AR(1) correction makes no sense because the AR(1) process is nearly nonstationary, so I use the equation estimated with ordinary least squares. The equations generally capture the timing and size of movements in investment, including the boom of the 1990s. Investment exceeded fitted values most notably in 1971 and in 1985-1986, and fell short of fitted values in 1994 and again in 2001.

Equations for investment allow us to assess the contributions of various factors to past fluctuations in investment. This subsection explores the role of three key factors over the last 30 years: demand shocks, or the business cycle; the real cost of funds; and growth in labor hours at full employment.

5.5.1 Demand Shocks

To isolate the effects of demand shocks on investment, I created an alternative measure of output growth equal to growth in technology plus the average difference in

[30] Using OLS, the estimated value of ψ (with t-statistic) is -0.307 (-2.2) for computers, 0.202 (1.1) for software, -0.192 (-2.3) for telecommunications equipment, and -0.096 (-2.4) for other equipment. Using AR1, the estimated value of ψ (with t-statistic) is -0.441 (-2.3) for structures.

Table 6: Investment Equations with Test of the Effect of Productivity

	Computers	Software	Telecom. Equipment	Other Equipment	Structures
Lags on Cost of Capital	0 to 1	1 to 4	0 to 1	0 to 6	0 to 4
Lags on Output	1 to 5	1 to 7	1 to 10	0 to 7	1 to 12
$\sum \zeta_{m,i}$	1	1	1	0.725	0.604
(t-statistic)				(22.8)	(16.3)
Lags on Output				8 to 12	13 to 24
$\sum \zeta_{m,i}$				0.275	0.396
α	0.258	0.261	0.257	0.255	0.269
(t-statistic)	(34.7)	(40.1)	(38.5)	(79.6)	(116.8)
θ_m	0.425	0.253	0.234	0.309	0.289
(t-statistic)	(4.3)	(3.2)	(3.8)	(7.9)	(7.3)
ψ_m	-0.037	-0.051	-0.049	0.057	-0.074
(t-statistic)	(-0.3)	(-3.1)	(-0.7)	(1.1)	(-1.1)
AR correction	AR(1)	AR(2)	AR(1)	AR(1)	none
R^2	0.983	0.999	0.940	0.937	0.717
Sample Period	1968:q1-2003:q1	1968:q3-2003:q1	1967:q2-2003:q1	1970:q1-2003:q1	1970:q3-2003:q1

"Other equipment" and "Structures" are defined in the notes to Table 4.

growth between output at full employment and technology.[31] Essentially, I set the first two terms in equation (15) equal to their sample average. The estimated effect of demand shocks on investment is then the difference between fitted values of investment using growth in actual output and fitted values of investment using growth in the alternative measure of output (see Figure 7).

The fall in demand accompanying the recent recession has been quite large, despite the small decline in GDP compared to declines in GDP in other recessions. The reason is that output has grown slowly despite very strong growth in TFP. I interpret this as evidence of negative demand shocks. Even so, the rise and fall of investment over the past decade was so large that removing fluctuations in investment caused by demand shocks eliminates a smaller fraction of that cycle in investment than of past cycles of investment (see Figure 8).

[31] Demand shocks are measured as fluctuations in output that exceed fluctuations in technology. Thus, a moderate rise in output during a period when underlying technology is falling rapidly would count as a positive demand shock. A sharp rise in output that only matches a sharp rise in underlying technology would imply the absence of a demand shock.

5.5.2 The Real Cost of Funds

To assess the role of fluctuations in the real cost of funds on investment, I started with the alternative measure of output growth described above, and then held the real cost of funds at its average value from 1967 to early 2003 of 5.4 percent. The impact of fluctuations in the real cost of funds on investment is then the difference between the fitted values of investment described in the previous section and those fitted values as recalculated with the real cost of funds at 5.4 percent (see Figure 9).

A low real cost of funds played a significant role in the investment boom of the 1990s, and also boosted investment during the early 1970s, while a high real cost of funds held investment down during 1979 and the first half of the 1980s. Together, demand shocks and fluctuations in the real cost of funds explain two-thirds of the rise in investment's share of potential GDP during the 1990s boom and nearly three-quarters of the fall in investment's share of potential GDP during the bust of 2000-2003 (see Figure 10). However, they leave unexplained a surge in investment from 1979 to 1986.

5.5.3 Demographics

Demographics, captured by growth in labor hours at full employment, play an important role in a putty-clay model of investment. Firms invest not only to meet cyclical fluctuations in output but to meet growth in output at full employment. Growth in labor hours at full employment is an important driver of that growth. In fact, absent cyclical shocks, the number of expansion machines equals net growth in workers. Thus, the more rapid the growth in labor hours, the more businesses will invest. One can think of this either as investing to help produce the higher output expected to be supplied by that additional labor or as taking advantage of the greater supply of labor to boost output. Growth models, e.g., Solow (1956), often assume that, for a given interest rate, investment depends on growth in labor hours.

To remove the effects of fluctuations in full-employment labor hours, I replaced labor hours at full employment with an alternative series that grows at the historical average measured from 1967 to the first quarter of 2003. The alternative series for output at full employment is found by multiplying output at full employment by the ratio of the alternative hours series to labor hours at full employment. The impact of demographics on investment is then the difference between the fitted values described in the previous section less those values recalculated using the alternative series for output at full employment. Both sets of fitted values eliminate the effects of demand shocks and fluctuations in the real cost of funds from its sample average.

Strong growth in labor hours made an important contribution to investment in the late 1970s and early 1980s (see Figure 11). On the other hand, slow growth in labor hours, after adjustment for normal responses to the unemployment rate, exacerbated downturns in investment during the 1990-1991 recession and during the recent recession. In fact, from 2000 to early 2003, slow growth in hours caused nearly as large a drop in investment as the fall in demand.

One can think of slow growth in labor hours from 2000 to 2003 as reducing the usual partial offset to a negative accelerator effect. Normally, when an adverse shock to demand causes output to slow, businesses assume that this shock will eventually fade. So, at the same time that lower demand pushes output down, the prospect for future output growth improves, because businesses assume the economy will move back toward full employment. Those improved prospects for future growth are thus a partial offset to the accelerator effect of lower demand, which would otherwise be much larger. Between 2000 and 2003, the gap between actual output and output at full employment widened much less than it normally does during a cyclical downturn, reducing the size of that offset.

To see this mathematically, rearrange equation (15) as

$$I_t = \alpha \frac{p_t}{v_t'} \bar{Y}_t \times \left[\frac{\Delta Y_t}{Y_{t-1}} - \frac{\Delta A_t}{A_{t-1}} + (\theta - 1) \left(\frac{\Delta Y_t}{Y_{t-1}} - \frac{\Delta \bar{Y}_t}{\bar{Y}_{t-1}} \right) + \frac{R_t}{\bar{N}_t} \right].$$

The first two terms in brackets, the difference in growth between output and technology, are the full accelerator effect. The third term in brackets, the difference in growth between actual output and output at full employment multiplied by (θ-1), is the offset to the full accelerator effect. The full accelerator effect shows what the impact of a slowdown in growth would be if not partially offset by expectations of more rapid growth in the future.

Table (7) shows that the net accelerator effect was much larger during the 2000-2003 period than during earlier downturns in investment. The full accelerator effect, the drop in investment during 2000-2003 that would have occured if businesses assumed that changes in the ratio of output to technology were permanent, was similar to that in prior recessions. However, the offset from the difference in growth between actual output and output at full employment was much smaller during the 2000-2003 period than during earlier downturns in investment. Slow growth in labor hours at full employment meant that a smaller gap than normal opened between actual output and output at full employment, meaning a smaller anticipated cyclical rebound in growth of output. The smaller increase in the output gap is reflected in a smaller rise in the unemployment rate than during previous recessions.

5.6 Out-of-Sample Properties

One good test of the model is to see whether it can explain the boom and bust in investment over the past decade when data for that period is excluded from the sample. To conduct that test, I repeated the procedures discussed in Section 4,

Table 7: Accelerator Effects on Business Fixed Investment (Percent of CBO's July 2003 Estimate of Potential GDP)

Peak Trough	1973:q3 1975:q4	1979:q3 1983:q1	1989:q3 1992:q1	2000:q2 2003:q1
Change in BFI's Share of Nominal Potential GDP	-1.8	-1.9	-1.9	-3.2
Net Accelerator Effect	-1.3	-1.6	-1.9	-2.9
Full Accelerator Effect	-4.2	-5.2	-4.8	-5.0
Offset From Expected Cyclical Rebound	2.9	3.6	2.9	2.2
Memo: Change in Unemployment Rate	3.5	4.5	2.1	1.8

In this table, business fixed investment excludes farm tractors, other agricultural machinery, mining and oilfield machinery, farm structures, and mining exploration, shafts, and wells.
Peak and trough refer to business fixed investment excluding mining and farming as a percent of potential GDP.
Accelerator effects may not add up due to rounding.

leaving out any data beyond the first quarter of 1993.[32] Doing so had little effect on labor hours at full employment, technology at full employment, or service lives and depreciation lives of the aggregate categories of capital. Removing the last 10 years from the data reduced the average historical ratio of dividends to after-tax profits from 0.556 to 0.509, and boosted the average historical values for the nominal corporate cost of funds and expected inflation.

The most important changes in the data were changes in the limits for the logits used in constructing the fitted cost shares. Following the rule I used in constructing those limits meant reducing the upper limits for both computers and software and raising the lower limit for structures. The change in the upper limit for the software share, from 0.17 to 0.10, was the largest of those changes.

Lastly, the investment equations were estimated over sample periods ending in the first quarter of 1993. For total investment excluding mining and farming, the fitted values for the second quarter of 1993 to the first quarter of 2003 are remarkably close to fitted values from equations estimated over the entire sample, and do a reasonably good job of capturing the boom and bust in investment (see Figure 12). Remarkably, the average absolute error in total investment from the fourth quarter of 1970 to the first quarter of 2003 is actually smaller using fitted values from the shortened sample (0.30 percent of potential GDP) than using fitted values from the full sample (0.31 percent of GDP).

[32] For the shortened sample period, I used the data as they are now reported, including all revisions made since 1993. I did not use the data as they appeared in early 1993.

However, errors for individual categories of investment are noticeably worse over the 1993-2003 period. In fact, they are so much worse that the sum of the absolute errors for each of the five categories of investment averages 0.68 percent of potential GDP over 1970:q4 to 2003:q1 for fitted values from the shortened sample, compared with 0.53 percent of potential GDP for fitted values from the full sample.

That deterioration of fit occurs because the reduced upper limits in the logit functions for computers and software fail to allow for the continuing shift of overall investment toward these categories during the 1990s and early 2000s. The forecasted share of computers and software in overall investment is much too small outside of the sample period (Figure 13). The results suggest that incorrect forecasts for the fitted shares have little effect on one's forecast for overall nominal investment, although real investment could be affected if price deflators for different categories of investment are different.

5.7 Implications

Together, adverse demand shocks, the rise in the real cost of funds, and slow growth in labor hours at full employment can explain all of the slowdown in investment between 2000 and 2003 (see Figure 14). Those factors also explain why the recent downturn in investment relative to potential GDP was so deep compared with prior downturns, even though the downturn in actual GDP was mild. In addition to a rise in the cost of funds and a slowdown in growth in output less growth in productivity as severe as in prior downturns, a smaller-than-normal widening of the gap between actual output and output at full employment caused businesses to expect a slower cyclical rebound than after prior recessions. Thus, I find that businesses have not cut back on investment because of an irrational pessimism, but instead have only reacted to changes in the fundamentals the way they always have. Without those adverse changes in fundamentals, investment would actually have risen as a share of potential GDP from the end of 1999 to early 2003.

Two factors account for much of the variation left in the investment share of GDP after the effects of demand shocks, fluctuations in the cost of funds, and growth of labor hours at full employment are removed. First, the shift in the mix of new capital toward computers and software has gradually reduced the service life of the stock, producing a gradual rise in replacement demand over time.[33]

Second, favorable tax treatment from mid-1981 through 1986 boosted investment during that period. Of course, if one were assessing the overall impact of changes in tax law on investment, one would also need to account for any possible effects of tax incentives on the real cost of funds and on demand. While the overall impact of tax incentives on investment is ambiguous without further analysis, the direct effect was

[33] The reduction in service life also boosts the cost of capital, producing a partial offset to higher replacement demand.

positive.

An important implication for the future is that investment is unlikely to regain the share of potential GDP it had in the late 1990s. Figure 14 shows that, were demand shocks to be eliminated and labor hours growth return to its historical average, business fixed investment excluding mining and farming would be a smaller share of potential GDP than during 1998-2000. Even so, those developments would push investment well above its level in early 2003.

Another important implication for the future is that slower growth in the labor force as baby boomers begin to retire will slow investment, all else equal. In a world of putty-clay capital, slower growth in the number of workers means fewer expansion machines. Unless the cost of capital falls, investment will grow more slowly.

References

Arrow, Kenneth J. (1962). "The Economic Implications of Learning by Doing." *Review of Economic Studies*, Vol. 29, No. 3: 155-173.

Bischoff, Charles W. (1971). "The Effect of Alternative Lag Distributions." In *Tax Incentives and Capital Spending*, edited by Gary Fromm. Washington, D.C.: Brookings Institution.

Caballero, Ricardo J. (1999). "Aggregate Investment." In *Handbook of Macroeconomics: Volume 1B*, edited by John B. Taylor and Michael Woodford. Amsterdam: Elsevier.

Chirinko, Robert S. (1993). "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications." *Journal of Economic Literature*, Vol. 31, No. 4: 1875-1911.

Clark, J. Maurice (1917). "Business Acceleration and the Law of Demand: A Technical Factor in Economic Cycles." *Journal of Political Economy*, Vol. 25, No. 3: 217-235.

Congressional Budget Office (2001). *CBO's Method for Estimating Potential Output: An Update* (August).

Cooper, Russell, John Haltiwanger, and Laura Power (1999). "Machine Replacement and the Business Cycle: Lumps and Bumps." *American Economic Review*, Vol. 89, No. 4: 921-946.

Doms, Mark, and Timothy Dunne (1998). "Capital Adjustment Patterns in Manufacturing Plants." *Review of Economic Dynamics*, Vol. 1, No. 2: 409-429.

Fraumeni, Barbara M. (1997). "The Measurement of Depreciation in the U.S. National Income and Product Accounts." *Survey of Current Business*, Vol. 77, No. 7: 7-23.

Gilchrist, Simon and John C. Williams (1998). "Investment, Capacity, and Output: A Putty-Clay Approach." Finance and Economics Discussion Series No. 1998-44 (Washington, D.C.: Federal Reserve Board).

Gravelle, Jane G. (1994). *The Economic Effects of Taxing Capital Income.* Cambridge, Mass.: The MIT Press.

Hall, Robert E. and Dale W. Jorgenson (1967). "Tax Policy and Investment Behavior." *American Economic Review*, Vol. 57, No. 3: 391-414.

Hall, Robert E. and Dale W. Jorgenson (1971). "Application of the Theory of Optimum Capital Accumulation." In *Tax Incentives and Capital Spending*, edited by Gary Fromm. Washington, D.C.: Brookings Institution.

Harper, Michael J., Ernst R. Berndt, and David O. Wood (1989). "Rates of Return and Capital Aggregation Using Alternative Rental Prices." In *Technology and Capital Formation*, edited by Dale W. Jorgenson and Ralph Landau. Cambridge, Mass.: The MIT Press.

Johansen, Leif (1959). "Substitution Versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis." *Econometrica*, Vol. 27, No. 2: 157-176.

Jorgenson, Dale W. (1963). "Capital Theory and Investment Behavior." *American Economic Review*, Vol. 53, No. 2: 247-259.

Kiley, Michael (2001). "Business Investment in the Federal Reserve Board's U.S. Model (FRB/US): Specification and Implications." (draft, Federal Reserve Board, April 2001).

Kuh, Edwin (1965). "Cyclical and Secular Labor Productivity in United States Manufacturing." *Review of Economics and Statistics*, Vol. 47, No. 1: 1-12.

Kydland, Finn E. and Edward C. Prescott (1982). "Time to Build and Aggregate Fluctuations." *Econometrica*, Vol. 50, No. 6: 1345-1370.

Laurence H. Meyer & Associates, Ltd. (1995). *The WUMM Model Book*. St. Louis: Laurence H. Meyer & Associates.

Okun, Arthur M. (1981). *Prices and Quantities.* Washington, D.C.: The Brookings Institution.

Phelps, Edmund S. (1963). "Substitution, Fixed Proportions, Growth and Distribution." *International Economic Review*, Vol. 4, No. 3: 265-288.

Solow, Robert M. (1956). "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, Vol. 70, No. 1: 65-94.

Solow, Robert M. (1962). "Substitution and Fixed Proportions in the Theory of Capital." *Review of Economic Studies*, 29, 3: 207-218.

Tevlin, Stacey and Karl Whelan (2000). "Explaining the Investment Boom of the 1990s." Finance and Economics Discussion Series No. 2000-11 (Washington, D.C.: Federal Reserve Board).

U.S. Department of Labor, Bureau of Labor Statistics (1983), *Trends in Multifactor Productivity*, Bulletin 2178 (September).

U.S. Department of Labor, Bureau of Labor Statistics (2001), *Revisions to Capital Inputs for the BLS Multifactor Productivity Measures*, October 16 (available at http://stats.bls.gov/web/mprcaptl.htm).

Figure 1. Labor's Share of Corporate GDP
(Percent of Corporate GDP Excluding Indirect Taxes)

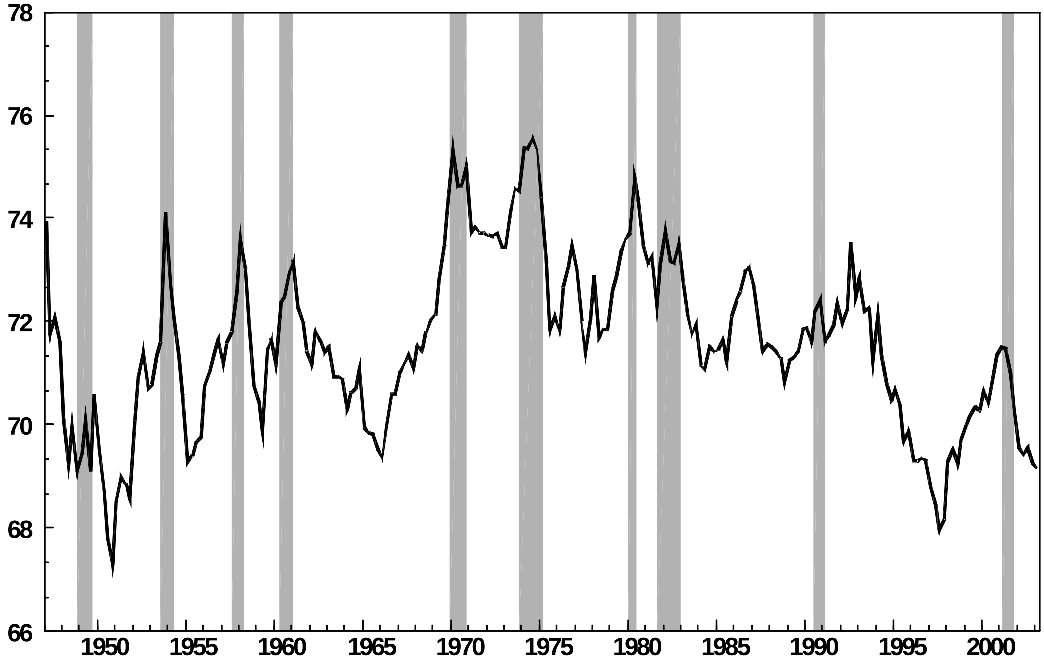
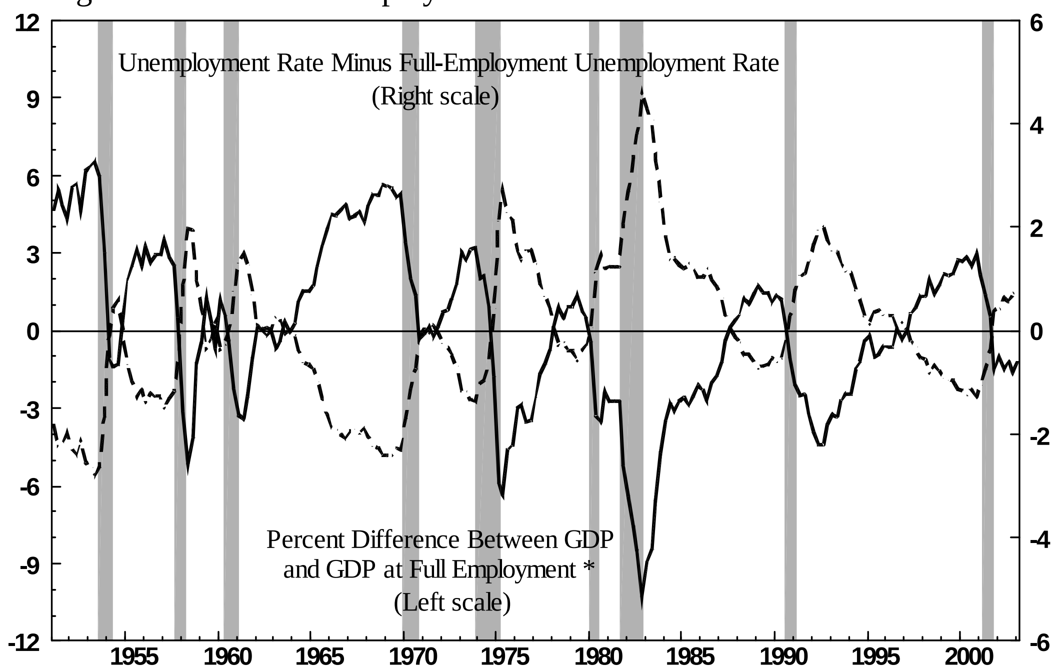


Figure 2. GDP and Unemployment



* GDP is for nonfarm business, households, and institutions.

Figure 3. Ratio of Actual GDP to Potential GDP
(Nonfarm Business, Households, and Institutions)

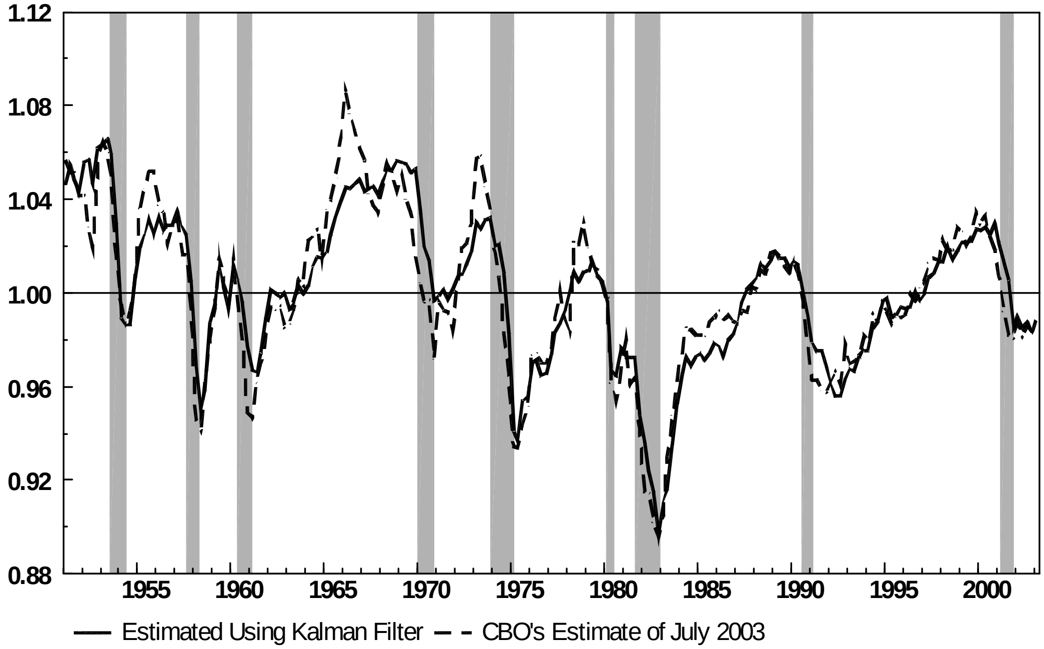


Figure 4. Growth of Two Measures of Potential GDP
(Percent Change from Year Ago; Nonfarm Business, Households, and Institutions)

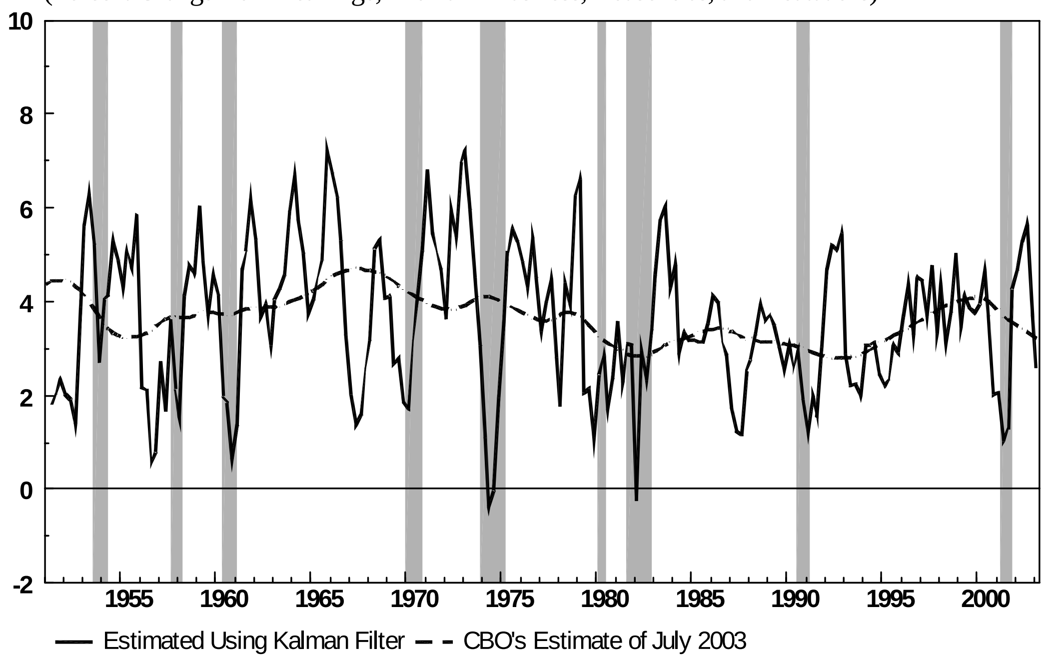


Figure 5. Cost Share for Software
(Percent of Capital Costs of All New Non-Replacement Capital)

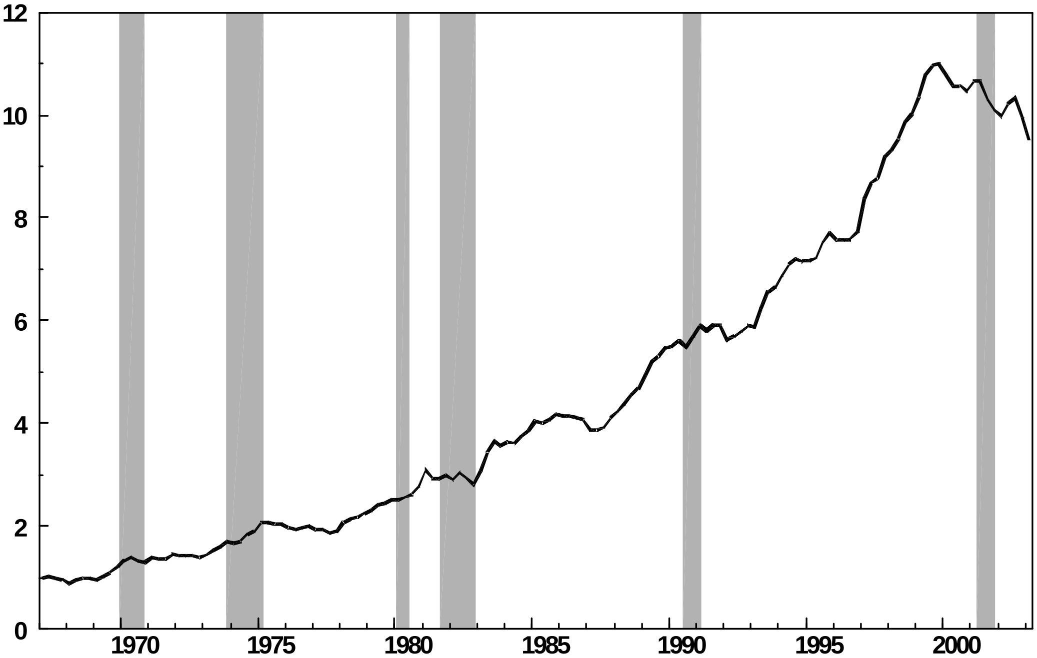


Figure 6. Business Fixed Investment, Except Mining and Farming
(Percent of CBO's July 2003 Estimate of Potential GDP)

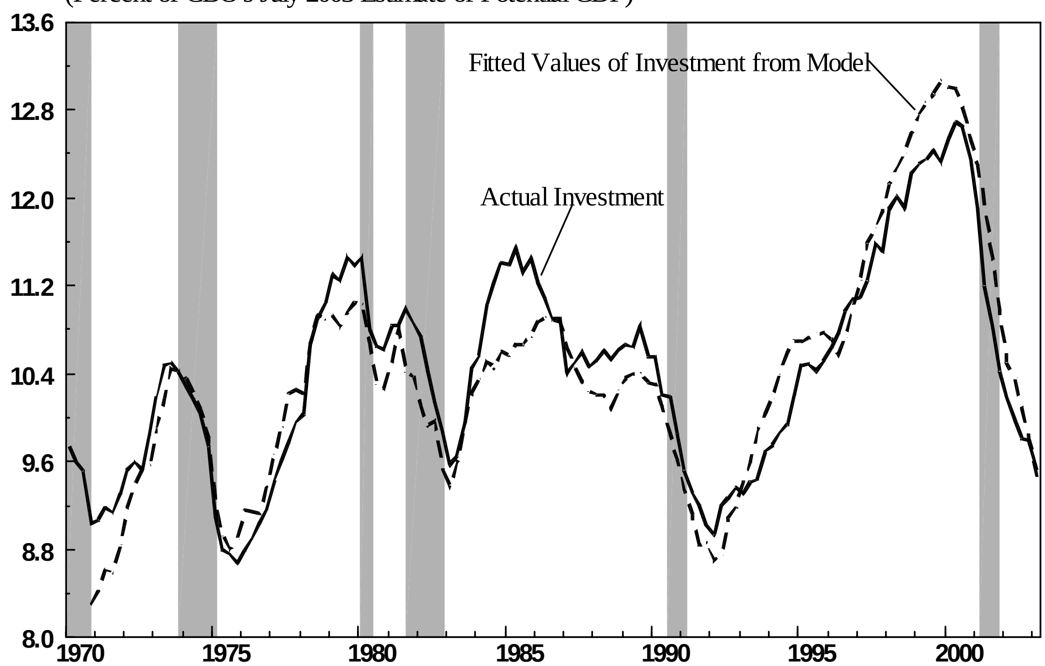


Figure 7. Effect of Demand on Business Fixed Investment
(Percent of CBO's July 2003 Estimate of Potential GDP)

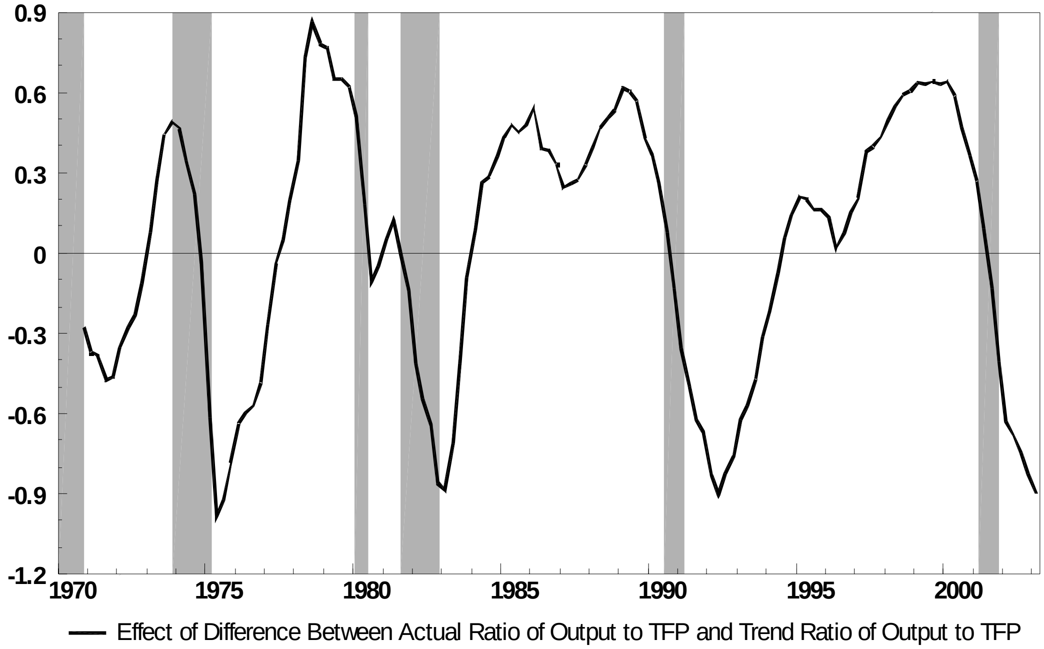


Figure 8. Removing Demand Effects from Business Fixed Investment
(Percent of CBO's July Estimate of Potential GDP)

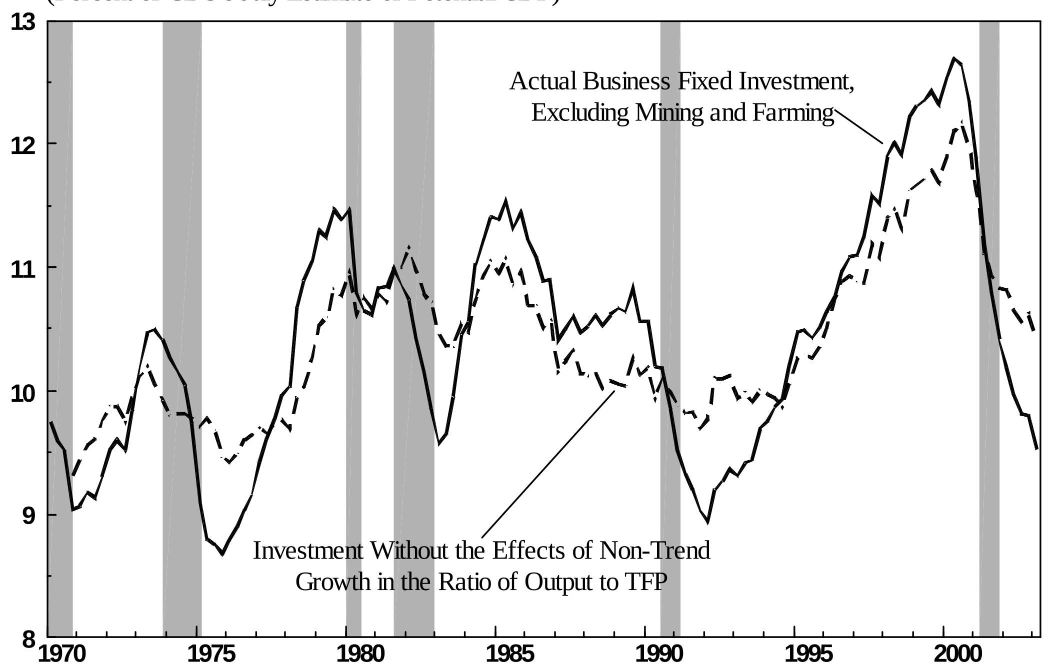


Figure 9. Effect of the Cost of Funds on Business Fixed Investment
(Percent of CBO's July 2003 Estimate of Potential GDP)

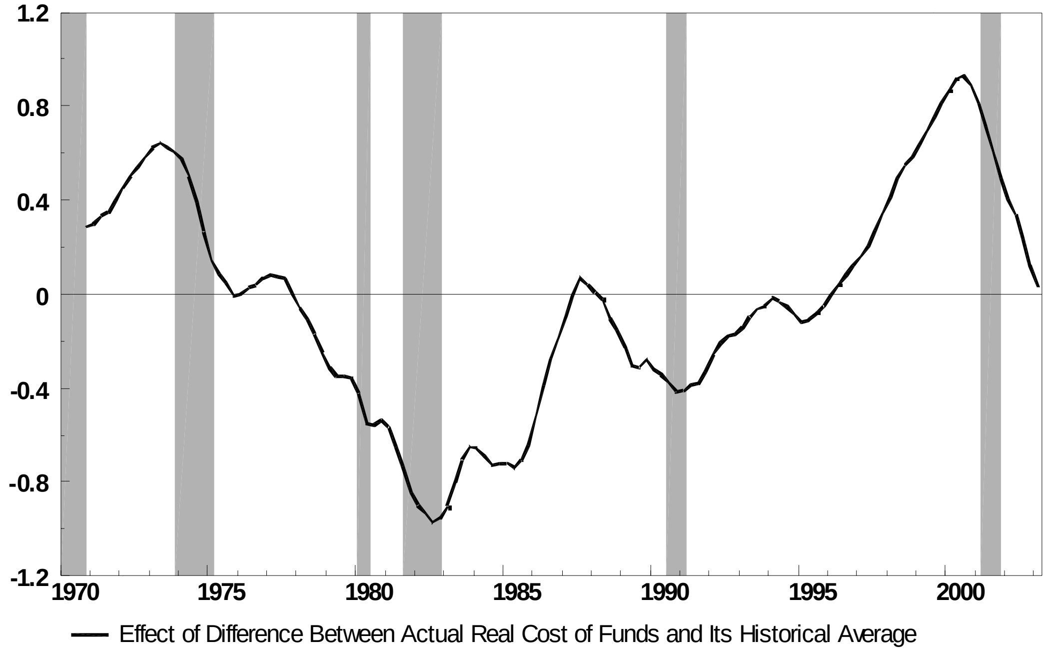


Figure 10. Business Fixed Investment, Without Demand and Cost Effects
(Percent of CBO's July 2003 Estimate of Potential GDP)

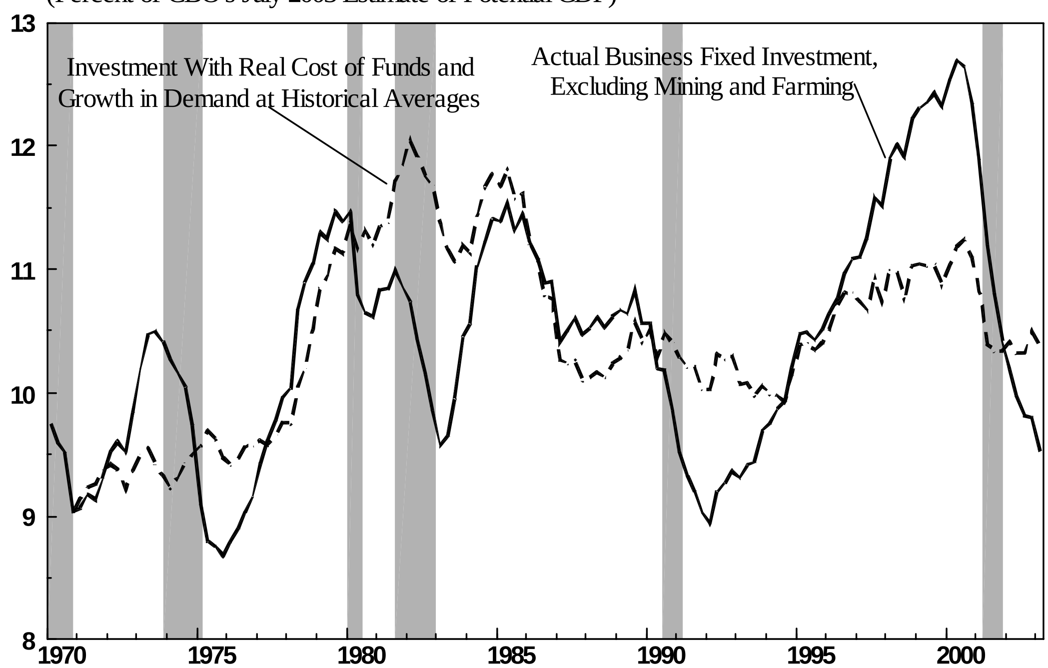


Figure 11. Effect of Full-Employment Hours on Business Fixed Investment
(Percent of CBO's July 2003 Estimate of Potential GDP)

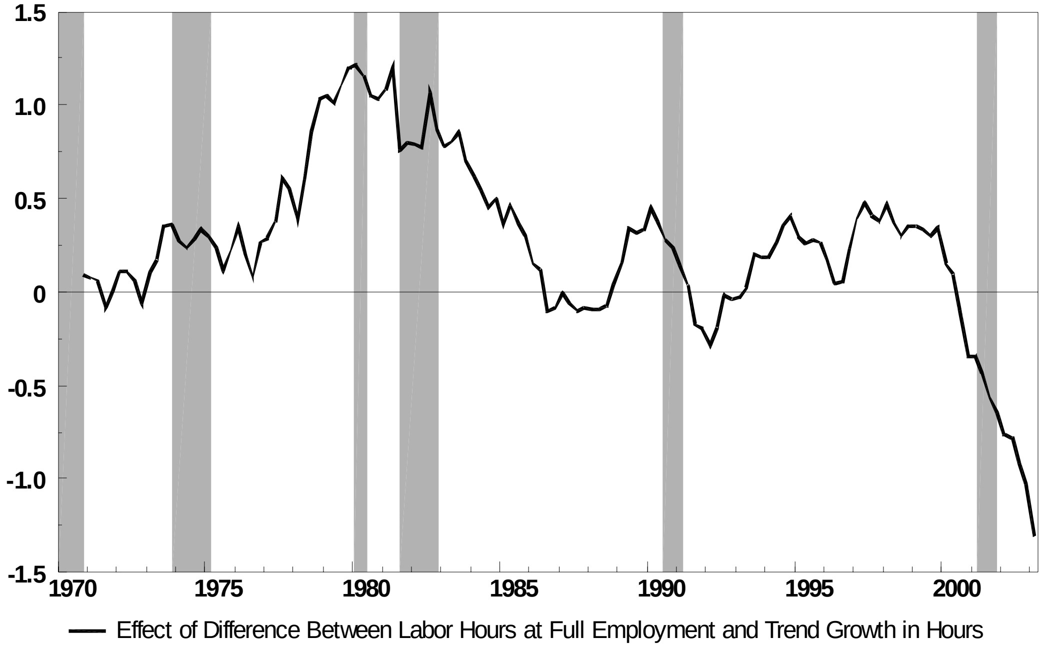
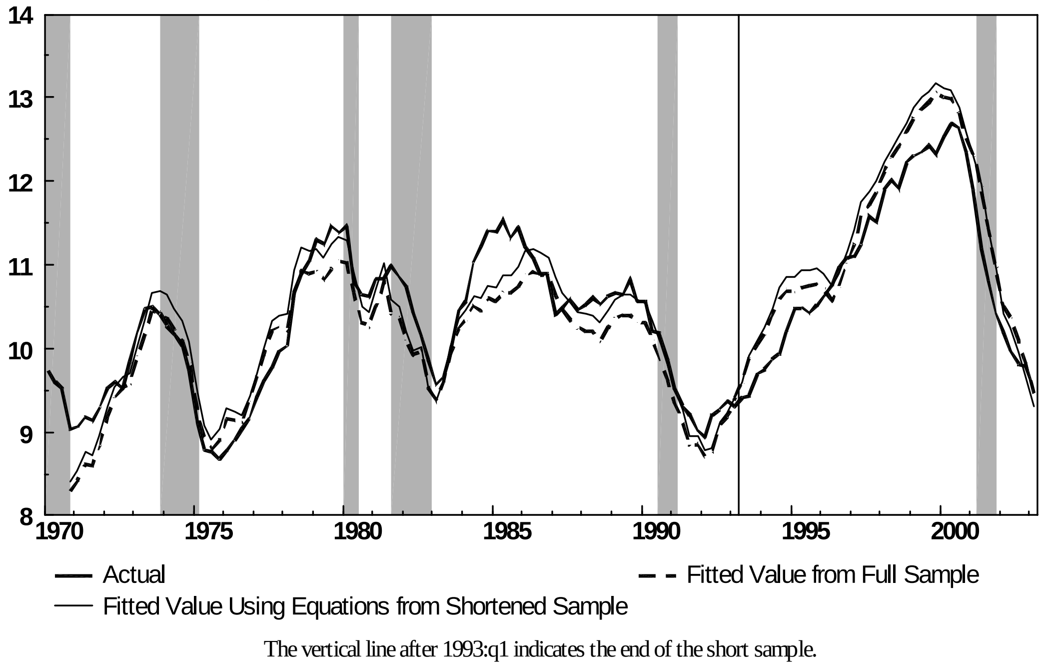
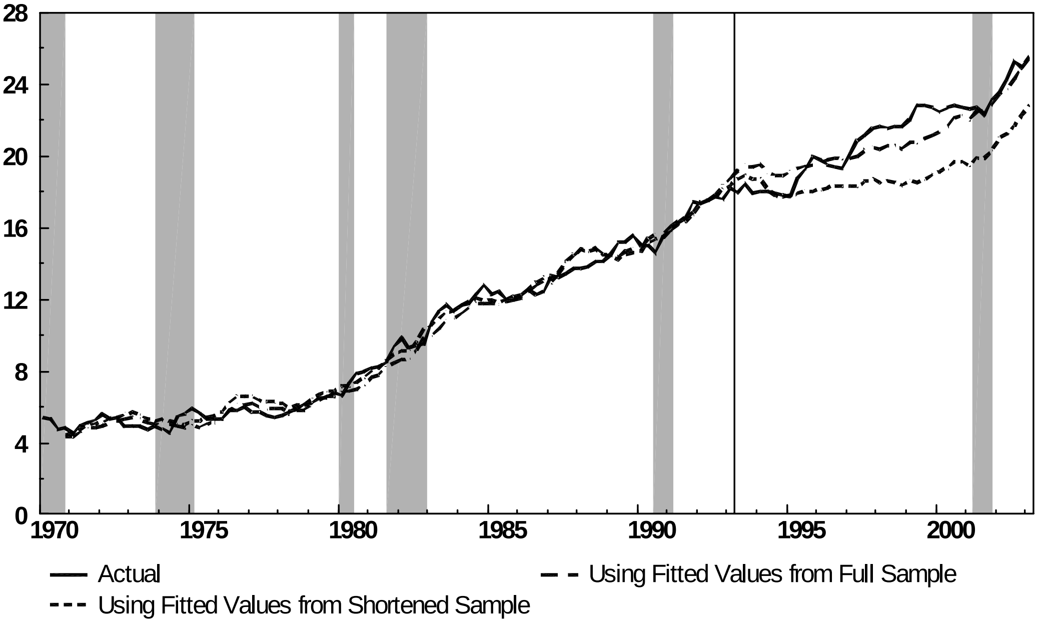


Figure 12. Business Fixed Investment, Excluding Mining and Farming
(Percent of CBO's July 2003 Estimate of Potential GDP)



The vertical line after 1993:q1 indicates the end of the short sample.

Figure 13. Investment in Computers and Software
(Percent of Total Business Fixed Investment, Excluding Mining and Farming)



The vertical line after 1993:q1 indicates the end of the short sample.

Figure 14. BFI Without Effects of Demand, Cost of Funds, or Demographics
(Percent of CBO's July 2003 Estimate of Potential GDP)