How Structural Are Structural Parameters?

Jesús Fernández-Villaverde

Duke University, NBER, and CEPR

Juan F. Rubio-Ramírez

Duke University and Federal Reserve Bank of Atlanta

March 5, 2007

1. Introduction

This paper studies the following problem: how stable over time are the so-called "structural parameters" of dynamic stochastic general equilibrium (DSGE) models? To answer this question, we estimate a medium-scale DSGE model with real and nominal rigidities using U.S. data. In our model, we allow for parameter drifting and rational expectations of the agents with respect to this drift. We document that there is strong evidence that parameters change within our sample. In particular, we illustrate variations in the parameters describing the monetary policy reaction function and in the parameters characterizing the pricing behavior of firms and households. Moreover, we show how the movements in the parameters are correlated with the evolution of inflation and are consistent with alternative sources of information. Our results cast doubts on some of the justifications for the empirical implementation of DSGE models, at least in their current form.

Our findings are important because DSGE models are at the core of modern macroeconomics. They promise to be a laboratory that researchers can employ to match theory with reality, to design economic policies, and to evaluate welfare. The allure of DSGE models has captured the imaginations of many, inside and outside academia. In universities, a multitude of economists implements DSGE models in their many varieties and fashions. More remarkable still, a burgeoning number of policy-making institutions are estimating DSGE models for policy analysis and forecasting (An and Schorfheide, 2006). The Federal Reserve Board, the European Central Bank, the International Monetary Fund, and the central banks of Austria, Canada, Germany, Italy, Norway, Spain, and Sweden are at the front of the tide, but many other institutions are keen to jump on the bandwagon. In addition, the profession is accumulating experience of the good forecasting record of DSGE models, even when compared with judgmental predictions from staff economists (Laforte and Windle, 2006).

At the center of DSGE models, we have the "structural parameters" that define the preferences and technology of the economy. Usually, we call these parameters "structural" in the sense of Hurwicz (1962): they are invariant to interventions, including shocks by nature. The structural character of the parameters is responsible for much of the appeal of DSGE models. Since the parameters are fully interpretable from the perspective of economic theory and invariant to policy interventions, DSGE models avoid the Lucas critique and can be used to quantitatively evaluate policy.

Our point of departure is that, at least at some level, it is hard to believe that the "structural parameters" of DSGE models are really structural given the class of interventions we are interested in. Let us think, for instance, about technology. Most DSGE models specify a stable production function, perhaps subject to productivity growth. Except in a few papers (Young, 2004), the features of the technology, like the elasticity of output to capital, are constant over time. But this constant elasticity is untenable in a world where technological change is purposeful. We can expect that changes in the relative price of capital versus labor will induce changes in the new technologies developed and that those may translate into different elasticities of output to inputs. Similar arguments can be made along pretty much every dimension of a modern DSGE model.

The previous argument is not sufficient to dismiss the practice of estimating DSGE models with constant parameter values. Simplifying assumptions, like stable parameters, are required to make progress in economics. However, as soon as we realize the possible changing nature of "structural" parameters, we weaken the justifications for inference exercises underlying the program of DSGE modelling. The separation between what is "structural" and what is reduced-form becomes much more ambiguous.[1]

The possibility but not the necessity of parameter drifting motivates the main question of this paper: how much evidence of parameter drifting in DSGE models is in the data? If the answer is that we find much support for drifting (where the metric to decide "much" needs to be discussed), we would need to re-evaluate the usefulness of our estimation exercises or at least modify our models to account for parameter variation. Moreover, parameter drifting may also be interpreted as a sign of model misspecification and, possibly, as a guide for improving our models. If the answer is negative, i.e., if we find little evidence of parameter drifting, we would increase our confidence in DSGE modelling as a way to tackle relevant policy discussions.

Beyond addressing our substantive question, this paper also develops new tools for the estimation of dynamic equilibrium models with parameter drifting. We show how the combination of perturbation methods and the particle filter allows the efficient estimation of this

[1] Indeed, Hurwicz (1962) himself emphasized the contingency of the definition of structural parameter: "...the *concept* of structure *is relative to the domain of modifications anticipated*", "If two individuals differ from regard to modifications they are willing to consider, they will probably differ with regard to the relations accepted as structural," and "...this relativity of the concept of structure is due to the fact that it represents *not a property of the material system under observation*, but rather a property of the anticipations of those asking for predictions concerning the state of the system" (italics in the original).

class of economies. Indeed, all the required computations can be implemented in a good PC in a reasonable amount of time. We hope that those tools may be put to good use in other applications, not necessarily in general equilibrium, that involve time-varying parameters in essential ways.

Our main results are as follows. First, we offer compelling proof of changing parameters in the Fed's behavior. Monetary policy became appreciably more aggressive in its stand against inflation after Volcker's appointment. This result agrees with Clarida, Galí, and Gertler (2000), Lubick and Schorfheide (2004), and Boivin (2006). Our contribution is that we re-derive the result within in the context of a model where agents understand and act upon the fact that monetary policy can change over time.

Second, we expose the instability of the parameters controlling the level of nominal rigidity and indexation of prices and wages. Those changes are strongly correlated with changes in inflation in an intuitive way: lower rigidities correlate with higher inflation and higher rigidities with lower inflation. Our finding suggests that a more thorough treatment of nominal rigidities, possibly through state-dependent pricing models, may have high payoffs in terms of data fitting and policy analysis.

We want to be up-front about some of the shortcomings of our exercise. First, and foremost, we face the limitation of the data. With 184 quarterly observations of the U.S. economy, there is a tight bound on how much we can learn from the data (Ploberger and Phillips, 2003, frame the problem of empirical limits for time series models precisely in terms of information bounds.) The main consequence of the limitations of the short sample size is relatively imprecise estimates.

The second limitation, forcefully emphasized by Sims (2001), is that we do not allow for changing volatilities in the innovations of the model, which is itself a particular form of parameter drift. If the innovations are heteroskedastic (and we have argued ourselves in Fernández-Villaverde and Rubio-Ramírez, 2007, that there is notable evidence of stochastic volatility in the U.S. data), the estimation may attempt to pick up the changing variance by spurious changes in the structural parameters. At the same time, Cogley and Sargent (2005) find that there is still variation in the parameters of a VAR even after controlling for heteroskedasticity. We are currently working on an extension of the model where we allow parameter drifting and changing volatilities.

In our work, we build upon an illustrious tradition of estimating models with parameter drifting. One classic reference is Cooley and Prescott (1976), where the authors studied the estimation of regression parameters that are subject to permanent and transitory shocks. Unfortunately, the techniques in this tradition are within the context of the Cowles Commission's framework and, hence, are of little direct application to our investigation.

Our paper is also linked with a growing body of research that shows signs of parameter drifting on dynamic models. Since the estimation of this class of models is a new undertaking, the evidence is scattered. One relevant literature estimates VARs with time-varying parameters and/or stochastic volatility. Examples include Uhlig (1997), Bernanke and Mihov (1988), Cogley and Sargent (2005), Primiceri (2005), and Sims and Zha (2006). The consensus emerging from these papers is that there is evidence of time variation in the parameters of a VAR, although there is a dispute about whether the variation comes mainly from changes in the autoregressive components or from stochastic volatility. This evidence, however, is only suggestive, since a DSGE model with constant parameters may be compatible with a time-varying VAR (Cogley and Sbordone, 2006).

A second literature has estimated equilibrium models with variation in some parameters, but it has been much less ambitious in the extent of the fluctuations allowed. Fernández-Villaverde and Rubio-Ramírez (2007) and Justiniano and Primiceri (2006) demonstrate the importance of stochastic volatility to account for U.S. data using a DSGE model. King (2005) works with a simple RBC economy with parameter drift in four parameters. However, his approach relies on particular properties of his model and it is too cumbersome to be of general applicability. Boivin (2006) estimates a parameter-drifting Taylor rule with real-time data. He corroborates previous findings of changes in the rule coefficients obtained with final data. Benati (2006), elaborating on an argument by Woodford (2006), questions the indexation mechanisms introduced in New Keynesian models and shows evidence that they are not structural to changes in monetary policy rules. Oliner, Rudebusch, and Sichel (1996) find unstable parameters even in investment models with more intricate representations of capital spending than those found in current DSGE models.

There are also numerous papers that tell us much about parameter drifting, albeit in an indirect way. A common practice when estimating models has been to divide the sample into two periods, usually before and after 1979, and argue that there are significant differences in the inference results. One celebrated representative of this method is Clarida, Galí, and Gertler (2000), a paper we will discuss later in more detail.

Finally, a literature that shares many connections with our analysis is the one that deals with DSGE models that have a Markov-switching process in some aspect of the environment, like monetary or fiscal policy (Davig and Leeper, 2006a and 2006b, Chung, Davig, and Leeper, 2006, and Farmer, Waggoner, and Zha, 2006). The stated motivation of these papers is that Markov switches may help us understand the dynamics of the economy better. So far, none of these papers has produced an estimated model.

The rest of the article is organized as follows. First, in section 2, we discuss different ways to think about parameter drifting in dynamic equilibrium models. In section 3, we develop two simple examples of parameter drift that motivate our investigation. Section 4 spells out a medium-scale model of the U.S. economy and discusses how to take this model to the data. Section 5 introduces parameter drifting and explains how to adapt the approach in section 4 to handle this situation. We report our results in section 6. Section 7 concludes. An appendix provides the interested reader with some technical details.

2. Parameter Drifting and Dynamic Equilibrium Models

There are at least three ways to think about parameter drifting in an estimated DSGE model. The simplest approach, which we call the pure econometric interpretation, is to consider parameter drifting as a convenient phenomenon to fit the data better or as the consequence of a capricious nature that agents in the model neither understand nor forecast. Despite its simplicity, this interpretation violates the spirit of rational expectations: not having free parameters that the researcher can play with. Consequently, we will not investigate this case further.

The second way to think about parameter drifting is as a characteristic of the environment that the agents understand and act upon. Let us come back to our example of the production function. Imagine that the aggregate technology is given by a Cobb-Douglas function $Y_t = AK_t^{\alpha_t} L_t^{1-\alpha_t}$ where output Y_t is produced with capital K_t and labor L_t given some technology level A and some share parameter α_t. The only difference with the standard environment is that α_t is indexed by time (neither the realism nor the empirical justification of our example is crucial for the argument, although we could argue in favor of both features). Let us also assume that α_t evolves over time as a random walk with reflecting boundaries at 0 and 1, to ensure that the production function satisfies standard properties. We could imagine that such drift comes about because the new technologies developed have a random requirement of capital. The solution of the agent's problems are decision rules that have as one of their arguments the current α_t. Why? First, because α_t determines current prices. Second, because α_t helps to forecast future values α_{t+j} and hence to predict future prices. This interpretation is our favorite one, and it will frame our reading of the results in section 6.

The final perspective about parameter drifting is as a telltale of model misspecification. This point, raised by Cooley (1971) and Rosenberg (1968), is particularly cognate when estimating DSGE models. These models are complex constructions. To make them useful for policy purposes, researchers add many mechanisms that affect the dynamics of the economy: sticky prices and wages, adjustment costs, etc. In addition, DSGE models require many parametric assumptions: the utility function, the production function, the adjustment costs, the distribution of shocks, etc. If we seriously misspecified the model along at least one dimension, parameter drifting may appear as the only possibility left to the model to fit the data. Our example in section 3 illustrates this point in detail. We will exploit this possibility in our empirical results and assess how the drift in the parameters determining the degree of nominal rigidity in the economy implies that time-dependent models of pricing decisions may be flawed.

3. Two Examples

In this section, we present two simple examples that generate parameter drifting in estimated DSGE models. We have chosen the examples to illustrate our points as clearly as possible and not based on their empirical relevance or plausibility. However, the examples are not far-fetched: they deal with recurrent themes in the literature and are linked, albeit we do not explore this connection to its fullest, to relevant features of the economy.

3.1. Parameter Drift as a Consequence of Changing Policies

The first example deals with the changes in the elasticity of monetary policy to different variables. It is common to postulate that the monetary authority uses open market operations to set the short-run nominal interest rate R_t according to a Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{y_t}{\widehat{y}_t}\right)^{\gamma_y} \right)^{1-\gamma_R} \exp\left(\sigma_m \varepsilon_{mt}\right)$$

The variable Π represents the target levels of inflation of the monetary authority, R the steady-state gross return of capital, y_t is output, and $\widehat{y}_t$ some measure of target output. The term ε_{mt} is a random shock to the systematic component of monetary policy and is distributed according to $\mathcal{N}(0,1)$.

In an influential contribution, Clarida, Galí, and Gertler (2000) drove the attention of the profession to changes in the elasticity parameter γ_Π before and after Volcker's appointment as Fed Chairman in 1979. They document, with a slightly different specification of the Taylor rule, that γ_Π more than doubles after 1979. This finding has been corroborated in many studies and found resilient to modifications in the empirical specification (see the results and references in Lubick and Schorfheide, 2004). The division of the sample between the time before and after 1979 has also been exploited in papers such as Boivin and Giannoni (2006), who find that the point estimates of the structural parameters also substantially vary between the two periods.

Changes in the policy coefficients are one particular example of parameter drift. We can think of them as consequences of modifications in priorities by the policy-makers or as consequences of changes in the perception of the effectiveness of monetary policy, a point keenly defended by Sargent (1999). Once we recognize that there is much evidence of the parameter γ_Π drifting over time, it is natural to pursue the consequences of agents realizing that these changes are a possibility and acting upon them. Also, such an environment may capture some of the insights of Sims (1980) about the difference between a change in policy regime (in our Taylor rule, a change in the way the interest rate is determined) and the evolution of the policy within one regime, which can be interpreted in our context as the drift of the parameters of the rule.

3.2. Parameter Drift as a Telltale of Model Misspecification

Our second example revisits several of the themes in Browning, Hansen, and Heckman (1999). We explore the consequences for inference of an econometrician estimating a model with infinitely lived agents when the data have actually been generated by an overlapping generations model. We show how our estimate of the discount factor will be a function of the true discount factor, the elasticity of output to capital, and the (changing) age distribution of the population. This example is relevant because variations in the age structure of the U.S. population have been continuous due both to changes in fertility and in mortality.

3.2.1. An Artificial World

We begin by creating a simple artificial world. In each period t, there are two generations of households alive, young and old. Each household maximizes the life utility

$$\log c_t^t + \beta \mathbb{E}_t \log c_{t+1}^t$$

where the superindex denotes that the household was born in period t, the subindex the period in which it consumes, and $\mathbb{E}_t$ is the conditional expectations operator. The discount factor, β, captures the preference for current consumption. We pick a log utility function to simplify the algebra below.

Households work when young and get a wage w_t for a unit of time that they supply inelastically. Households live off their savings when they are old. The period budget constraints are given by $c_t^t + s_t = w_t$ and $c_{t+1}^t = R_{t+1}s_t$, where s_t is the household savings and R_{t+1} the gross return on capital. From the first order condition of households, we have that $c_t^t = \frac{1}{1+\beta}w_t$ and $c_t^t = \frac{\beta}{1+\beta}w_t$.

In each period, a number n_t of new households is born. For the moment, we will assume only that l_t is the realization of some random process. Nothing of substance for our argument is lost by assuming that the size of the new generation is exogenous.

The production side of the economy is defined by a Cobb-Douglas production function $y_t = k_t^\alpha l_t^{1-\alpha}$ where k_t is the total amount of capital in the economy and l_t the total amount of labor. If we assume total depreciation in the economy, again to simplify the algebra, and impose the condition $l_t = n_t$, we get by competitive pricing $w_t = (1-\alpha) k_t^\alpha n_t^{-\alpha}$ and $R_t = \alpha k_t^{\alpha-1} n_t^{1-\alpha}$.

Now, all that remains is some accounting. Total consumption in the economy in period t, C_t, is equal to the total consumption of the old generation plus the total consumption of the young generation. The old consume all of their income, which is equal to the total capital income of the economy, $R_t k_t = \alpha k_t^\alpha n_t^{1-\alpha}$. The young consume a fraction $\frac{1}{1+\beta}$ of their total income, which is equal to the total labor income of the economy $w_t l_t = (1-\alpha) k_t^\alpha n_t^{1-\alpha}$. Then total consumption is:

$$C_t = \frac{1+\alpha\beta}{1+\beta} k_t^\alpha n_t^{1-\alpha}$$

By the aggregate resource constraint, investment (or, equivalently, capital in period $t+1$) is

$$I_t = k_{t+1} = \frac{(1-\alpha)\beta}{1+\alpha\beta} C_t$$

Finally, we can find per capita consumption c_t^{pc} as:

$$c_t^{pc} = \frac{C_t}{n_t + n_{t-1}}$$

3.2.2. An Econometrician

Let us now suppose that we have an econometrician who aims to estimate with T observations generated from our economy, a version of the model with a representative infinitely lived agent. To do so, the econometrician postulates that the agent has a utility function:

$$\max_{\{c_t^{pc}\}} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \left[\prod_{i=0}^{t} (1+\gamma_t) \right] \log c_t^{pc}$$

where γ_t is the (random) growth rate of the population between periods $t-1$ and t:

$$1 + \gamma_t = \frac{n_t + n_{t-1}}{n_{t-1} + n_{t-2}}$$

and $\gamma_0 = 0$. This utility function is the same as in the canonical presentation of the RBC model in Cooley and Prescott (1995) except that the growth rate of the population is stochastic instead of constant. The production side of the economy is the same as before, $y_t = k_t^\alpha l_t^{1-\alpha}$. Thus, the only difference between the artificial world we have created and the model the econometrician estimates is that, instead of having two generations alive in each moment, the econometrician estimates a model with a representative agent.

What are the consequences on the estimated parameters? Imagine that the econometrician knows α and that the depreciation factor is 1. Then, a simple procedure to estimate the only remaining unknown parameter in the model, the discount factor β, is to build the the population moment:

$$\frac{1}{c_t^{pc}} = \beta \mathbb{E}_t \left(1 + \gamma_{t+1}\right) \frac{R_{t+1}}{c_{t+1}^{pc}}$$

and substitute the expectation by the sample mean:

$$\widehat{\beta}_T = \frac{\frac{1}{T-1}\sum_{t=0}^{T-1}\frac{1}{c_t^{pc}}}{\frac{1}{T-1}\sum_{t=0}^{T-1}\left(1+\gamma_{t+1}\right)\frac{R_{t+1}}{c_{t+1}^{pc}}}$$

We study how this expression evolves over time. First, note that, by substituting the expressions found before, we get:

$$\left(1+\gamma_{t+1}\right)\frac{R_{t+1}}{c_{t+1}^{pc}} = \frac{\left(n_{t+1}+n_t\right)^2}{n_t+n_{t-1}}\frac{\alpha}{1-\alpha}\frac{1+\beta}{\beta}\frac{1}{C_t}$$

Then:

$$\widehat{\beta}_T = \beta\frac{1-\alpha}{\alpha}\frac{1}{1+\beta}\frac{\sum_{t=0}^{T-1}\left(n_t+n_{t-1}\right)\frac{1}{C_t}}{\sum_{t=0}^{T-1}\frac{\left(n_{t+1}+n_t\right)^2}{n_t+n_{t-1}}\frac{1}{C_t}}$$

We want to work on the previous expression. First, we substitute aggregate consumption for its value in terms of capital and labor:

$$\widehat{\beta}_T = \beta\frac{1}{1+\beta}\frac{1-\alpha}{\alpha}\frac{\sum_{t=0}^{T-1}\left(n_t+n_{t-1}\right)\frac{1}{k_t^{\alpha}n_t^{1-\alpha}}}{\sum_{t=0}^{T-1}\frac{\left(n_{t+1}+n_t\right)^2}{n_t+n_{t-1}}\frac{1}{k_t^{\alpha}n_t^{1-\alpha}}}$$

The only remaining endogenous element in this equation is k_t. To eliminate it, we recursively substitute k_{t-i} to find:

$$k_t = \left[\frac{(1-\alpha)\beta}{1+\alpha\beta}n_{t-1}^{1-\alpha}\prod_{i=1}^{t-1}\left(\frac{(1-\alpha)\beta}{1+\alpha\beta}n_{t-1-i}^{1-\alpha}\right)^{\alpha i}\right]k_0^{\alpha t}$$

Then:

$$\widehat{\beta}_T = \beta\frac{1}{1+\beta}\frac{1-\alpha}{\alpha}\frac{\sum_{t=0}^{T-1}\frac{n_t+n_{t-1}}{n_t^{1-\alpha}}\left(\left[n_{t-1}^{1-\alpha}\prod_{i=1}^{t-1}\left(\frac{(1-\alpha)\beta}{1+\alpha\beta}n_{t-1-i}^{1-\alpha}\right)^{\alpha i}\right]k_0^{\alpha t}\right)^{-\alpha}}{\sum_{t=0}^{T-1}\frac{\left(n_{t+1}+n_t\right)^2}{\left(n_t+n_{t-1}\right)n_t^{1-\alpha}}\left(\left[n_{t-1}^{1-\alpha}\prod_{i=1}^{t-1}\left(\frac{(1-\alpha)\beta}{1+\alpha\beta}n_{t-1-i}^{1-\alpha}\right)^{\alpha i}\right]k_0^{\alpha t}\right)^{-\alpha}}$$

which delivers a $\widehat{\beta}_T$, which is biased and drifts over time according to the evolution of the population. This expression is composed of three parts. First, the true parameter, β, second the deterministic bias,

$$\frac{1}{1+\beta}\frac{1-\alpha}{\alpha}$$

and finally the term involving the n_t's and k_0, which fluctuates over time.

Without further structure on population growth over time, it is difficult to say much about $\widehat{\beta}_T$. In the simple case where $\gamma_t = \gamma$ is constant, as $T \to \infty$, the only factor dominating is:

$$\widehat{\beta}_T \simeq \beta \frac{1}{1+\beta} \frac{1-\alpha}{\alpha} (1+\gamma)^{-2} \tag{1}$$

To explore the behavior of $\widehat{\beta}_T$ in the general case where γ_t varies, we simulate the model and estimate the parameter recursively with data from an economy with $\alpha = 0.3$ and $\beta = 0.96$. The growth rates of population are 2, 4, 3, 1, 2, and 5 percent each for 50 periods (i.e., for period 1 to 50, growth rate is 2 percent, for period 51 to 100, the growth rate is 4 percent and so forth). We plot our results in figure 2.3.1 where we can see the evolution over time of $\widehat{\beta}_T$ and how it inherits the properties of γ_t. To facilitate comparison with (1), we superimpose the value of (1) that would be implied if the growth rate in a period stayed constant over time. The graph shows how $\widehat{\beta}_T$ converges to (1) within each period.

4. The Baseline Model

We will structure our investigation around a baseline New Keynesian business cycle model. We pick this model because it is the paradigmatic representative of the DSGE economies estimated by practitioners. Since on other occasions (for example, Fernández-Villaverde, 2005), we have gone on the record criticizing the problems of this framework, we do not feel obliged to repeat those shortcomings here. Suffice it to say as a motivation that given the level of interest by policy-making institutions in this model, it is difficult to see a more appropriate vessel for our exploration.

The New Keynesian model is quite well known (see the book-length description in Woodford, 2003). Consequently, we will be short in our presentation, and we will omit some of the technical aspects. On the other hand, for concreteness and to make our quantitative results below meaningful, we need to discuss some aspects of the model in certain detail. The interested reader can get the whole description of the model at a complementary technical appendix posted at `www.econ.upenn.edu/~jesusfv/benchmark_DSGE.pdf`. In this section, to fix ideas, we will introduce the model without changes in the parameters. In section 5, we will introduce the parameter change over time.

4.1. Households

The basic structure of the economy is as follows. A representative household consumes, saves, holds money, supplies labor, and sets its own wages subject to a demand curve and Calvo's pricing. The final output is manufactured by a final good producer, which uses as inputs a continuum of intermediate goods manufactured by monopolistic competitors. The intermediate good producers rent capital and labor to manufacture their good. Also, these intermediate good producers face the constraint that they can only change prices following a Calvo's rule. Finally, there is a monetary authority that fixes the one-period nominal interest rate through open market operations with public debt. Long-run growth is induced by the presence of two unit roots, one in the level of neutral technology and one in the investment-specific technology. These stochastic trends will allow us to estimate the model with the raw, undetrended data.

We have a continuum of households in the economy indexed by j. The households maximize the following lifetime utility function, which is separable in consumption, c_{jt}, real money balances, m_{jt}/p_t, and hours worked, l_{jt}:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t d_t \left\{ \log\left(c_{jt} - hc_{jt-1}\right) + \upsilon \log\left(\frac{m_{jt}}{p_t}\right) - \varphi_t \psi \frac{l_{jt}^{1+\vartheta}}{1+\vartheta} \right\}$$

where β is the discount factor, h is the parameter that controls habit persistence, ϑ is the inverse of Frisch labor supply elasticity, d_t is a shock to intertemporal preference with the law of motion:

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t} \text{ where } \varepsilon_{d,t} \sim \mathcal{N}(0,1),$$

and φ_t is a labor supply shock with the law of motion:

$$\log \varphi_t = \rho_\varphi \log \varphi_{t-1} + \sigma_\varphi \varepsilon_{\varphi,t} \text{ where } \varepsilon_{\varphi,t} \sim \mathcal{N}(0,1).$$

Households trade on the whole set of Arrow-Debreu securities, contingent on idiosyncractic and aggregate events. Our notation a_{jt+1} indicates the amount of those securities that pay one unit of consumption in event $\omega_{j,t+1,t}$ purchased by household j at time t at (real) price $q_{jt+1,t}$. To save on notation, we drop the explicit dependence on the event. Summing over different individual assets, we can price securities contingent only on aggregate states. Households also hold an amount b_{jt} of government bonds that pay a nominal gross interest rate of R_t and invest x_t. Then, the $j-th$ household's budget constraint is given by:

$$\begin{aligned} & c_{jt} + x_{jt} + \frac{m_{jt}}{p_t} + \frac{b_{jt+1}}{p_t} + \int q_{jt+1,t} a_{jt+1} d\omega_{j,t+1,t} \\ = \; & w_{jt} l_{jt} + \left(r_t u_{jt} - \mu_t^{-1} \Phi\left[u_{jt}\right] \right) k_{jt-1} + \frac{m_{jt-1}}{p_t} + R_{t-1} \frac{b_{jt}}{p_t} + a_{jt} + T_t + \digamma_t \end{aligned}$$

where w_{jt} is the real wage, r_t the real rental price of capital, $u_{jt} > 0$ the intensity of use of capital, $\mu_t^{-1} \Phi\left[u_{jt}\right]$ is the physical cost of the use of capital in resource terms, μ_t is an investment-specific technological shock to be described momentarily, T_t is a lump-sum transfer, and $\digamma_t$ are the profits of the firms in the economy. We assume that $\Phi\left[1\right] = 0$, Φ' and $\Phi'' > 0$.

Investment x_{jt} induces a law of motion for capital

$$k_{jt} = \left(1-\delta\right) k_{jt-1} + \mu_t \left(1 - V\left[\frac{x_{jt}}{x_{jt-1}}\right] \right) x_{jt}$$

where δ is the depreciation rate and $V\left[\cdot\right]$ is a quadratic adjustment cost function such that $V\left[\Lambda_x\right] = 0$, where Λ_x is the growth rate of investment along the balance growth path. Note our capital timing: we index capital by the time its level is decided. The investment-specific technological shock follows an autoregressive process:

$$\mu_t = \mu_{t-1} \exp\left(\Lambda_\mu + z_{\mu,t}\right) \text{ where } z_{\mu,t} = \sigma_\mu \varepsilon_{\mu,t} \text{ and } \varepsilon_{\mu,t} \sim \mathcal{N}(0,1)$$

The first order conditions with respect to c_{jt}, b_{jt}, u_{jt}, k_{jt}, and x_{jt} are:

$$\begin{aligned} d_t \left(c_{jt} - h c_{jt-1}\right)^{-1} - b\beta \mathbb{E}_t d_{t+1} \left(c_{jt+1} - h c_{jt}\right)^{-1} &= \lambda_{jt}, \\ \lambda_{jt} &= \beta \mathbb{E}_t \{ \lambda_{jt+1} \frac{R_t}{\Pi_{t+1}} \}, \\ r_t &= \mu_t^{-1} \Phi'\left[u_{jt}\right], \\ q_{jt} &= \beta \mathbb{E}_t \left\{ \frac{\lambda_{jt+1}}{\lambda_{jt}} \left(\left(1-\delta\right) q_{jt+1} + r_{t+1} u_{jt+1} - \mu_{t+1}^{-1} \Phi\left[u_{jt+1}\right] \right) \right\}, \text{ and} \\ 1 = q_{jt} \mu_t \left(1 - V\left[\frac{x_{jt}}{x_{jt-1}}\right] - V'\left[\frac{x_{jt}}{x_{jt-1}}\right] \frac{x_{jt}}{x_{jt-1}} \right) &+ \beta \mathbb{E} q_{jt+1} \mu_{t+1} \frac{\lambda_{jt+1}}{\lambda_{jt}} V'\left[\frac{x_{jt+1}}{x_{jt}}\right] \left(\frac{x_{jt+1}}{x_{jt}}\right)^2, \end{aligned}$$

where λ_{jt} is the lagrangian multiplier associated with the budget constraint and q_{jt} is the marginal Tobin's Q, the lagrangrian multiplier associated with the investment adjustment constraint normalized by λ_{jt}.

The first order condition with respect to labor and wages is more involved. The labor employed by intermediate good producers to be described below is supplied by a representative, competitive firm that hires the labor supplied by each household j. The labor supplier aggregates the differentiated labor of households with the production function:

$$l_t^d = \left(\int_0^1 l_{jt}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \tag{2}$$

where η controls the elasticity of substitution among different types of labor and l_t^d is the aggregate labor demand.

The labor "packer" maximizes profits subject to the production function (2), taking as given all differentiated labor wages w_{jt} and the wage w_t. From his maximization problem we get:

$$l_{jt} = \left(\frac{w_{jt}}{w_t} \right)^{-\eta} l_t^d \qquad \forall j \tag{3}$$

Then, to find the aggregated wage, we use again the zero profit condition $w_t l_t^d = \int_0^1 w_{jt} l_{jt} dj$ to deliver:

$$w_t = \left(\int_0^1 w_{jt}^{1-\eta} dj \right)^{\frac{1}{1-\eta}}.$$

Households set their wages following a Calvo's setting. In each period, a fraction $1-\theta_w$ of households can change their wages. All other households can only partially index their wages by past inflation. Indexation is controlled by the parameter $\chi_w \in [0,1]$. This implies that if the household cannot change its wage for τ periods, her normalized wage after τ periods is $\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} w_{jt}$.

Since we assume complete markets and separable utility in labor (see Erceg *et al.*, 2000), we will concentrate on a symmetric equilibrium where $c_{jt} = c_t$, $u_{jt} = u_t$, $k_{jt-1} = k_t$, $x_{jt} = x_t$, $\lambda_{jt} = \lambda_t$, $q_{jt} = q_t$, and $w_{jt}^* = w_t^*$. In anticipation of that equilibrium, and after a fair amount of manipulation, we arrive at the recursive equations:

$$f_t = \frac{\eta-1}{\eta} \left(w_t^* \right)^{1-\eta} \lambda_t w_t^\eta l_t^d + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta-1} f_{t+1}$$

and:

$$f_t = \psi d_t \varphi_t \left(\frac{w_t}{w_t^*}\right)^{\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta(1+\vartheta)} f_{t+1}.$$

that determine the evolution of wages.

Then, in every period, a fraction $1-\theta_w$ of households set w_t^* as their wage, while the remaining fraction θ_w partially index their price by past inflation. Consequently, the real wage index evolves:

$$w_t^{1-\eta} = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta} w_{t-1}^{1-\eta} + (1-\theta_w)\, w_t^{*1-\eta}.$$

4.2. The Final Good Producer

There is one final good produced using intermediate goods with the following production function:

$$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}}. \tag{4}$$

where ε controls the elasticity of substitution.

Final good producers are perfectly competitive and maximize profits subject to the production function (4), taking as given all intermediate goods prices p_{ti} and the final good price p_t. Following the same steps as for wages, we find the input demand functions associated with this problem are:

$$y_{it} = \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} y_t^d \qquad \forall i,$$

where y_t^d is the aggregate demand and the zero profit condition $p_t y_t^d = \int_0^1 p_{it} y_{it} di$ to deliver:

$$p_t = \left(\int_0^1 p_{it}^{1-\varepsilon} di\right)^{\frac{1}{1-\varepsilon}}.$$

4.3. Intermediate Good Producers

There is a continuum of intermediate goods producers. Each intermediate good producer i has access to a technology represented by a production function:

$$y_{it} = A_t k_{it-1}^{\alpha} \left(l_{it}^d\right)^{1-\alpha} - \phi z_t$$

where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the "packed" labor input rented by the firm, the parameter ϕ corresponds to the fixed cost of production, and where A_t follows:

$$A_t = A_{t-1} \exp\left(\Lambda_A + z_{A,t}\right) \text{ where } z_{A,t} = \sigma_A \varepsilon_{A,t} \text{ and } \varepsilon_{A,t} \sim \mathcal{N}(0,1)$$

The fixed cost ϕ is scaled by the variable $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$. We can think of z_t as a weighted index of the two technology levels A_t and μ_t, where the weight is given by the share of capital in the production function. The product ϕz_t guarantees that economic profits are roughly equal to zero in the steady state. Also, we rule out the entry and exit of intermediate good producers. Note that z_t evolves over time as $z_t = z_{t-1} \exp\left(\Lambda_z + z_{z,t}\right)$ where $z_{z,t} = \frac{z_{A,t} + \alpha z_{\mu,t}}{1-\alpha}$ and $\Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha}$. We will see below that Λ_z will be the mean growth rate of the economy.

Intermediate goods producers solve a two-stages problem. First, given w_t and r_t, they rent l_{it}^d and k_{it-1} in perfectly competitive factor markets in order to minimize real costs, which implies a marginal cost of:

$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha} \left(\frac{1}{\alpha}\right)^{\alpha} \frac{w_t^{1-\alpha} r_t^{\alpha}}{A_t}$$

The marginal cost does not depend on i: all firms receive the same shocks and all firms rent inputs at the same price.

Second, intermediate good producers choose the price that maximizes discounted real profits. To do so, they consider that they are under the same pricing scheme as households. In each period, a fraction $1-\theta_p$ of firms can change their prices. All other firms can only index their prices by past inflation. Indexation is controlled by the parameter $\chi \in [0,1]$, where $\chi = 0$ is no indexation and $\chi = 1$ is total indexation.

The problem of the firms is then:

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \left(\beta \theta_p\right)^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left\{ \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau} \right) y_{it+\tau} \right\}$$

subject to

$$y_{it+\tau} = \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} y_{t+\tau}^d,$$

where the marginal value of a dollar to the household is treated as exogenous by the firm. Since we have complete markets in securities, this marginal value is constant across households and, consequently, $\lambda_{t+\tau}/\lambda_t$ is the correct valuation on future profits.

We write the solution of the problem in terms of two recursive equations in g_t^1 and g_t^2:

$$g_t^1 = \lambda_t mc_t y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}}\right)^{-\varepsilon} g_{t+1}^1$$
$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}}\right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) g_{t+1}^2$$

and $\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$ where:

$$\Pi_t^* = \frac{p_t^*}{p_t}$$

Given Calvo's pricing, the price index evolves:

$$p_t^{1-\varepsilon} = \theta_p \left(\Pi_{t-1}^\chi\right)^{1-\varepsilon} p_{t-1}^{1-\varepsilon} + (1-\theta_p)\, p_t^{*1-\varepsilon}$$

or, dividing by $p_t^{1-\varepsilon}$,

$$1 = \theta_p \left(\frac{\Pi_{t-1}^\chi}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\, \Pi_t^{*1-\varepsilon}$$

4.4. The Government

The government sets the nominal interest rates according to the Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}}\right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t) \tag{5}$$

through open market operations that are financed with lump-sum transfers T_t to ensure that the government budget is balanced period by period. The variable Π represents the target levels of inflation (equal to inflation in the steady state), R steady-state gross return of capital, and Λ_{y^d} the steady-state gross growth rate of y_t^d. With a bit of abuse of language, we will refer to the term $\frac{y_t^d}{y_{t-1}^d}/\Lambda_{y^d}$ as the *growth gap*. The term m_t is a random shock to monetary policy that follows $m_t = \sigma_m \varepsilon_{mt}$ where ε_{mt} is distributed according to $\mathcal{N}(0,1)$. We introduce the previous period interest rate, R_t, to match the smooth profile of the interest rate over time observed in the U.S.

4.5. Aggregation

To close the model, we derive an aggregate supply equation. First, we begin with the aggregate demand:

$$y_t^d = c_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1}$$

Then, using the production function for intermediate good producers, the fact that all the firms have the same optimal capital-labor ratio, and market clearing (both in the output and the input markets), we have:

$$y_t^d = \frac{A_t\left(u_t k_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - \phi z_t}{v_t^p}$$

where:

$$v_t^p = \int_0^1 \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} di$$

is the aggregate loss of efficiency induced by price dispersion. By the properties of the index under Calvo's pricing:

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + \left(1-\theta_p\right)\Pi_t^{*-\varepsilon}.$$

Finally, we integrate labor demand over all households j to get:

$$\int_0^1 l_{jt}dj = l_t = \int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj l_t^d$$

where l_t is the aggregate labor supply of households. Hence if we define:

$$v_t^w = \int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj$$

we have:

$$l_t^d = \frac{1}{v_t^w} l_t$$

and:

$$v_t^w = \theta_w \left(\frac{w_{t-1}}{w_t}\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta} v_{t-1}^w + \left(1-\theta_w\right)\left(\Pi_t^{w*}\right)^{-\eta}.$$

4.6. Equilibrium

A definition of equilibrium in this economy is standard and the equations that characterize it are determined by the first order conditions of the household, the first order conditions of the firms, the Taylor rule of the government, and market clearing.

To undertake our quantitative analysis, we must approximate the equilibrium dynamics of the economy. Ours is a large model (even the version without parameter drifting has 19 state variables). Moreover, we will need to solve the model repeatedly during our estimation process. We have argued elsewhere (Fernández-Villaverde, Rubio-Ramírez, and Santos, 2006) that there is much to be gained from a nonlinear estimation of the model, both in terms of accuracy and in terms of identification. As we will discuss later, this is particularly true if we want to allow the agents in the economy to insure themselves against future changes in the parameters of the model. Hence, we need a nonlinear solution method that is fast and accurate. In previous work (Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006), we have found that a second order perturbation around the deterministic steady-state of the model fulfills the previous desiderata .

But before solving the model, we need to clear up some technical issues. First, since we have growth induced by technological change, most of the variables are growing in average. To achieve the right accuracy in the computation, we make the variables stationary and solve the model in those transformed variables. Hence, we define $\widetilde{c}_t = \frac{c_t}{z_t}$, $\widetilde{\lambda}_t = \lambda_t z_t$, $\widetilde{r}_t = r_t \mu_t$, $\widetilde{q}_t = q_t \mu_t$, $\widetilde{x}_t = \frac{x_t}{z_t}$, $\widetilde{w}_t = \frac{w_t}{z_t}$, $\widetilde{w}_t^* = \frac{w_t^*}{z_t}$, $\widetilde{k}_t = \frac{k_t}{z_t \mu_t}$, and $\widetilde{y}_t^d = \frac{y_t^d}{z_t}$. Also note that $\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{w^*} = \Lambda_{y^d} = \Lambda_z$. Second, we need to choose functional forms for $\Phi[\cdot]$ and $V[\cdot]$. For $\Phi[u]$ we pick $\Phi[u] = \Phi_1 (u-1) + \frac{\Phi_2}{2}(u-1)^2$. Since in the steady state we have $u = 1$, then $\widetilde{r} = \Phi'[1] = \Phi_1$ and $\Phi[1] = 0$. The investment adjustment cost function is $V\left[\frac{x_t}{x_{t-1}}\right] = \frac{\kappa}{2}(\frac{x_t}{x_{t-1}} - \Lambda_x)^2$. Then, along the balanced growth path, $V[\Lambda_x] = V'[\Lambda_x] = 0$.

We will perform our perturbation in logs. For each variable var_t, we define $\widehat{var}_t = \log var_t - \log var$, as the log deviation with respect to the steady state. Then, the states of the model $\overline{S}_t$ are given by:

$$\overline{S}_t = \begin{pmatrix} \widehat{\Pi}_{t-1}, \widehat{\widetilde{w}}_{t-1}, \widehat{g}_{t-1}^1, \widehat{g}_{t-1}^2, \widehat{\widetilde{k}}_{t-1}, \widehat{R}_{t-1}, \widehat{\widetilde{y}}_{t-1}^d, \widehat{\widetilde{c}}_{t-1}, \widehat{v}_{t-1}^p, \widehat{v}_{t-1}^w, \\ \widehat{\widetilde{q}}_{t-1}, \widehat{\widetilde{f}}_{t-1}, \widehat{\widetilde{x}}_{t-1}, \widehat{\widetilde{\lambda}}_{t-1}, \widehat{\widetilde{z}}_{t-1}, z_{\mu,t-1}, \widehat{d}_{t-1}, \widehat{\varphi}_{t-1}, z_{A,t-1} \end{pmatrix}',$$

and the exogenous shocks are $\varepsilon_t = (\varepsilon_{\mu,t}, \varepsilon_{d,t}, \varepsilon_{\varphi,t}, \varepsilon_{A,t}, \varepsilon_{m,t})'$.

As a first step, we parameterize the matrix of variances-covariances of the exogenous shocks as $\overline{\Omega}(\chi) = \chi\Omega$, where $\overline{\Omega}(1) = \Omega$, a diagonal matrix. However, nothing really depends on that assumption, and we could handle an arbitrary matrix of variances-covariances. Then, we take a perturbation solution around $\chi = 0$, i.e., around the deterministic stead- state of the model.

From the output of the perturbation, we can build the law of motion for the states:

$$\overline{S}_{t+1} = \Psi_{s1}\left(\overline{S}'_t, \varepsilon'_t\right)' + \frac{1}{2}\left(\overline{S}'_t, \varepsilon'_t\right)\Psi_{s2}\left(\overline{S}'_t, \varepsilon'_t\right)' + \Psi_{s3} \tag{6}$$

where Ψ_{s1} is a 1×24 vector and Ψ_{k2} is a 24×24 matrix. The term $\Psi_{s1}\left(\overline{S}'_t, \varepsilon'_t\right)'$ constitutes the linear solution of the model, $\left(\overline{S}'_t, \varepsilon'_t\right)\Psi_{s2}\left(\overline{S}'_t, \varepsilon'_t\right)'$ is the quadratic component, and Ψ_{s3} is a 1×24 vector of constants added by the second order approximation that corrects for precautionary behavior. Some of the entries of the matrices Ψ_{si} will be zero.

From the same output, we find the law of motion for the observables

$$\mathcal{Y}^T = \left(\triangle\log\mu_t^{-1}, \triangle\log y_t, \triangle\log x_t, \triangle\log l_t, \log\Pi_t, \log R_t\right)'$$

Now, define:

$$S_t = \left(\overline{S}'_t, \overline{S}'_{t-1}, \varepsilon'_{t-1}\right).$$

We keep track of the past states, $\overline{S}'_{t-1}$, because some of the observables in the measurement equation below will appear in first differences. Then, we get to the observation equation:

$$\mathcal{Y}^T = \Psi_{o1}\left(S'_t, \varepsilon'_t\right)' + \frac{1}{2}\left(S'_t, \varepsilon'_t\right)\Psi_{o2}\left(S'_t, \varepsilon'_t\right)' + \Psi_{o3} \tag{7}$$

where Ψ_{o1} and Ψ_{o3} 1×48 matrices and Ψ_{o2} is a 48×48 matrix.

While the law of motion for states is unique (or at least equivalent to a class of different choices of states, all of which have the same implications for the dynamics of the model), the observation equation depends on what we assume the researcher actually observes. In our case, we have chosen the first differences of the relative price of investment, output, investment, and hours, and the log of inflation and the interest rate. Unfortunately, we do not know much about the right choice of observables and how they may affect our estimation results (for one of the few articles on this topic, see Boivin and Giannoni, 2006).

4.7. The Likelihood Function

Equations (6) and (7) constitute the state space representation of our model. One convenient properties of this representation is that we can exploit it to evaluate the likelihood of a DSGE model, an otherwise challenging task. The likelihood, $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$, is the probability that the model assigns to a sequence of realizations of the observable $\mathbb{Y}^T$ given parameter values:

$$\Psi = \left\{\beta, h, \upsilon, \vartheta, \delta, \eta, \varepsilon, \alpha, \phi, \theta_w, \chi_w, \theta_p, \chi_p, \Phi_2, \gamma_R, \gamma_y, \gamma_\Pi, \Pi, \Lambda_\mu, \Lambda_A, \rho_d, \rho_\varphi, \sigma_\mu, \sigma_d, \sigma_A, \sigma_m, \sigma_\varphi\right\}.$$

Note that Φ_1 is not included in Ψ because it is a function of all the other parameters in the economy to ensure that $\widetilde{r} = \Phi_1$. With $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$, we can estimate Ψ by maximizing the likelihood or by combining it with a prior density for the model parameter to form a posterior distribution.

But how do we evaluate the likelihood $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$? Given the Markov structure of our state space representation, we begin by factorizing the likelihood function as:

$$\mathcal{L}\left(\mathbb{Y}^T;\Psi\right) = \prod_{t=1}^{T} \mathcal{L}\left(\mathbb{Y}_t|\mathbb{Y}^{t-1};\Psi\right)$$

Then, conditioning on the states:

$$\mathcal{L}\left(\mathbb{Y}^T;\Psi\right) = \int \mathcal{L}\left(\mathbb{Y}_1|S_0;\Psi\right) dS_0 \prod_{t=2}^{T} \int \mathcal{L}\left(\mathbb{Y}_t|S_t;\Psi\right) p\left(S_t|\mathbb{Y}^{t-1};\Psi\right) dS_t \tag{8}$$

If we know S_t, computing $\mathcal{L}\left(\mathbb{Y}_t|S_t;\Psi\right)$ is relatively easy using the measurement equation (7) of the state space representation. Conditional on S_t, equation (7) is a change of variables from ε_t to $\mathcal{Y}^T$ and, hence, we can use it to compute probabilities using the change of variable formula. Similarly, if we know S_0, we can employ the transition (6) and measurement equation (7) of the state space representation of the model to compute $\mathcal{L}\left(\mathbb{Y}_1|S_0;\Psi\right)$.

Consequently, knowledge of the sequence $\{p\left(S_t|\mathbb{Y}^{t-1};\Psi\right)\}_{t=1}^{T}$ and of $p\left(S_0;\Psi\right)$ allows the evaluation of the likelihood of the model. Evaluating (or at least drawing from) $p\left(S_0;\Psi\right)$ is usually straightforward, although often costly (Santos and Peralta-Alva, 2005). More involved is to characterize the sequence of conditional distributions $\{p\left(S_t|\mathbb{Y}^{t-1};\Psi\right)\}_{t=1}^{T}$ and to compute the integrals in (8).

An algorithm for doing so (but not the only one!; see the technical appendix to Fernández-

Villaverde and Rubio-Ramírez, 2007 for alternatives) is to use a simulation technique known as the particle filter (see the review in Doucet, de Freitas, and Gordon, 2001). Fernández-Villaverde and Rubio-Ramírez (2005 and 2007) have shown that the particle filter can be successfully applied to the estimation of nonlinear and/or non-normal DSGE models. The particle filter is a sequential Monte Carlo method that replaces the $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$ by an empirical distribution of draws generated by simulation. The bit of magic of the particle filter is that the simulation is generated through a procedure known as sequential importance resampling (SIR). SIR guarantees that the Monte Carlo method achieves sufficient accuracy in a reasonable amount of time, something that cannot be achieved without resampling (Arulampalam *et al.*, 2002). The appendix describes in further detail the working of the particle filter.

4.8. A Bayesian Approach

We will confront our model with the data using Bayesian methods. The Bayesian paradigm is a powerful and flexible perspective for the estimation of DSGE models (see the survey by An and Schorfheide, 2006). First, Bayesian analysis is a coherent approach to inference based on a clear set of axioms. Second, the Bayesian approach handles in a natural way misspecification and lack of identification, both serious concerns in the estimation of DSGE models (Canova and Sala, 2006). Moreover, it has desirable small sample and asymptotic properties, even when evaluated by classical criteria (Fernández-Villaverde and Rubio-Ramírez, 2004). Third, the use of priors is a flexible procedure to introduce presample information that the researcher may have and to reduce the dimensionality problem associated with number of parameters. This property will be especially attractive in our application, since parameter drifting will increase the practical number of dimensions of our model.

The Bayesian approach combines the likelihood of the model $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$ with a prior density for the parameters $p(\Psi)$ to form a posterior:

$$p\left(\Psi|\,\mathbb{Y}^T\right) \propto \mathcal{L}\left(\mathbb{Y}^T;\Psi\right)p(\Psi)$$

The posterior summarizes the uncertainty regarding the parameters, and it can be used for point estimation. For example, under a quadratic loss function, our point estimates will be the mean of the posterior.

Since the posterior is also difficult to characterize, we generate draws from it using a Metropolis-Hastings algorithm. We use the resulting empirical distribution to obtain point estimates, standard deviations, etc. We describe this algorithm in the appendix.

5. Parameter Drifting

Now we can deal with parameter drifting. Since the extension to other cases of parameter drifting is straightforward, we present only one example of drift within our model.

Motivated by the first example in section 3, we will investigate the situation where the Taylor rule is specified as:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_{Rt}} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_{\Pi t}} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}}\right)^{\gamma_{yt}}\right)^{1-\gamma_{Rt}} \exp\left(m_t\right) \tag{9}$$

Note the difference with the specification in (5): now the elasticities of the response of the interest rate $\left\{\gamma_{Rt}, \gamma_{\Pi t}, \gamma_{yt}\right\}$ are indexed by time.

We will postulate that the parameters follow an AR(1) in logs to ensure that the parameter is positive:

$$\log \gamma_{Rt} = (1-\rho_R) \log \overline{\gamma}_R + \rho_R \log \gamma_{Rt-1} + \varepsilon_{Rt} \tag{10}$$

$$\log \gamma_{Rt} = (1-\rho_\Pi) \log \overline{\gamma}_\Pi + \rho_\Pi \log \gamma_{\Pi t-1} + \varepsilon_{\Pi t} \tag{11}$$

$$\log \gamma_{Rt} = \left(1-\rho_y\right) \log \overline{\gamma}_y + \rho_y \log \gamma_{yt-1} + \varepsilon_{yt} \tag{12}$$

where $\{\varepsilon_{Rt}, \varepsilon_{\Pi t}, \varepsilon_{yt}\}$ are i.i.d. normal shocks and Q is a 3×3 matrix of covariances.[2] We allow for arbitrary correlation in the innovations, since it is plausible that the reasons why the monetary authority becomes more (less) responsive to inflation are the same reasons it will become less (more) responsive to the growth gap. Also, we could generalize the changes in parameters by allowing changes in Π or in the variance of m_t (R and Λ_{y^d} are not chosen by the monetary authority but given by preferences and technology parameters of the model and Π).

[2] The autoregressive coefficients $\left\{\rho_R, \rho_{\Pi t}, \rho_y\right\}$ and the matrix Q become in this formulation the new "structural parameters." We are also skeptic about their true structural nature, but, to avoid the infinite regression problem, we will ignore our doubts for the moment.

Our specification of parameter drift emphasizes the continuity of the change process, in opposition to the discrete changes in the parameters emphasized by the Markov-switching process (Davig and Leeper, 2006a and 2006b). We do not have a strong prior preference for one version or the other. We like our specification because it is parsimonious and easy to handle, and it captures phenomena such the Fed's gradual learning about the behavior of the economy.

According to our favorite interpretation of parameter drifting, we will assume that agents understand that policy evolves over time following (10)-(12). Consequently, they react to it and make their decisions based on the current values of γ_t and on the fact that γ_t will evolve over time.

The drift of the parameters implies that the economy will travel through zones where the Taylor principle is not satisfied. However, this may not necessarily mean that the equilibrium is not unique. In the context of Markov-regime changes in the coefficients of the Taylor rule, Davig and Leeper (2006a) have developed what they call the generalized Taylor principle. Davig and Leeper argue that a unique equilibrium survives if the Taylor rule is sufficiently active when the economy is in the active policy regime or if the expected length of time the economy will be in the nonactive policy regime is sufficiently small. To keep this paper focused, we will not dwell on generating results equivalent to Davig and Leeper's in our environment. Suffice it to say that one further advantage of the Bayesian approach is that we can handle restrictions on the parameter drifting with the use of the priors. For example, we can implement a reflecting boundary on (10) by putting a zero prior to the possibility of violating that boundary. Also, in our empirical analysis, we estimate $\overline{\gamma}_R$ as being bigger than one. This suggests that the Taylor principle will be satisfied, at least on average.

Our formulation of parameter drifting has one important drawback: we do not model explicitly why the parameters change over time. In section 3, we discussed that changes in the policy parameters could be a reflection of changing political priorities or evolving perceptions about the effectiveness of policy. A more complete model would include explicit mechanisms through which we discipline the movement of the parameters over time. Many of those mechanisms can be incorporated into our framework, since we are rather flexible with the type of functional forms for the parameter drift that we can handle.

The model in section 4 carries on except with the modification of (9) and the fact that all the conditional expectations now incorporate the process (10). Thus, the states of the model with parameter drifting are:

$$\overline{S}_t = \left(\begin{array}{c} \widehat{\Pi}_{t-1}, \widehat{\widetilde{w}}_{t-1}, \widehat{g}^1_{t-1}, \widehat{g}^2_{t-1}, \widehat{\widetilde{k}}_{t-1}, \widehat{R}_{t-1}, \widehat{\widetilde{y}}^d_{t-1}, \widehat{\widetilde{c}}_{t-1}, \widehat{v}^p_{t-1}, \widehat{v}^w_{t-1}, \widehat{\widetilde{q}}_{t-1}, \widehat{\widetilde{f}}_{t-1}, \\ \widehat{\widetilde{x}}_{t-1}, \widehat{\widetilde{\lambda}}_{t-1}, \widehat{\widetilde{z}}_{t-1}, z_{\mu,t-1}, \widehat{d}_{t-1}, \widehat{\varphi}_{t-1}, z_{A,t-1}; \gamma_{Rt-1}, \gamma_{\Pi t-1}, \gamma_{yt} \end{array} \right)',$$

where we have included $\gamma_{Rt}, \gamma_{\Pi t}$, and γ_{yt} as three additional states. We will follow the convention of separating drifting parameters from the other states with a ";" since they are an object of interest by themselves. Similarly, we apply the particle filter to evaluate the likelihood of the model and the Metropolis-Hastings to draw from the posterior.

6. Empirical Analysis

This section presents our empirical analysis. First, we report the point estimates of the model when we keep all parameters fixed over the estimation. This estimation sets a natural benchmark for the rest of the study. Second, we discuss the results of an exercise where we allow the parameters of the Taylor rule of the monetary authority to change over time. Third, we analyze the evolution of the parameters that control the level of price and wage rigidities. In the interest of space, we select these two exercises as particularly illustrative of the procedure we propose. However, we could have performed many other exercises within the framework of our methodology.

We estimate the model using six time series for the U.S.: 1) the relative price of investment with respect to the price of consumption, 2) real output per capita growth, 3) real gross investment per capita growth, 4) hours worked per capita, 5) the CPI and 6) the federal funds rate. Our sample goes from 1955:Q1 to 2000:Q4. We stop our sample at the end of 2000 because of the absence of good information on the relative price of investment after that time. To make the observed series compatible with the model, we need to compute both real output and real gross investment in consumption units. For that purpose, we use the relative price of investment defined as the ratio of an investment deflator and a deflator for consumption. The consumption deflator is constructed from the deflators of nondurable goods and services reported in the NIPA. Since the NIPA investment deflators are poorly measured, we rely on the investment deflator constructed by Fisher (2006), a series that ends at 2000:Q4. The appendix provides further information on the construction of the data.

6.1. Point Estimation

Before reporting results, we specify priors for the model's parameters. We adopt flat priors for all parameters. We impose boundary constraints only to make the priors proper and to rule out parameter values that are either incompatible with the model (i.e., a negative value for a variance, Calvo parameters outside the unit interval) or extremely implausible (the response to inflation in the Taylor rule being bigger than 100). The looseness of such constraints is shown by the fact that the simulations performed below never travel even close to those bounds. Also, we fix three parameters, $\{\upsilon, \phi, \Phi_2\}$. The parameter controlling money demand υ is irrelevant for equilibrium dynamics because the government will supply as much money as required to implement the nominal interest rate determined by the Taylor rule. We fix the parameter ϕ to zero, since we do not have information on pure profits by firms (in the absence of entry/exit of firms, there are no serious implications for equilibrium dynamics). The parameter of the investment adjustment cost, Φ_2, is set to 0.001 because it was difficult to identify.

Our choice of flat priors is motivated by the fact that, with this prior, the posterior is proportional to the likelihood function.[3] Consequently, our Bayesian results can be interpreted as a classical exercise where the mode of the likelihood function (the point estimate under an absolute value loss function for estimation) is the maximum likelihood estimate. Also, a researcher who prefers to use more informative priors can always reweight the draws from the posterior to accommodate his favorite priors (Geweke, 1998).[4] We repeated our estimation with an informative prior without finding important differences in the results.

Table 6.1 summarizes our results. Most of our point estimates coincide with the typical findings of other estimation exercises. Hence, we comment only on a few of them. We have a high degree of habit persistence, h is 0.88, and we have a Frisch elasticity of labor supply of 0.74 (1/1.36), well within the bounds of findings of the recent microeconomic literature (Browning, Hansen, and Heckman, 1999). The estimates of elasticities of substitution ε and η are around 8, implying average mark-ups of around 14 percent.

[3]There is a small qualifier: the bounded support of the priors. We can fix this small difference by thinking about those bounds as frontiers of admissible parameter values in a classical perspective.

[4]We do not argue that our flat priors are uninformative. After a reparameterization of the model, a flat prior may become highly curved. Also, if we wanted to compare the model with, for example, a VAR, we would need to elicit our priors more carefully.

TABLE 6.1: Point Estimates

Parameter	Point Estimate	Parameter	Point Estimate
β	0.9999	γ_R	0.7900
h	0.8773	γ_y	0.1904
ψ	8.9420	γ_Π	1.2596
ϑ	1.3586	Π	1.0078
δ	0.0149	Λ_μ	0.0100
α	0.2550	Λ_A	0.0005
ε	7.9570	ρ_d	0.9506
η	7.9650	ρ_φ	0.9420
κ	7.6790	σ_μ	0.1010
θ_p	0.9067	σ_d	0.0600
χ_p	0.1505	σ_A	0.0072
θ_w	0.4506	σ_m	0.0030
χ_w	0.8492	σ_φ	0.0700

The Calvo parameter for price adjustment, θ_p, is a relatively high 0.91, while the indexation level χ_p, is 0.15. It is tempting to compare the results from our exercise with microeconomic evidence to determine the average duration of prices (Bils and Klenow, 2004, or Nakamura and Steinsson, 2006). However, the comparison is difficult because we have partial indexation: prices change every quarter for all producers, a fraction θ_p because producers can reoptimize and a fraction $1 - \theta_p$ because of indexation. The Calvo parameter for wage adjustment, θ_w, is 0.45, while the indexation, χ_w, is 0.85.

The policy parameters $\{\gamma_R, \gamma_\Pi, \gamma_y, \Pi\}$ are also quite standard. The Fed smooths the interest rate over time (γ_R is estimated to be 0.79), and responds actively to inflation (γ_R is 1.25) and weakly to the output growth gap (γ_y is 0.19). We estimate that the Fed has a target for quarterly inflation of 0.78 percent (or around 3 percent yearly).

The growth rates of the investment-specific technological change, Λ_μ, and of the neutral technology, Λ_A, imply that most of the growth in the U.S. economy (83 percent) is induced by improvements in the capital-producing technology. This result corroborates the importance of modelling this biased technological change for understanding growth and fluctuations that Greenwood, Herkowitz, and Krusell (1997 and 2000) have so forcefully defended. The estimated long-run growth rate of the economy, $(\Lambda_A + \alpha\Lambda_\mu) / (1-\alpha)$ is 0.4 percent per quarter, or 1.6 percent annually, roughly the observed mean in the sample. Also, the standard deviation σ_μ is much higher than σ_A.

Our estimation serves different roles. First, it validates our model as a promising laboratory for our exercises with parameter drifting. Since in the benchmark case we obtain results compatible with the literature and with the basic growth properties of the U.S. economy, we know that the results with parameter drifting will indeed come from that feature of the estimation. Second, we use our point estimates to initialize the parameters in the exercises with parameter drifting.

In the next two subsections, we will report the results of estimating the model when we allow one parameter to vary at a time. We do this for convenience. First, allowing several parameters to move simultaneously makes the computation and estimation of the model much more costly. Second, the information in the sample is limited, and it is difficult to obtain stable estimates otherwise. Third, especially in our second exercise, our objective is not so much to have the richest possible model to fit the data well but to show that as soon as you let parameters change over time, strong signs of misspecification appear. We will continue the exploration of joint moves of parameters in the near future.

6.2. Evolution of Policy Parameters

Our first exercise is to study the evolution of the policy parameters in the Taylor rule. This investigation evaluates how much evidence there is in the data of a changing monetary policy over time. As we discussed in section 3, the literature has extensively debated this topic (Clarida, Galí, and Gertler, 2000, Cogley and Sargent, 2001, Lubick and Schorfheide, 2004, Sims and Zha, 2006, Boivin, 2006, just to cite a few papers). However, the empirical methods applied so far are unsatisfactory because they rely either on divisions of the sample that do not let the agents in the model forecast the changes in policy or on the estimation of reduced forms.

Arguably, the most interesting parameter is $\gamma_{\Pi t-1}$, since this parameter controls how aggressively the monetary authority responds to inflation. In addition, $\gamma_{\Pi t-1}$ is intimately linked with the issue of multiplicity of equilibria and the possibility of monetary policy being a source of instability. Figure 6.2.1 plots our point estimate of the evolution of $\gamma_{\Pi t-1}$ over time. We report the smoothed values of $\gamma_{\Pi t-1}$ using the whole sample (Godsill, Doucet, and West, 2004). We find it convenient, for expositional purposes, to eliminate some of the quarter-to-quarter variation of the parameter. To accomplish this goal, in figure 6.2.2, we graph the trend of the evolution of the parameter where we compute the trend using a Hodrick-Prescott filter. We emphasize that this trend is only a device to read the graph more clearly and lacks a formal statistical interpretation.

In both figures 6.2.1 and 6.2.2, we see how $\gamma_{\Pi t-1}$ starts low, slightly above 1 during the 1950s, 1960s, and early 1970s, with periods when it was even below 1. However, in the mid-1970s, and especially after Volcker's appointment as Chairman of the Board of Governors, $\gamma_{\Pi t-1}$ soared. The response to inflation reached its peak in the early 1980s, where it was as high as 6 in one quarter. After that, $\gamma_{\Pi t-1}$ slowly decreases during the 1990s, perhaps reflecting the Fed's more permissive attitude to accommodate the strong productivity growth of the Internet boom.

Since our model has parameter drifting, it is not straightforward to compare these numbers with estimates obtained in fixed-parameter models. However, we clearly confirm the findings of Clarida, Galí, and Gertler (2000), Lubick and Schorfheide (2004), and Boivin (2006) that monetary policy became much more active in the last 25 years. Our finding is also consistent with the results of figure 12 in Cogley and Sargent (2001), where they trace the evolution of the activism coefficient as measured by a parameter-drifting VAR.

Another parameter of importance is the inflation target of the monetary authority, Π. Histories like those in Taylor (1998), Sargent (1999), or Primiceri (2006) explain that the inflation target may have changed over time as a reflection of the Fed's varying beliefs about the trade-off between unemployment and inflation. Figure 6.2.3 plots the evolution of the target over time. From the start of the sample until the early 1970s and, later, for the 1990s, Π hovers around 1.004 or, in annual terms, around 1.6 percent. This number is close to the informal target or comfort zone that describes the Fed's behavior according to many commentators. During the intermediate years, the inflation target increases, reflecting perhaps the views the Fed had about the possibility of exploiting the Phillips curve or illustrating the information lags regarding the changing features of the economy emphasized by Orphanides (2002). We also find intriguing the similarity of figure 6.2.3 to Romer and Romer's (2002) hypothesis, based on narrative accounts and internal greenbook forecast of the Fed, that monetary policy in the U.S. has gone through a long cycle of moderation, aggressiveness, and renewed temperance.

Our estimates of the evolution of the inflation target provide a reality check on our procedure. In our figure 6.2.4, we plot the inflation target versus the realized inflation and, in figure 6.2.5, the HP-trend of both series. If the estimation is working properly, part of the variation in the inflation target needs to be accounted for, in a purely mechanical fashion, by changes in inflation. That is precisely what we see: as inflation increases and then falls during the late 1960s and the 1970s, the target inflation estimated goes up and down (although in a smaller quantity to leave room for other shocks to the economy).

Consequently, we trust our results not only because they come from the estimation of a coherent DSGE model, but also because they are consistent with the findings of the existing literature that uses alternative estimation procedures, with narrative accounts of monetary policy, and with the reality check explained above.

6.3. Evolution of Price and Wage Rigidities

A key set of parameters in the New Keynesian model we are estimating are those determining the extent of price and wage rigidities, $\left\{\theta_p, \chi_p, \theta_w, \chi_w\right\}$. These four parameters generate the nominal rigidity in the economy required to match the impulse response functions documented by VARs (Christiano, Eichenbaum, and Evans, 2005).

Given their importance in the model, it is unfortunate that these parameters have only a tenuous link with microeconomic foundations. Even if the Calvo adjustment probabilities are the reduced form of a convex adjustment cost model, the environment that produces this reduced form has changed over the years in our sample: we have gone from periods of high inflation and low response of the monetary authority to raising prices to periods of much lower inflation and a much more aggressive attitude toward inflation by the Fed. Also, the U.S. economy has experienced a notable level of deregulation, increasing competition in internal markets from international trade, and lower unionization rates. The justification of the indexation parameters or their relation to the Calvo adjustment probabilities is even less clear. Why do agents index their prices and wages? And if they do, to which quantity? Past inflation? Current inflation? Steady-state inflation? Wage inflation?

Consequently, it is natural to examine the possibility that the parameters $\left\{\theta_p, \chi_p, \theta_w, \chi_w\right\}$ drift over time, both as a measure of how strong nominal rigidities have been in each different moment and as a tool to assess the extent of possible misspecification of the model along this dimension.

As in the case of policy parameters, we specify an AR(1) as the law of motion for the parameters:

$$\begin{aligned}
\theta_{pt} &= \left(1-\rho_{\theta_p}\right)\overline{\theta}_p + \rho_{\theta_p}\theta_{pt-1} + \varepsilon_{\theta_p t} \\
\chi_{pt} &= \left(1-\rho_{\chi_p}\right)\overline{\chi}_p + \rho_{\chi}\chi_{pt-1} + \varepsilon_{\chi_p t} \\
\theta_{wt} &= \left(1-\rho_{\theta_w}\right)\overline{\theta}_w + \rho_{\theta_w}\theta_{wt-1} + \varepsilon_{\theta_w t} \\
\chi_{wt} &= \left(1-\rho_{\theta_w}\right)\overline{\chi}_w + \rho_{\chi}\chi_{wt-1} + \varepsilon_{\chi_w t}
\end{aligned}$$

where $\left\{\varepsilon_{\theta_p t}, \varepsilon_{\chi_p t}, \varepsilon_{\theta_w t}, \varepsilon_{\chi_w t}\right\}$ are i.i.d. normal shocks.

We report first the experiment where we let θ_{pt}, the Calvo parameter of price changes, evolve over time. We find it more informative (and more directly comparable to the micro evidence) to report the average duration of the spell before the producers can reoptimize, $1\left(/1-\theta_{pt}\right)$, in quarter terms. Figure 6.3.1 plots that duration while figure 6.3.2 plots the HP-trend and, for comparison purposes, the HP-trend of the CPI. The figures reveal a clear pattern: average duration was high in the late 1950s, dropped quickly in the 1960s, and only started to pick up in the late 1970s, continuing with an upward trend until today.

Interestingly enough, the changes in the average duration of the spell before the producers can reoptimize are strongly correlated with changes in inflation. In figure 6.3.2 we see how times of increasing trend inflation (late 1960s, 1970s) are times of falling average duration and vice versa-how times of decreasing trend inflation (the 1980s and the 1990s) are times of increasing average duration.

Our second experiment regarding price rigidities is with χ_{pt}, the parameter that controls price indexation. Figure 6.3.3 plots the evolution of the parameter over the sample and figure 6.3.4 its HP-trend (again, with the HP-trend of the CPI superimposed). Indexation evolves in an opposite way to price duration: it starts low in the 1950s and 1960s, but raises very strongly during the late 1960s. Then, it drops dramatically in the mid-1970s, and stays low over the next 20 years (except for a temporary increase in the early 1980s). In the last part of the sample, during the 1990s, χ_{pt} steadily drops. The drop in indexation the second half of the 1970s may be accounted for by firms switching to more often optimal price adjustments and less automatic pricing rules. Firms were perhaps induced by the volatile inflation of those years, which made partial indexation a costly option. Mechanically, our estimation finds less indexation because inflation is less persistent in the 1970s.

We find illuminating to combine the evolution of the Calvo parameter θ_{pt} and of indexation χ_{pt}. We do so in figure 6.3.5 (for their levels) and in figure 6.3.6 (for their HP-trends). The comparison of both parameters shows that periods of high price rigidities are also periods of low indexation. The converse is also true, except for the mid 1970s. This result points out that adding indexation as an ad hoc procedure to increase the level of inflation inertia may hide important dynamics in price adjustments.

We repeat our two experiments for wages. Figure 6.3.7 (in levels) and figure 6.3.8 (in HP-trends, with inflation superimposed) plot the evolution of the average duration of the spell before the worker can reoptimize their wage, $1\left(/1-\theta_{wt}\right)$, in quarter terms. In this case the evidence is more difficult to interpret, with a big spike in the second half of the 1980s which is probably due to sampling uncertainty. However, we still see that, during the 1970s, as inflation went up, wage rigidity went down, and as inflation was tamed in the early 1980s, wages again became more rigid.

Figures 6.3.9 and 6.3.10 draw the evolution of wage indexation. Here, in comparison, the clarity of the result is embarrassing: wage indexation is nearly the perfect mirror of inflation. As we did for prices, we interpret this finding as the natural consequence of workers switching to more often wage reoptimizations that make indexation less of an interesting rule.[5] Less wage indexation is what the model needs to capture the higher volatility of inflation in the data.

For completeness, we finish our graphical display with figures 6.3.11 to 6.3.16, where we plot the evolution of the different parameters controlling nominal rigidities against others. Because of space constraints, we refrain from further discussion of the plots. However, the reader can appreciate that the similarity in the evolution of the parameters over time solidifies our confidence that we are uncovering a systematic pattern of relationships between nominal rigidities and inflation.

We consider our findings to be strong proof of the changing nature of the nominal rigidities in the economy and of a strong indication of model misspecification along the dimension of price and wage adjustment. Calvo's price adjustment cannot capture the evolution of the

[5] During the early 1970s, there was a raise in the prevalence of cost-of-living allowance (COLA) escalators in collective bargaining agreements (Hendricks and Kahn, 1985). This observation could be used to undermine our result. However, even at their peak, COLAs only covered 6 millions workers, a small percentage of the labor force. Moreover, it is difficult to map COLAs from the 1970s into our model since they had many contingent rules that make them quite different from the naïve indexation rules that we use. In fact, it could be even possible to think about a state-contingent COLA as an implicit form of reoptimization.

fundamentals that determine the pricing decisions of firms and households. Our results underscore that this problem is relevant empirically. Also, they suggest that the evidence in Klenow and Kryvtsov (2005) that the intensive margin of price changes accounts for 95 percent of the monthly variance of inflation may be a product of the sample period (1988-2003), where the low level of inflation limits identification because it eliminates the source of variation of the data. Indeed, in our figures 6.3.5 and 6.3.6, if we look at the period 1988-2000, we observe much less variation in the pricing parameters.

More generally, we read our results as favoring models of state-dependent pricing (Caballero and Engel, 1993, Caplin and Leahy, 1991 and 1997) over time-dependent pricing as in the Calvo-type model we have presented or, indirectly, Taylor's staggered contract economies (Taylor, 1980). The extra analytical difficulty implied by state-dependent models (Dotsey, King, and Wolman, 1999) may be a price we are forced to pay.

Another strand of the literature that may find our results interesting is the one that deals with sticky information (Mankiw and Reiss, 2002 or Sims, 2002). Higher inflation increases the incentives to gather information and hence, it is likely to imply more frequent price and wage adjustments.

Finally, our findings also have relevant implications for optimal policy design. First, if we interpret the evolution of parameters like θ_{pt} as exogenously given, it may be something that the monetary authority may condition its behavior on (we do not enter into a discussion of how it would estimate them in real time, we only raise this as a theoretical possibility). Second, if we interpret our results as showing that the measured amount of price rigidities are endogenous to monetary policy, optimal design becomes tougher than in the standard Ramsey exercises.

7. Conclusion

How structural are the structural parameters of DSGE models? Less so than we often claim. Our analysis indicates that there are large variations in the estimated values of several of the key parameters of a benchmark medium-scale macroeconomic model during our sample period.

We document changes in the response of the monetary authority to inflation and in the inflation target that confirm previous findings by other researchers. In particular, we report a move by the Fed toward a much more aggressive stand against raising prices. Our results are remarkable because they are derived in a context where agents understand that policy evolves over time and respond to that evolution.

We also uncover that the parameters controlling nominal rigidities drift in a substantial way, and more important, are strongly correlated with inflation. These findings cast doubts on the usefulness of models based on Calvo pricing and invite deeper investigations of state-dependent pricing models.

We do not want our work to be interpreted as a sweeping criticism of the estimation of DSGE models, because it is not. The literature has made impressive progress over the last years and has contributed much to improving our understanding of aggregate fluctuations and the effects of economic policy. We ourselves have been engaged in this research agenda and plan to continue doing so. We hope, instead, that our paper will be read as an invitation to further estimation of DSGE models with parameter drifting. This avenue is promising, both as a mechanism to incorporate richer dynamics and as a diagnostic tool for detecting gross misspecifications.

Finally, our skepticism about the structural nature of most "structural" parameters is not a call to perform reduced-form exercises. With Tom Sargent and Mark Watson (Fernández-Villaverde *et al.*, 2007), we have singled out some of the problems of estimating reduced form models. But there are many other papers emphasizing the weaknesses of reduced form inference, too many indeed to even bother with a list. Our position is that every empirical procedure has strengths and limitations. As Hurwicz (1962) warned us many years ago, just because we name something "structural," we should not believe we have taken the theoretical high-ground.

8. Appendix

This appendix offers further details about some of the technical aspects of the paper. First, we discuss some general computational aspects and elaborate on the solution of the model. Second, we describe the particle filter that we use to evaluate the likelihood function of the model. Third, we comment on the estimation procedure. Fourth, we close with the details of the construction of the data.

8.1. Computation of the Model

The most important feature of the algorithm to be described below to solve and estimate the model is that it can be implemented on a good desktop computer. We coded all programs needed for the perturbation of the model and the particle filter in Fortran 95 and compiled them in Intel Visual Fortran 9.1 to run on Windows-based machines (except some `Mathematica` programs to generate analytic derivatives). We use a Xeon Processor 5160 EMT64 at 3.00 GHz with 16 GB of RAM.

As described in the main text, the solution of the model is challenging because we have 19 state variables plus the drifting parameters that we allow in each empirical exercise. Moreover, we need to recompute the solution of the model for each new set of parameter values in the estimation. The only non-linear procedure that can accomplish this computation in a reasonable amount of time is perturbation (see Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006). We implement our solution by perturbing the equilibrium conditions of the rescaled version of the model (i.e., the one where we have already eliminated the two unit roots) around the deterministic steady state. This means that the solution is locally accurate regardless of the level of the technology of the economy. Also, note that the steady state will depend on the level of inflation targeted by the monetary authority.

We use `Mathematica` to compute the analytical derivatives and to generate Fortran 95 code with the corresponding analytical expression. Then, we load that output into a Fortran 95 code that evaluates the solution of the model for each parameter value as implied by the Metropolis-Hastings algorithm to be described below. The solution will have the form:

$$\left(\overline{S}'_{t+1}, J'_t\right)' = \Gamma_{s1}\left(\overline{S}'_t, \varepsilon'_t\right)' + \frac{1}{2}\left(\overline{S}'_t, \varepsilon'_t\right)\Gamma_{s2}\left(\overline{S}'_t, \varepsilon'_t\right)' + \Gamma_{s3}$$

where, recalling our notation, $\overline{S}_t$ are the states of the model, ε_t are the shocks, J_t is a vector of variables of interest in the model that are not states, and the Γ_{si}'s are matrices of the right size. With the solution of the model, and by selecting the appropriate rows, we can build the state space form:

$$\begin{aligned} \overline{S}_{t+1} &= \Psi_{s1}\left(\overline{S}_t', \varepsilon_t'\right)' + \frac{1}{2}\left(\overline{S}_t', \varepsilon_t'\right)\Psi_{s2}\left(\overline{S}_t', \varepsilon_t'\right)' + \Psi_{s3} \quad (13) \\ \mathcal{Y}^T &= \Psi_{o1}\left(S_t', \varepsilon_t'\right)' + \frac{1}{2}\left(S_t', \varepsilon_t'\right)\Psi_{o2}\left(S_t', \varepsilon_t'\right)' + \Psi_{o3} \quad (14) \end{aligned}$$

where $S_t = \left(\overline{S}_t', \overline{S}_{t-1}', \varepsilon_{t-1}'\right)$ and

$$\mathcal{Y}^T = \left(\triangle \log \mu_t^{-1}, \triangle \log y_t, \triangle \log x_t, \triangle \log l_t, \log \Pi_t, \log R_t\right)'.$$

8.2. Description of the Particle Filter

We provide now a short description of the particle filter. We will deliberately focus on the intuition of the procedure and we will gloss over many technical issues that are relevant for a successful application of the filter. We direct the interested reader to Fernández-Villaverde and Rubio-Ramírez (2007), where we discuss most of those issues in much detail, and to the excellent collection of articles in Doucet, de Freitas, and Gordon (2001). Also, note that we are presenting here only a basic sequential Monte Carlo filter. Researchers have defended many different routes to improve the efficiency of the simulation. Pitts and Shephard's (1999) auxiliary particle filter, which uses the observed data in the current period, is perhaps the most celebrated of those alternatives.

As we described in the main text, given the Markov structure of our state space representation, we can factorize the likelihood function as:

$$\mathcal{L}\left(\mathbb{Y}^T; \Psi\right) = \prod_{t=1}^{T} \mathcal{L}\left(\mathbb{Y}_t | \mathbb{Y}^{t-1}; \Psi\right)$$

and obtain the factorization:

$$\mathcal{L}\left(\mathbb{Y}^T; \Psi\right) = \int \mathcal{L}\left(\mathbb{Y}_1 | S_0; \Psi\right) dS_0 \prod_{t=2}^{T} \int \mathcal{L}\left(\mathbb{Y}_t | S_t; \Psi\right) p\left(S_t | \mathbb{Y}^{t-1}; \Psi\right) dS_t \quad (15)$$

Consequently, if we had the sequence $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$ and $p(S_0;\Psi)$, we could evaluate the likelihood of the model. Santos and Peralta-Alva (2005) show conditions under which we can draw the numerical solution of the model to approximate $p(S_0;\Psi)$. The two difficulties of evaluation (15) are then to characterize the sequence of conditional distributions $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$ and to compute the different integrals in the expression.

The particle filter begins from the observation that, if somehow we can get N draws of the form $\left\{\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}\right\}_{t=1}^{T}$ from the sequence $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$, we can appeal to a law of large numbers and substitute the integrals with a mean of the conditional likelihoods evaluated in the empirical draws:

$$\mathcal{L}\left(\mathbb{Y}^{T};\Psi\right) \simeq \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_1|s_{0|0}^{i};\Psi\right)\prod_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)$$

where our notation for the draws indicates in the subindex the conditioning set (i.e., $t|t-1$ means draw at moment t conditional on information until $t-1$) and the superindex denotes the index of the draw. The intuition of the procedure is that we substitute the exact but unknown sequence $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$ by its empirical counterpart.

How do we draw from $\{p(S_t|\mathbb{Y}^{t-1};\Psi)\}_{t=1}^{T}$? The second key idea of the particle filter is that we can extend importance sampling (Geweke, 1989) to a sequential environment. The following proposition, due in its original form to Rubin (1988), formalizes the idea:

Proposition 1. *Let* $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ *be a draw from* $p(S_t|\mathbb{Y}^{t-1};\Psi)$. *Let the sequence* $\{\widetilde{s}_t^{i}\}_{i=1}^{N}$ *be a draw with replacement from* $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ *where the resampling probability is given by*

$$q_t^i = \frac{\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)}{\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)},$$

Then $\{\widetilde{s}_t^{i}\}_{i=1}^{N}$ *is a draw from* $p(S_t|\mathbb{Y}^{t};\Psi)$.

The proposition 1 shows how to recursively use a draw $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ from $p(S_t|\mathbb{Y}^{t-1};\Psi)$ to get a draw $\left\{s_{t|t}^{i}\right\}_{i=1}^{N}$ from $p(S_t|\mathbb{Y}^{t};\Psi)$. This result is crucial. It allows us to incorporate the information in $\mathbb{Y}_t$ to change our current estimate of S_t. This is why this step is known in filtering theory as update (the discerning reader has probably already realized that this update is nothing more than an application of the Bayes' theorem).

The resampling step is key for the success of the filter. A naïve extension of Monte Carlo techniques will just draw a whole sequence of $\left\{\left\{s^i_{t|t-1}\right\}^N_{i=1}\right\}^T_{t=1}$ without stopping period by period to resample according to proposition 1. It is easy to show how this naïve scheme diverges. The reason is that all the sequences become arbitrarily far away from the true sequence of states, which is a zero measure set and the sequence that is closer to the true states dominates all the remaining ones in weight. Simple experiments show that this degeneracy appears even after very few steps.

Given $\left\{s^i_{t|t}\right\}^N_{i=1}$, we draw N exogenous shocks, something quite simple, since exogenous shocks in our model:

$$\varepsilon^i_{t+1} = \left(\varepsilon^i_{\mu,t+1}, \varepsilon^i_{d,t+1}, \varepsilon^i_{\varphi,t+1}, \varepsilon^i_{A,t+1}, \varepsilon^i_{m,t+1}\right)'$$

are normally distributed. Then, we apply the law of motion for states that relates the $s^i_{t|t}$ and the shocks ε^i_{t+1} to generate $\left\{s^i_{t+1|t}\right\}^N_{i=1}$. This step, known as forecast, put us back at the beginning of proposition 1, but with the difference that we have moved forward one period in our conditioning.

The following pseudocode summarizes the description of the algorithm:

Step 0, Initialization: `Set` $t \rightsquigarrow 1$. `Sample` N `values` $\left\{s^i_{0|0}\right\}^N_{i=1}$ from $p\left(S_0;\Psi\right)$.

Step 1, Prediction: `Sample` N `values` $\left\{s^i_{t|t-1}\right\}^N_{i=1}$ `using` $\left\{s^i_{t-1|t-1}\right\}^N_{i=1}$, `the law of motion for states and the distribution of shocks` ε_t.

Step 2, Filtering: `Assign to each draw` $\left(s^i_{t|t-1}\right)$ `the weight` q^i_t `in proposition 1.`

Step 3, Sampling: `Sample` N `times with replacement from` $\left\{s^i_{t|t-1}\right\}^N_{i=1}$ `using the probabilities` $\{q^i_t\}^N_{i=1}$. `Call each draw` $\left(s^i_{t|t}\right)$. `If` $t < T$ `set` $t \rightsquigarrow t+1$ `and go to step 1. Otherwise stop.`

Then, with the output of the algorithm, we just substitute into our formula

$$\mathcal{L}\left(\mathbb{Y}^T;\Psi\right) \simeq \frac{1}{N}\sum_{i=1}^N \mathcal{L}\left(\mathbb{Y}_1|s^i_{0|0};\Psi\right)\prod_{t=2}^T \frac{1}{N}\sum_{i=1}^N \mathcal{L}\left(\mathbb{Y}_t|s^i_{t|t-1};\Psi\right) \tag{16}$$

and get an estimate of the likelihood of the model. Del Moral and Jacod (2002) and Künsch

(2005) show weak conditions under which the right-hand side of the previous equation is a consistent estimator of $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$ and a central limit theorem applies.

8.3. Estimation Procedure

We mention in the main part of the text that the posterior of the model

$$p\left(\Psi|\,\mathbb{Y}^T\right) \propto \mathcal{L}\left(\mathbb{Y}^T;\Psi\right)p\left(\Psi\right)$$

is difficult, if not impossible, to characterize. However, we can draw from it and build its empirical counterpart using a Metropolis-Hastings algorithm. The algorithm is as follows:

Step 0, Initialization: `Set` $i \rightsquigarrow 0$ `and an initial` Ψ_i`. Solve the model for` Ψ_i `and build the state space representation. Evaluate prior` $p\left(\Psi_i\right)$ `and approximate` $\mathcal{L}\left(\mathbb{Y}^T;\Psi\right)$ `with (16). Set` $i \rightsquigarrow i+1$.

Step 1, Proposal draw: `Get a draw` Ψ_i^* `from a proposal density` $q\left(\gamma_{i-1},\gamma_i^*\right)$`.`

Step 2, Solving the Model: `Solve the model for` Ψ_i^* `and build the new state space representation.`

Step 3, Evaluating the proposal: `Evaluate` $p\left(\Psi_i^*\right)$ `and` $\mathcal{L}\left(\mathbb{Y}^T;\Psi_i^*\right)$ `with (16).`

Step 4, Accept/Reject: `Draw` $\chi_i \sim U\left(0,1\right)$`. If` $\chi_i \leq \frac{\mathcal{L}\left(\mathbb{Y}^T;\Psi_i^*\right)p\left(\Psi_i^*\right)q\left(\Psi_{i-1},\Psi_i^*\right)}{\mathcal{L}\left(\mathbb{Y}^T;\Psi_{i-1}\right)p\left(\Psi_{i-1}\right)q\left(\Psi_i^*,\Psi_{i-1}\right)}$ `set` $\Psi_i=\Psi_i^*$`, otherwise` $\Psi_i=\Psi_{i-1}$`.`

Step 5, Iteration: `If` $i<M$, `set` $i \rightsquigarrow i+1$ `and go to step 1. Otherwise stop.`

This algorithm requires us to specify a proposal density $q\left(\cdot,\cdot\right)$. We follow the standard practice and choose a random walk proposal, $\Psi_i^*=\Psi_{i-1}+\kappa_i$, $\kappa_i \sim \mathcal{N}\left(0,\Sigma_\kappa\right)$, where Σ_κ is a scaling matrix. This matrix is selected to get the appropriate acceptance ratio of proposals (Roberts, Gelman and Gilks, 1997).

To reduce the "chatter" of the problem, we will keep the innovations in the particle filter (i.e., the draws from the exogenous shock distributions and the resampling probabilities) constant across different passes of the Metropolis-Hastings algorithm. As pointed out by McFadden (1989) and Pakes and Pollard (1989), this is required to achieve stochastic equicontinuity, and even if this condition is not strictly necessary in a Bayesian framework, it does reduce the numerical variance of the procedure.

8.4. Construction of Data

As we mention in the text, we compute both real output and real gross investment in consumption units to make the observed series compatible with the model. We define the relative price of investment as the ratio of the investment deflator and the deflator for consumption. The consumption deflator is constructed from the deflators of nondurable goods and services reported in the NIPA. Since the NIPA investment deflators are poorly measured, we use the investment deflator constructed by Fisher (2006). For the real output per capita series, we first define nominal output as nominal consumption plus nominal gross investment. We define nominal consumption as the sum of personal consumption expenditures on nondurable goods and services, national defense consumption expenditures, federal nondefense consumption expenditures, and state and local government consumption expenditures. We define nominal gross investment as the sum of personal consumption expenditures on durable goods, national defense gross investment, federal government nondefense gross investment, state and local government gross investment, private nonresidential fixed investment, and private residential fixed investment. Per capita nominal output is defined as the ratio between our nominal output series and the civilian noninstitutional population between 16 and 65. Since we need to measure real output per capita in consumption units, we deflate the series by the consumption deflator. For the real gross investment per capita series, we divide our above mentioned nominal gross investment series by the civilian noninstitutional population between 16 and 65 and the consumption deflator. Finally, the hours worked per capita series is constructed with the index of total number of hours worked in the business sector and the civilian noninstitutional population between 16 and 65. Since our model implies that hours worked per capita are between 0 and 1, we normalize the observed series of hours worked per capita such that it is, on average, 0.33.

References

[1] An, S. and F. Schorfheide (2006). "Bayesian Analysis of DSGE Models." *Econometric Reviews,* forthcoming.

[2] Aruoba, S.B., J. Fernández-Villaverde, and J. Rubio-Ramírez (2006). "Comparing Solution Methods for Dynamic Equilibrium Economies.". *Journal of Economic Dynamics and Control* 30, 2477-2508.

[3] Arulampalam, A.S., S. Maskell, N. Gordon, and T. Clapp (2002). "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking". *IEEE Transactions on Signal Processing* 50, 174-188.

[4] Benati, L. (2006). "Investigating Inflation Persistence Across Monetary Regimes." *Mimeo*, European Central Bank.

[5] Bernanke, B.S. and I. Mihov (1998). "Measuring Monetary Policy." *Quarterly Journal of Economics* 113, 869-902.

[6] Bils, M. and P. Klenow (2004). "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112, 947-985.

[7] Boivin, J. (2006). "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data." *Journal of Money, Credit and Banking*, forthcoming.

[8] Boivin, J. and M. Giannoni (2006). "DSGE Models in a Data-Rich Environment." *NBER Working Paper* 12272.

[9] Browning, M., L.P. Hansen, and J.J. Heckman (1999). "Micro Data and General Equilibrium Models." in J.B. Taylor & M. Woodford (eds.), *Handbook of Macroeconomics*, volume 1, chapter 8, 543-633. Elsevier.

[10] Caballero, R. and E. Engel (1999). "Heterogeneity and Output Fluctuations in a Dynamic Menu-Costs Economy." *Review of Economic Studies* 60, 95-119.

[11] Canova, F. and L. Sala (2006). "Back to Square One: Identification Issues in DSGE Models." *Mimeo*, CREI.

[12] Caplin, A. and J. Leahy (1991). "State-Dependent Pricing and the Dynamics of Money and Output." *Quarterly Journal of Economics*, 106, 683-708.

[13] Caplin, A. and J. Leahy (1997). "Aggregation and Optimization with State-Dependent Pricing," *Econometrica* 65, 601-626.

[14] Chung, H., T. Davig, and E. Leeper (2006). "Monetary and Fiscal Policy Switching." *Mimeo,* Indiana University.

[15] Clarida, R., J. Galí, and M. Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics* 115, 147-180.

[16] Cogley, T. and T.J. Sargent (2001). "Evolving Post-World War II U.S. Inflation Dynamics." Macroeconomics Annual, volume 16.

[17] Cogley, T. and T.J. Sargent (2005). "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S." *Review of Economic Dynamics* 8, 262-302.

[18] Cogley, T. and A. Sbordone (2006). "Trend Inflation and Inflation Peristence in the New Keynesian Phillips Curve." *Mimeo*, University of California at Davis.

[19] Cooley, T.F. and E.C. Prescott (1976). "Estimation in the Presence of Stochastic Parameter Variation." *Econometrica* 44, 167-184.

[20] Cooley, T.F. and E.C. Prescott (1995), "Economic Growth and Business Cycles" in T. F. Cooley. (ed), *Frontiers of Business Cycle Research*. Princeton University Press.

[21] Davig, T. and E. Leeper (2006a). "Generalizing the Taylor Principle." *American Economic Review*, forthcoming.

[22] Davig, T. and E. Leeper (2006b). "Endogenous Monetary Policy Regime Change." *NBER International Seminar on Macroeconomics*, forthcoming.

[23] Del Moral P. and J. Jacod (2002), "The Monte-Carlo Method for Filtering with Discrete Time Observations. Central Limit Theorems." in T. J. Lyons and T. S. Salisbury (eds), *Numerical Methods and Stochastics*. The Fields Institute Communication*s,* American Mathematical Society.

[24] Dotsey, M., R.G. King, and A. Wolman (1999). "State Dependent Pricing and the General Equilibrium Dynamics of Money and Output." *Quarterly Journal of Economics* 655-690.

[25] Doucet. A., N. de Freitas, and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. Springer Verlag.

[26] Farmer, R., D. Waggoner, and T. Zha (2006). "Indeterminacy in a Forward Looking Regime Switching Model." *Mimeo*, UCLA.

[27] Fernández-Villaverde, J. and J. Rubio-Ramírez (2004). "Comparing Dynamic Equilibrium Models to Data: a Bayesian Approach." *Journal of Econometrics, 123,* 153-187.

[28] Fernández-Villaverde, J. and J. Rubio-Ramírez (2005). "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood." *Journal of Applied Econometrics,* 20, 891-910.

[29] Fernández-Villaverde, J. and J. Rubio-Ramírez (2007). "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies,* forthcoming.

[30] Fernández-Villaverde, J., J. Rubio-Ramírez, and M.S. Santos (2006). "Convergence Properties of the Likelihood of Computed Dynamic Models." *Econometrica* 74, 93-119.

[31] Fernández-Villaverde, J., J. Rubio-Ramírez, T.J. Sargent, and M. Watson (2006). "A,B,C's (and D)'s for Understanding VARs." *American Economic Review,* forthcoming.

[32] Fisher, J. (2006). "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks." *Journal of Political Economy* 114, 413-451.

[33] Geweke, J. (1989). "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 24, 1037-1399.

[34] Godsill, S.J., A. Doucet, and M. West (2004). "Monte Carlo Smoothing for Nonlinear Time Series." *Journal of the American Statistical Association* 99, 156-168.

[35] Greenwood, J, Z. Herkowitz, and P. Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review* 87, 342-362.

[36] Greenwood, J, Z. Herkowitz, and P. Krusell (2000). "The Role of Investment-Specific Technological Change in the Business Cycle." *European Economic Review* 44, 91-115.

[37] Justiniano A. and G.E. Primiceri (2005). "The Time Varying Volatility of Macroeconomic Fluctuations." *Mimeo*, Northwestern.

[38] Hendricks, W.E. and L. Khan (1985). *Wage Indexation in the United States: Cola or Uncola?* Cambridge: Ballinger Publishing Company.

[39] King, T.B. (2006). "Dynamic Equilibrium Models with Time-Varying Structural Parameters." *Mimeo*, Washington University.

[40] Künsch, H.R. (2005). "Recursive Monte Carlo Filters: Algorithms and Theoretical Analysis." *Annals of Statistics* 33, 1983-2021.

[41] Lubick, T. and F. Schorfheide (2004). "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review* 94, 190-217.

[42] Mankiw, G.N. and R. Reiss (2002). "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117, 1295-1328.

[43] McFadden, D.L. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57, 995-1026.

[44] Nakamura, E. and J. Steinsson (2006). "Five Facts About Prices: A Reevaluation of Menu Cost Models." *Mimeo*, Harvard University.

[45] Oliner, S. G. Rudebusch, and D. Sichel (1996) "The Lucas Critique Revisited: Assessing the Stability of Empirical Euler Equations for Investment." *Journal of Econometrics* 70, 291-316.

[46] Orphanides, A. (2002). "Monetary Policy Rules and the Great Inflation." American Economic Review 92, 115-120

[47] Pakes, A. and D. Pollard (1989). "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57, 1027-1057.

[48] Pitt, M.K. and N. Shephard (1999). "Filtering via Simulation: Auxiliary Particle Filters." *Journal of the American Statistical Association* 94, 590-599.

[49] Ploberger, W. and P.C.B. Phillips (2003). "Empirical Limits for Time Series Econometric Models." *Econometrica* 71, 627-673.

[50] Primiceri, G.E. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy". *Review of Economic Studies* 72, 821-852.

[51] Primiceri, G.E. (2006). "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy." *Quarterly Review of Economics* 121, 867-901

[52] Romer, C. and D. Romer (2002). "The Evolution of Economic Understanding and Postwar Stabilization Policy" in *Rethinking Stabilization Policy*, Federal Reserve Bank of Kansas City.

[53] Roberts, G., A. Gelman, and W. Gilks (1997). "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorthims." *Annals of Applied Probability* 7, 110-120.

[54] Rubin, D.B. (1988). "Using the SIR Algorithm to Simulate Posterior Distributions". in J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (eds), *Bayesian Statistics 3*, 395-402, Oxford University Press.

[55] Santos, M.S. and A. Peralta-Alva (2005). "Accuracy of Simulations for Stochastic Dynamic Models". *Econometrica* 73, 1939-1976.

[56] Sargent, T.J. (1999). *The Conquest of American Inflation*. Princeton University Press.

[57] Sims, C.A. (2001). "Discussion of Evolving Post-World War II U.S. Inflation Dynamics by T. Cogley and T.J. Sargent." NBER Macroeconomics Annual 16.

[58] Sims, C.A. (2002). "Implications of Rational Inattention." *Mimeo*, Princeton University..

[59] Sims, C.A. and T. Zha (2006). "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* vol 96, 54-81.

[60] Taylor, J. (1980). "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88, 1-23.

[61] Taylor, J. (1998). "Monetary Policy Guidelines for Unemployment and Inflation Stability." In *Inflation, Unemployment, and Monetary Policy*, R.M. Solown and J.B. Taylor (eds). MIT Press.

[62] Uhlig, H. (1997). "Bayesian Vector Autoregressions with Stochastic Volatility." *Econometrica* 65, 59-73.

[63] Woodford, M. (2006). "Interpreting Inflation Persistence: Comments on the Conference on Quantitative Evidence on Price Discrimination." *Journal of Money, Credit and Banking*, forthcoming.

[64] Young, A.T. (2004). "Labor's Share Fluctuations, Biased Technical Change, and the Business Cycle." *Review of Economic Dynamics* 7, 916-931.

Figure 2.3.1: Estimate of β versus $(1+g)^{-2}$

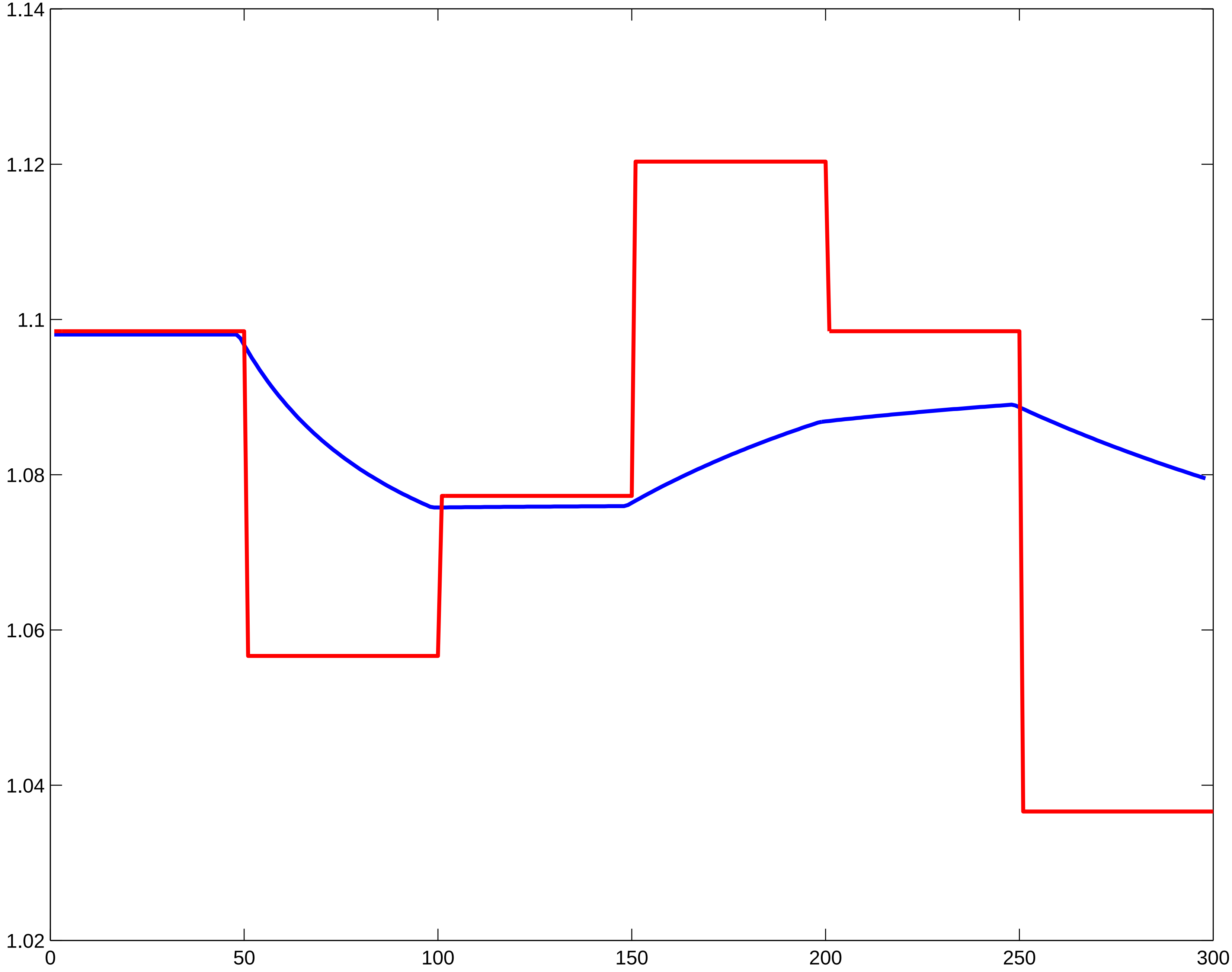


Figure 6.2.1: Evolution of Response to Inflation

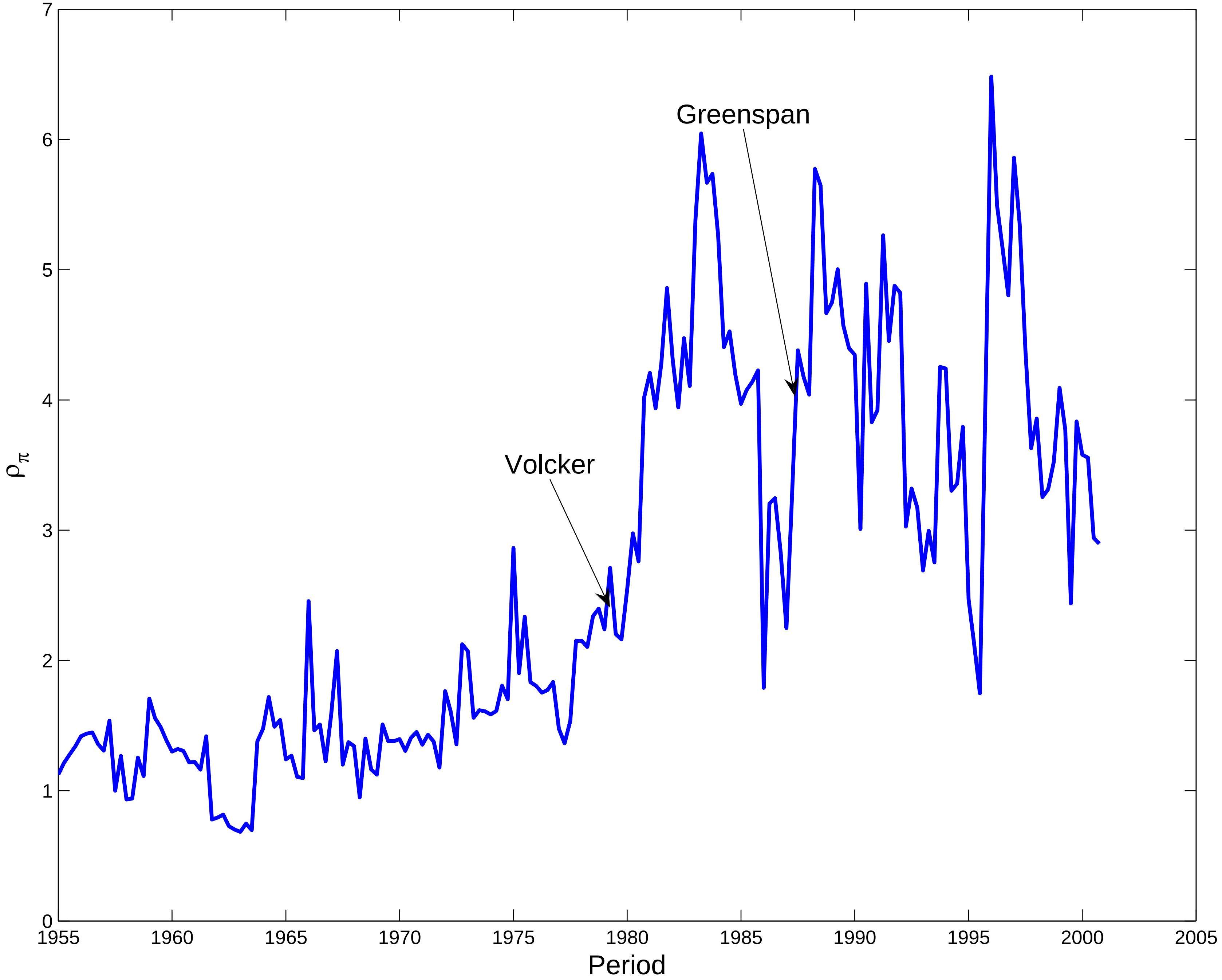


Figure 6.2.2: HP-Trend Evolution of Response to Inflation

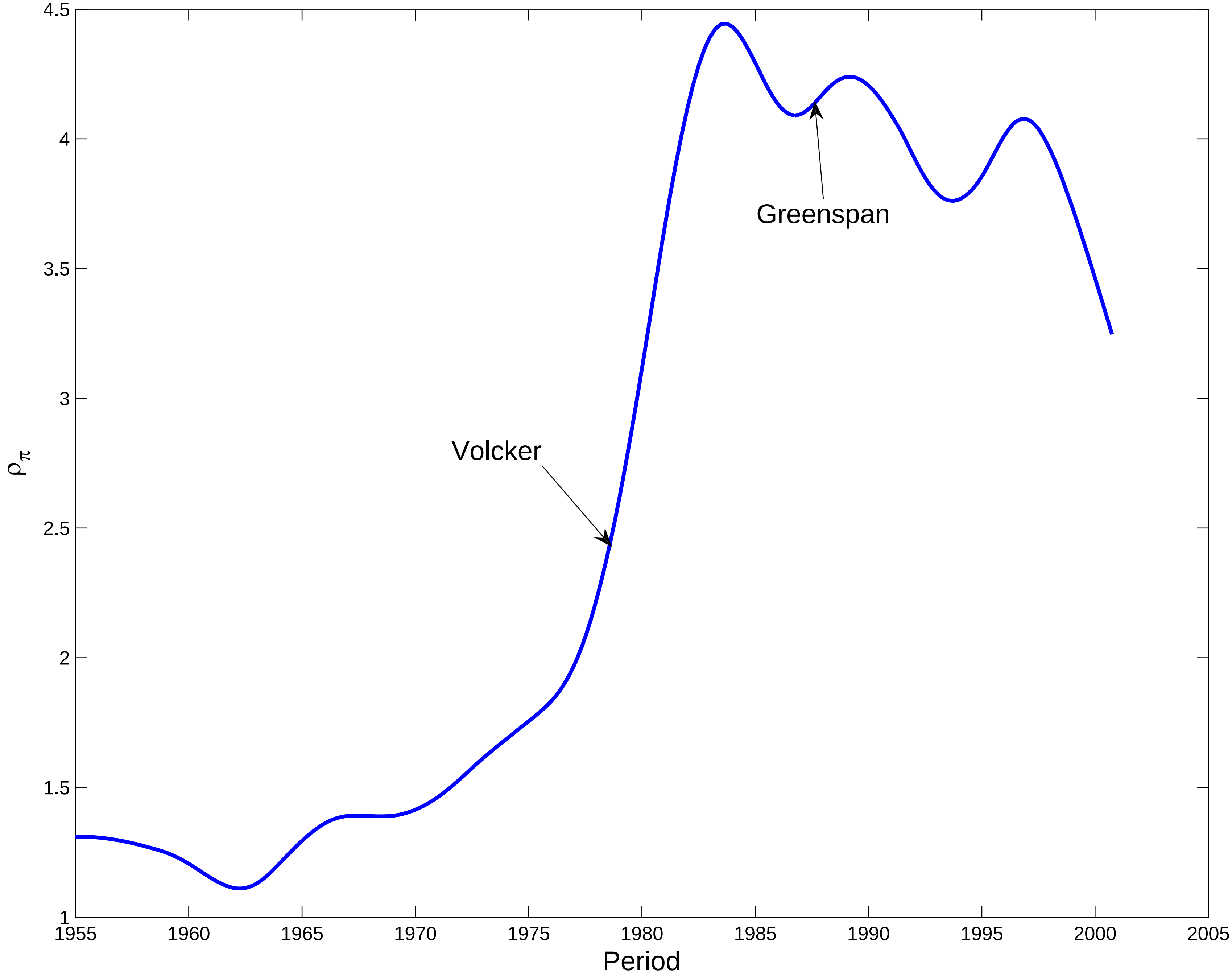
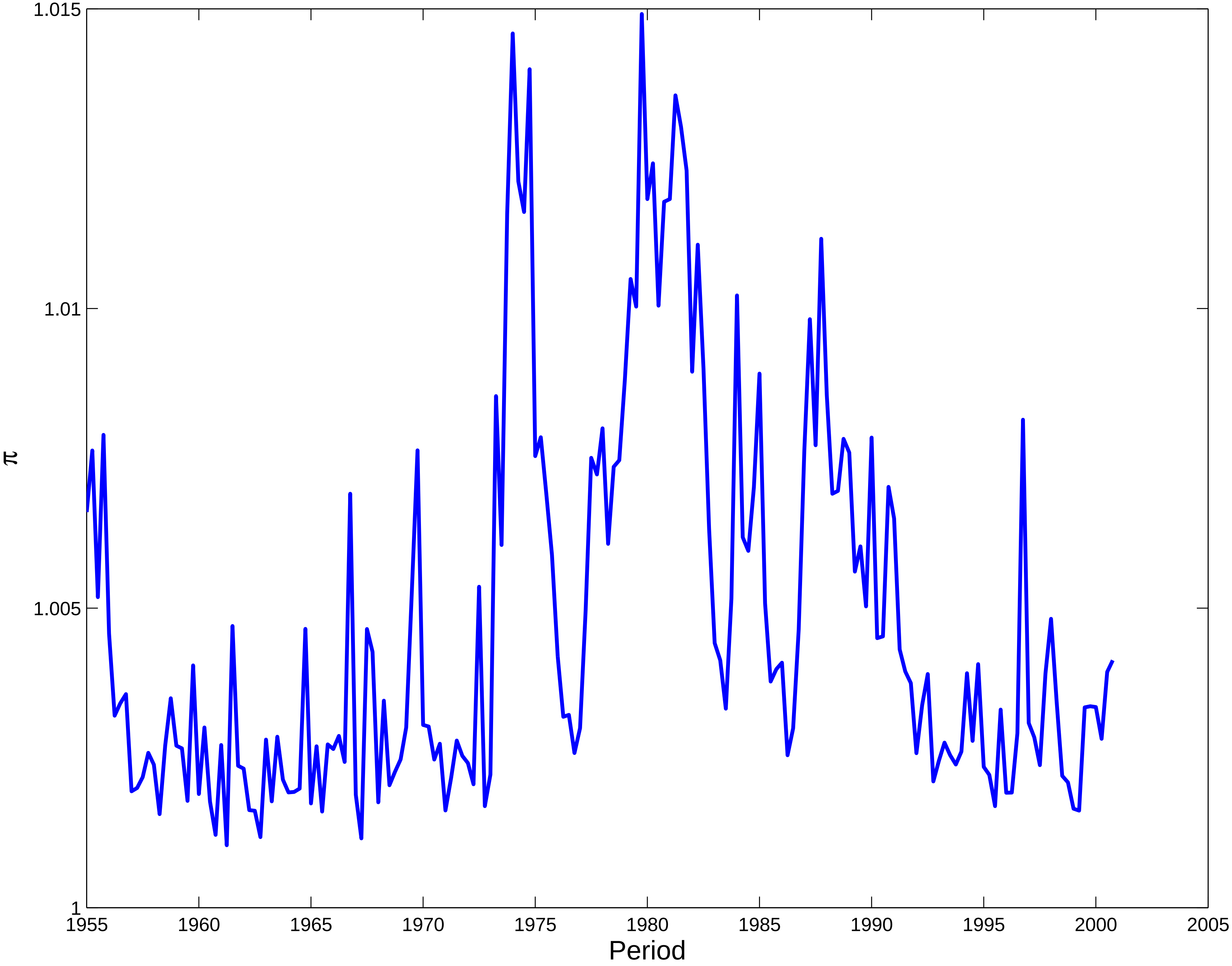


Figure 6.2.3: Evolution of Inflation Target



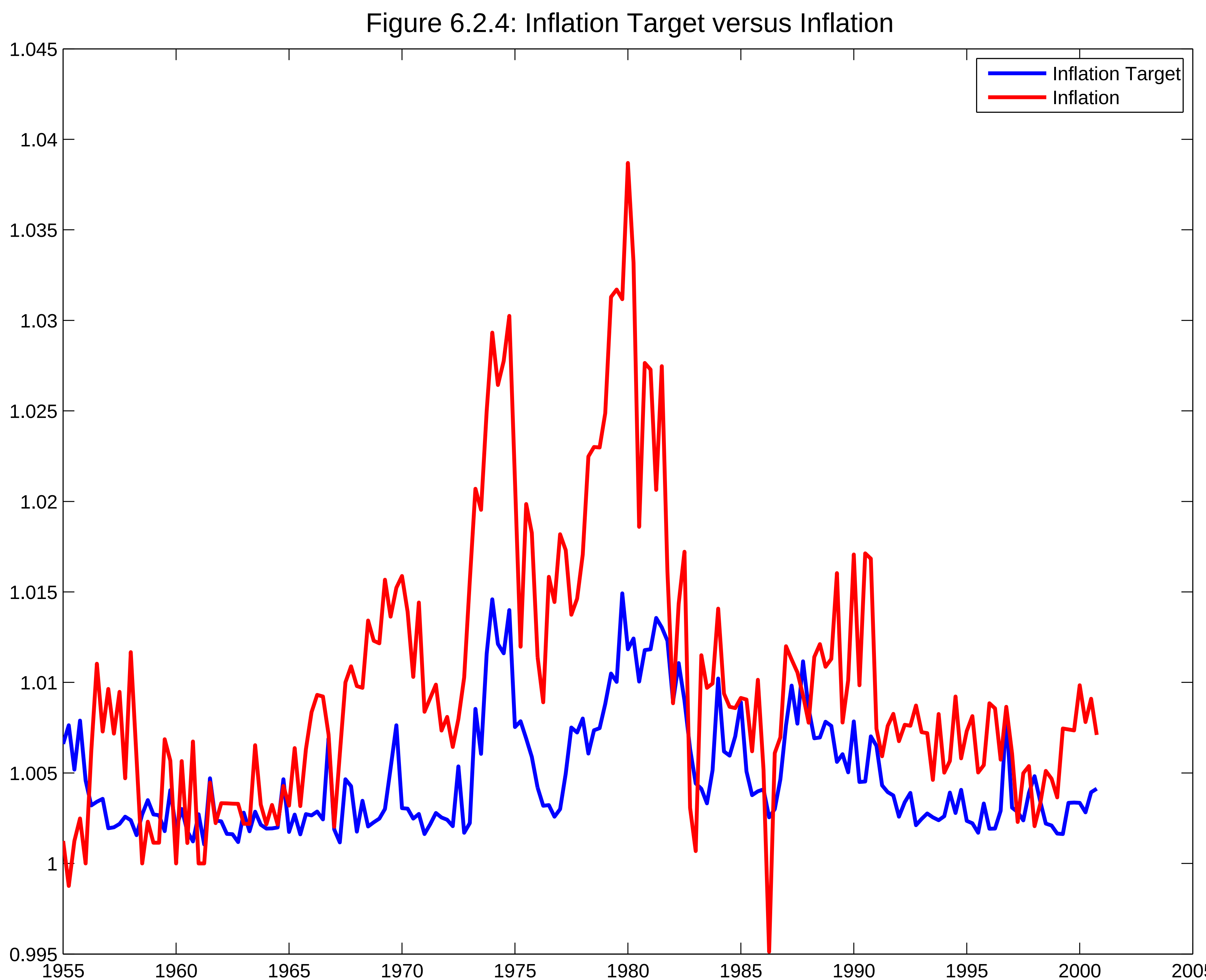


Figure 6.2.4: Inflation Target versus Inflation
Inflation Target
Inflation
1.045
1.04
1.035
1.03
1.025
1.02
1.015
1.01
1.005
1
0.995
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005

Figure 6.2.5: Inflation Target versus Inflation

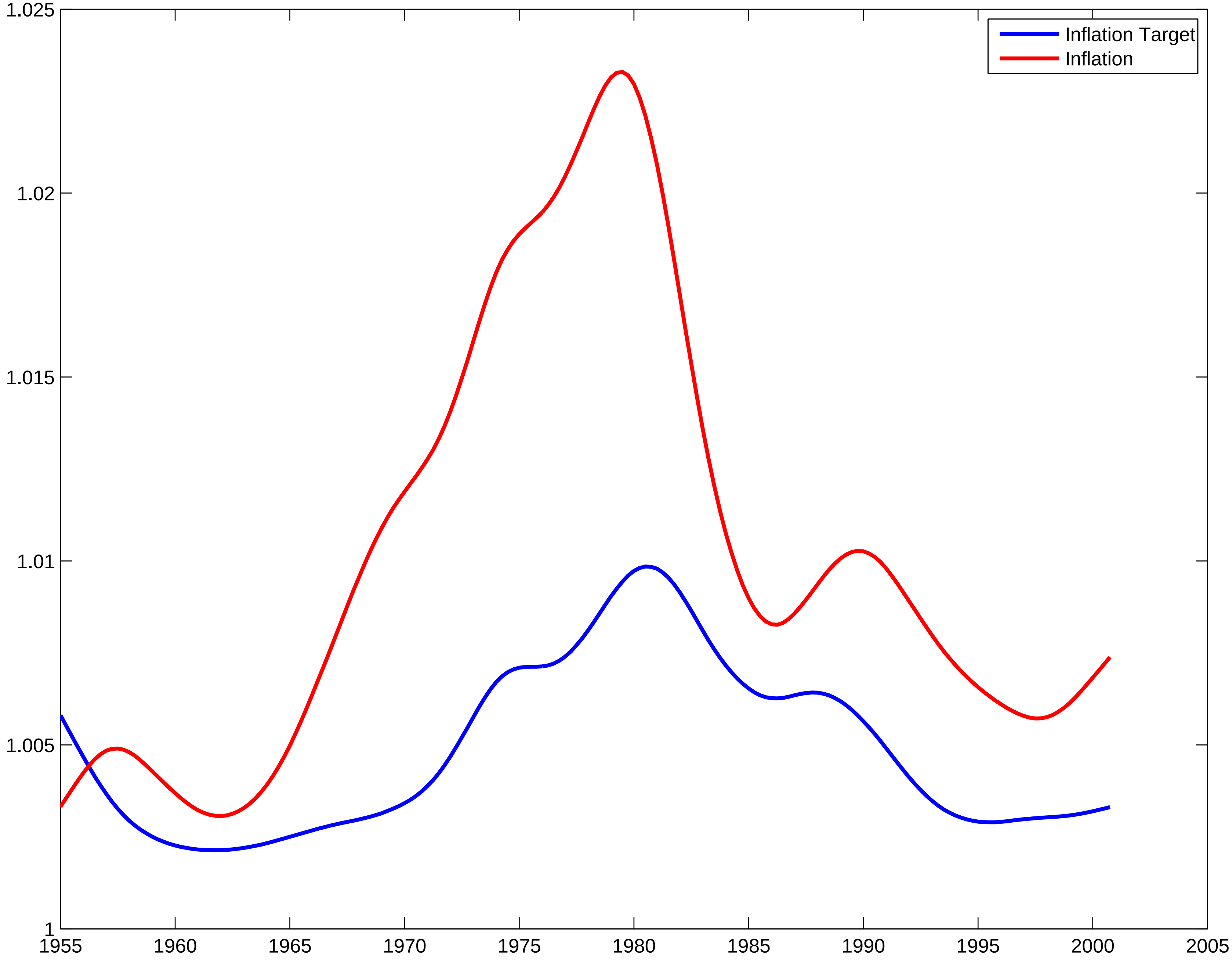


Figure 6.3.1: Average Price Duration

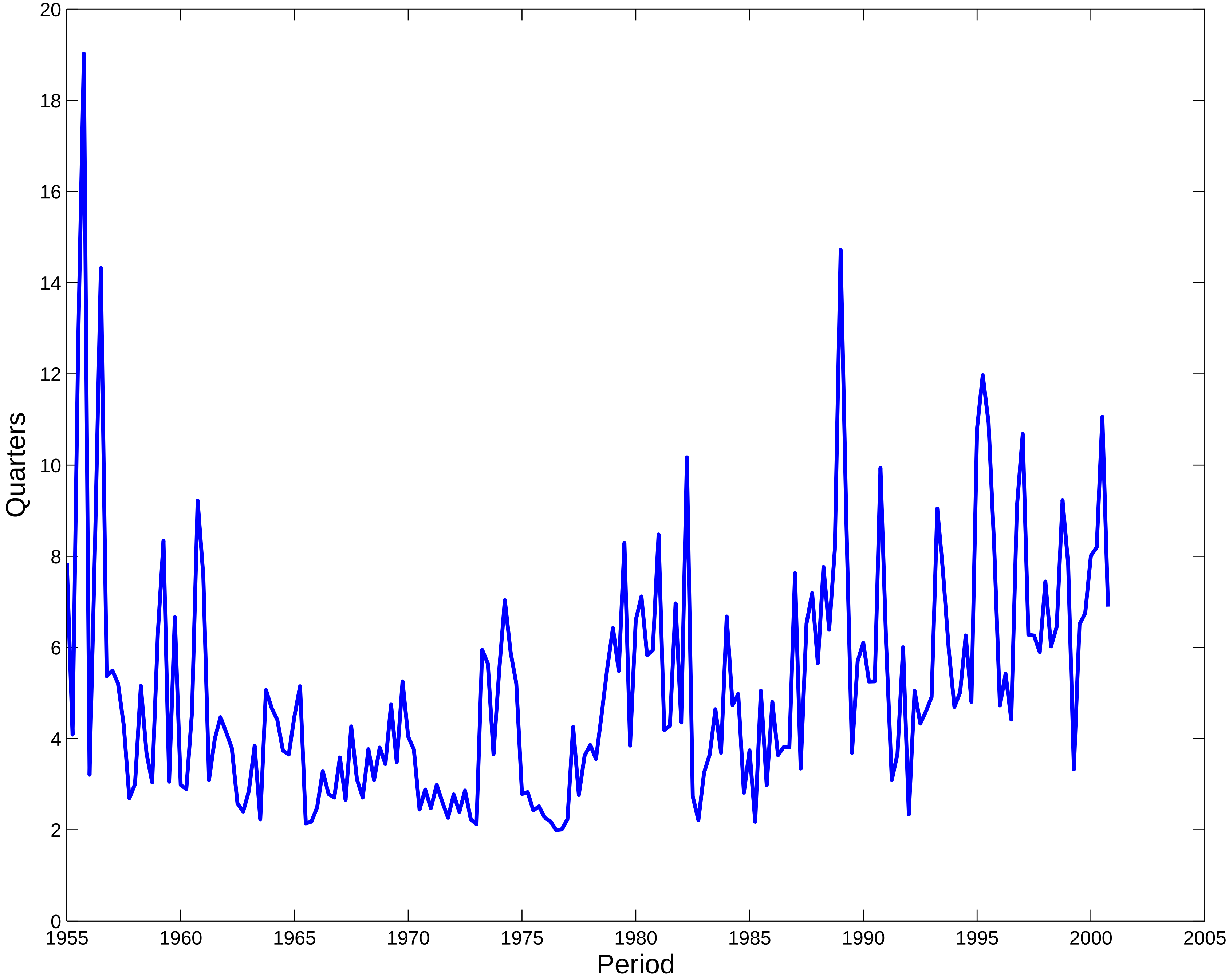


Figure 6.3.2: HP−Trend Price Rigidity versus HP−Trend Inflation

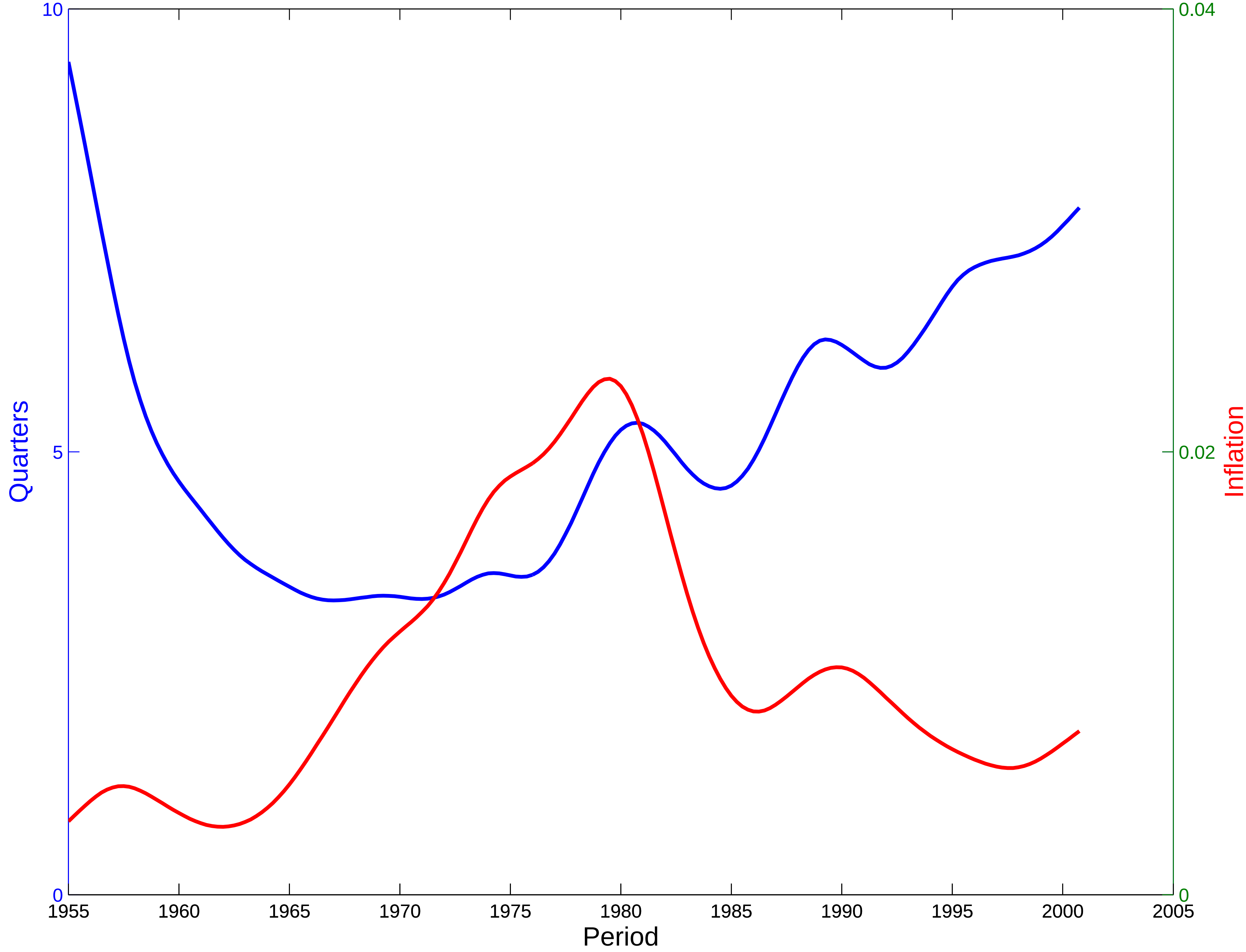


Figure 6.3.3: Price Indexation

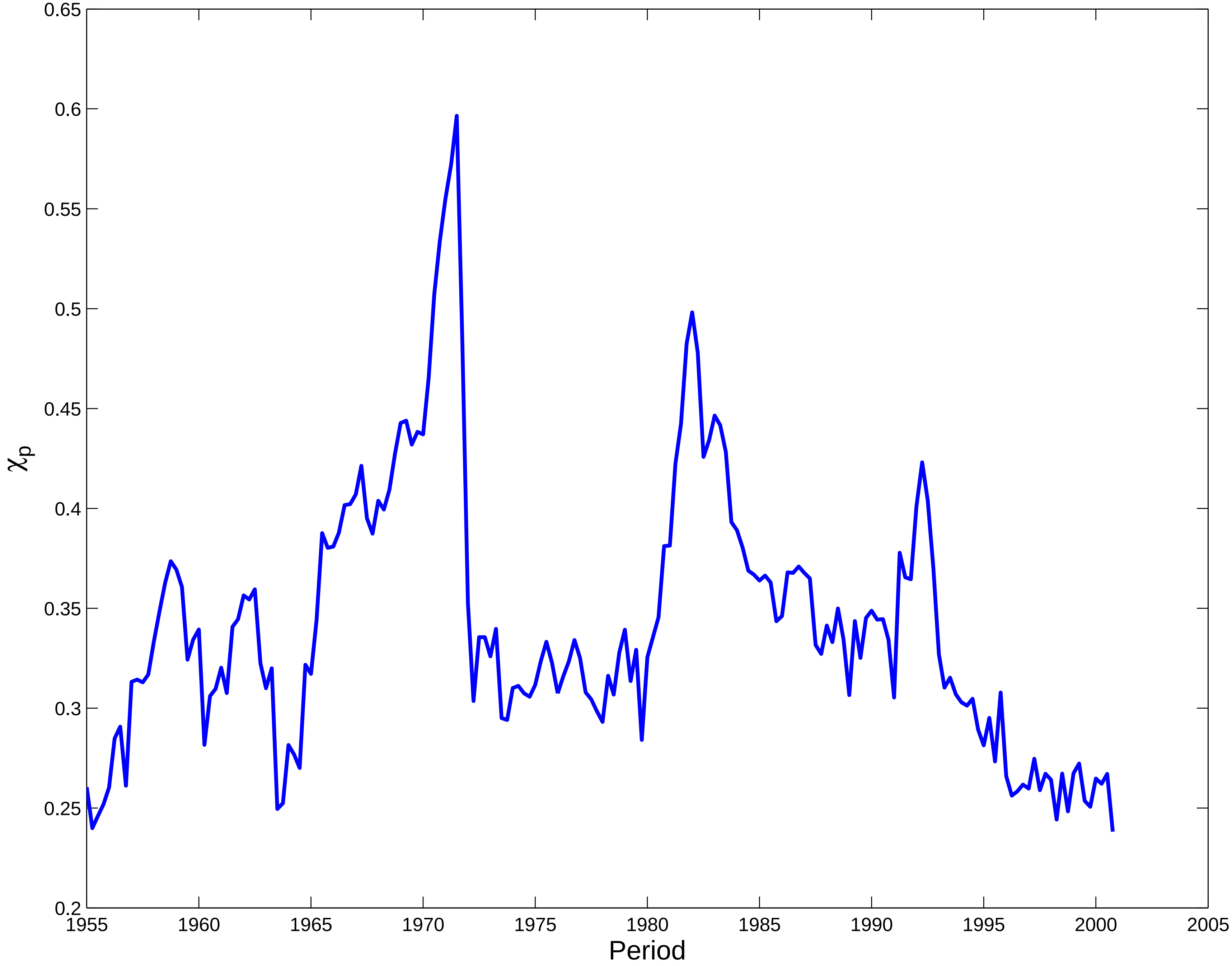
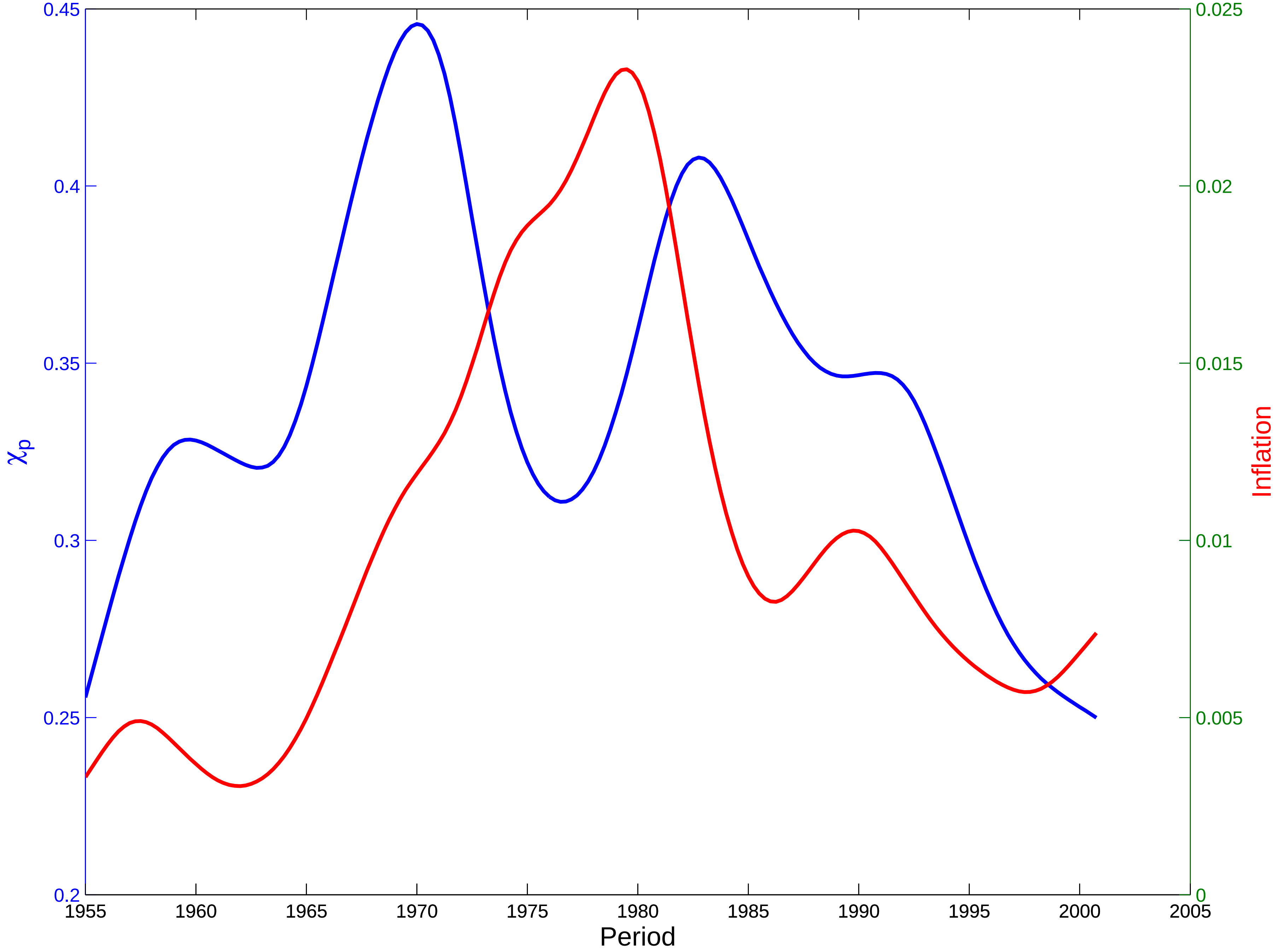


Figure 6.3.4: HP−Trend Price Indexation versus HP−Trend Inflation



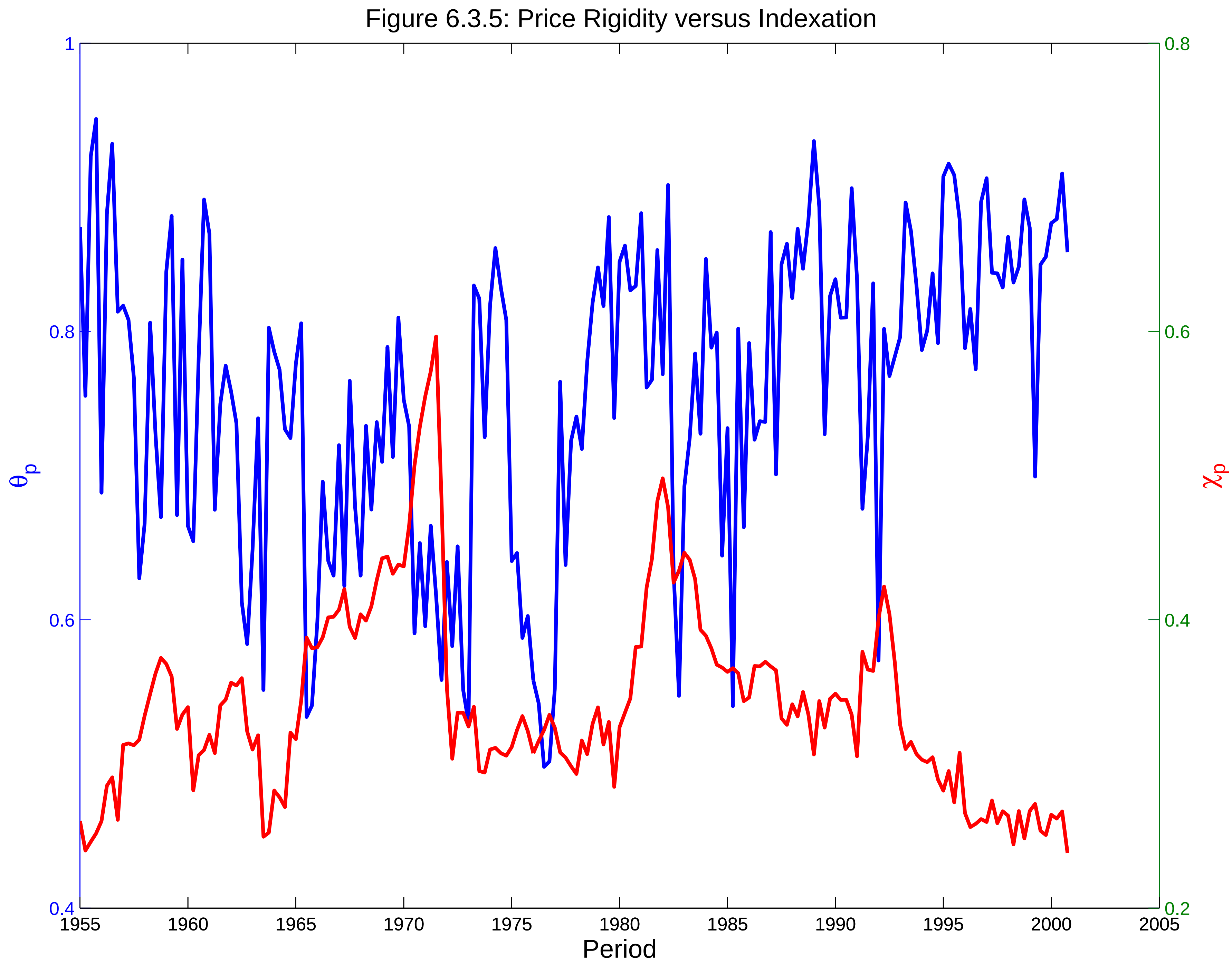


Figure 6.3.5: Price Rigidity versus Indexation
1
0.8
0.6
0.4
θ_p
0.8
0.6
0.4
0.2
χ_p
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period

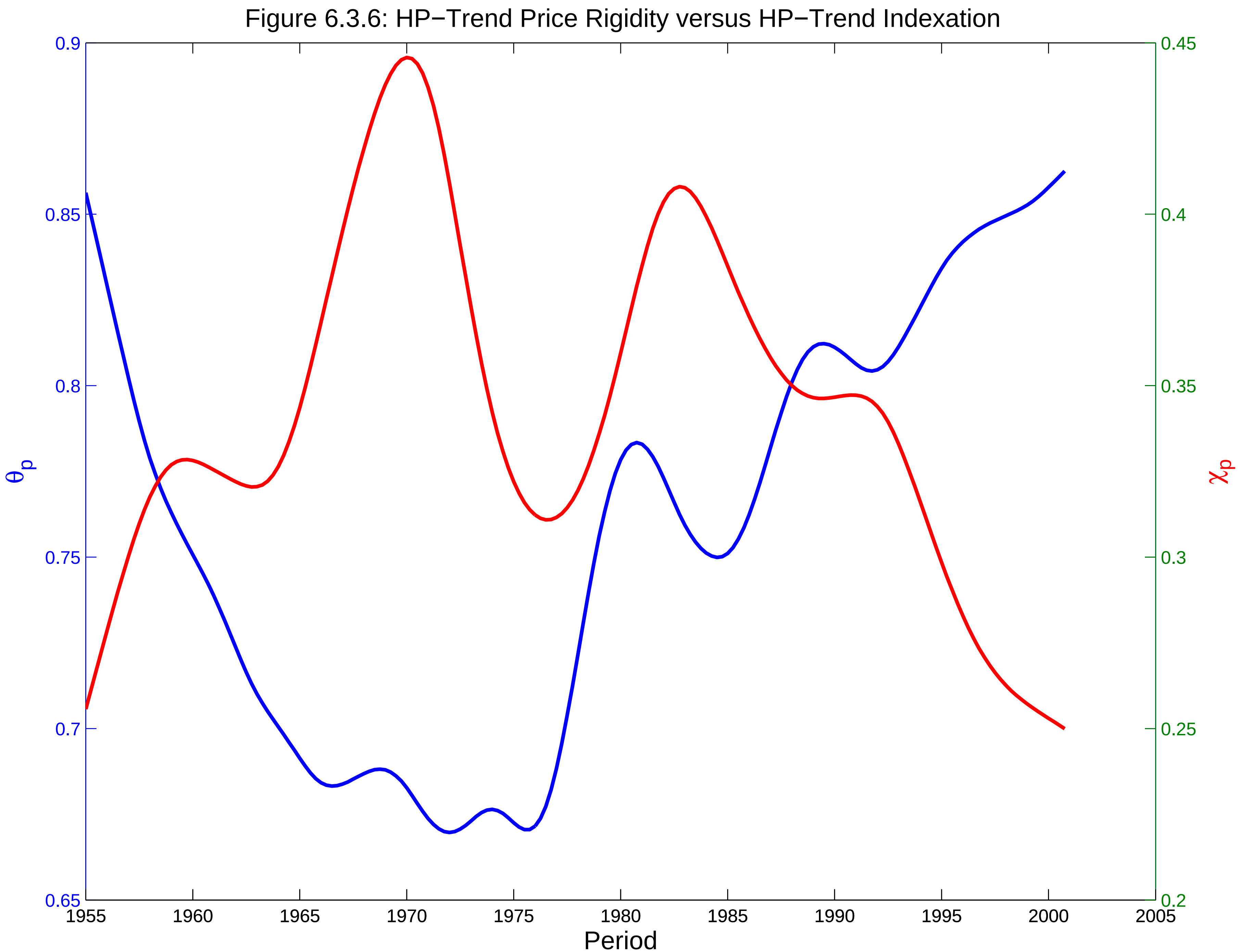


Figure 6.3.6: HP−Trend Price Rigidity versus HP−Trend Indexation
0.9
0.85
0.8
0.75
0.7
0.65
0.45
0.4
0.35
0.3
0.25
0.2
θp
χp
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period

Figure 6.3.7: Average Wage Duration

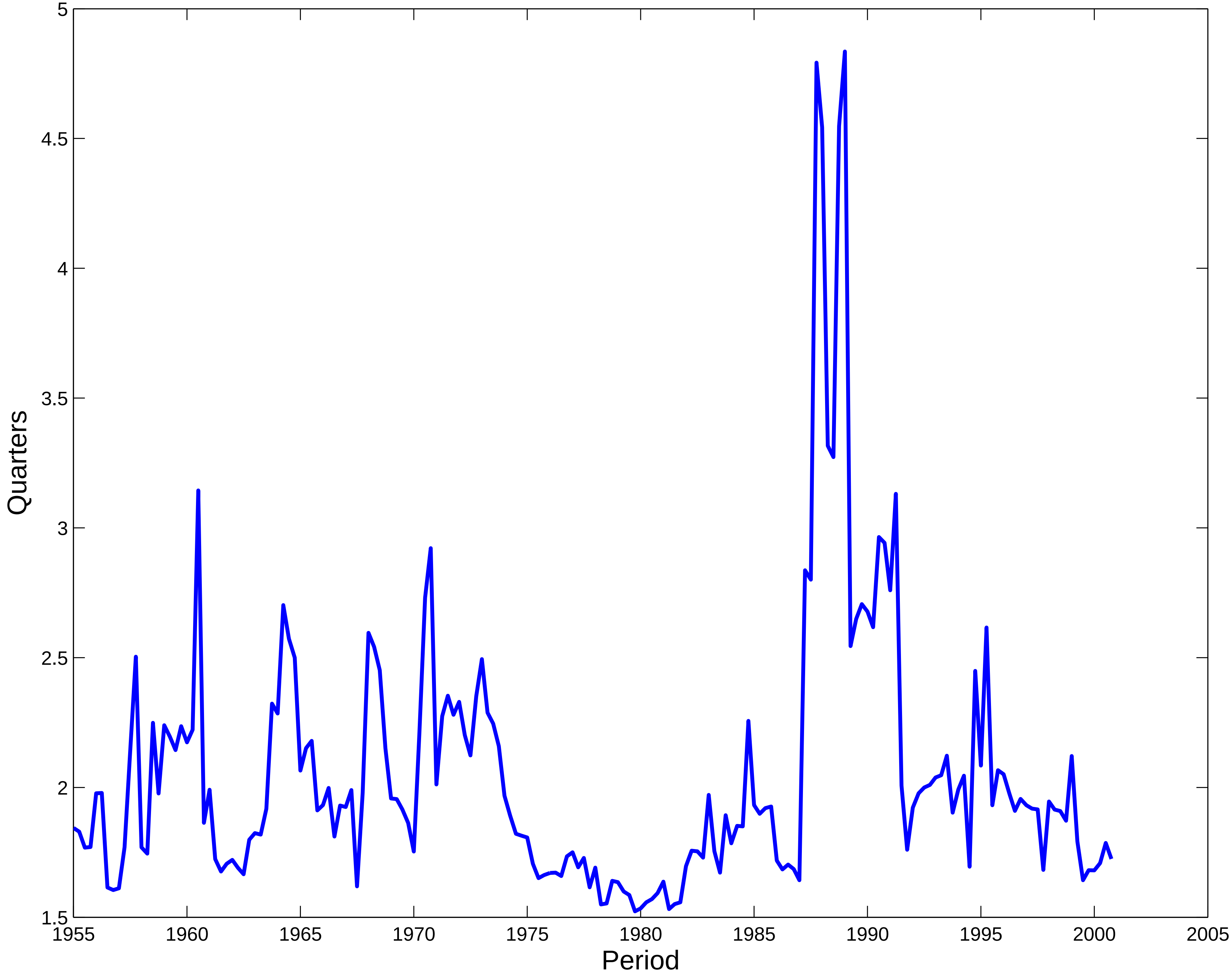


Figure 6.3.8: HP−Trend Wage Rigidity versus HP−Trend Inflation

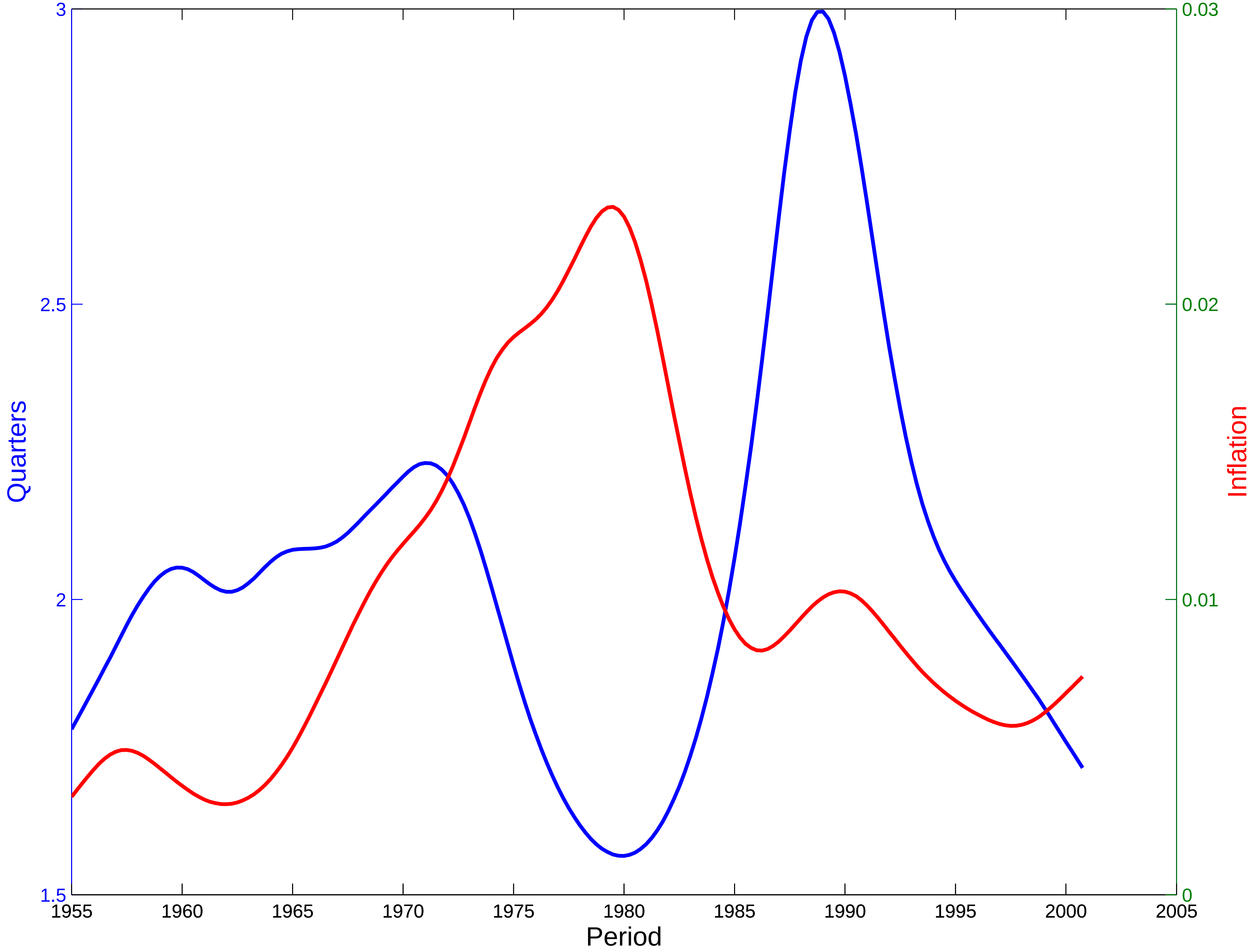
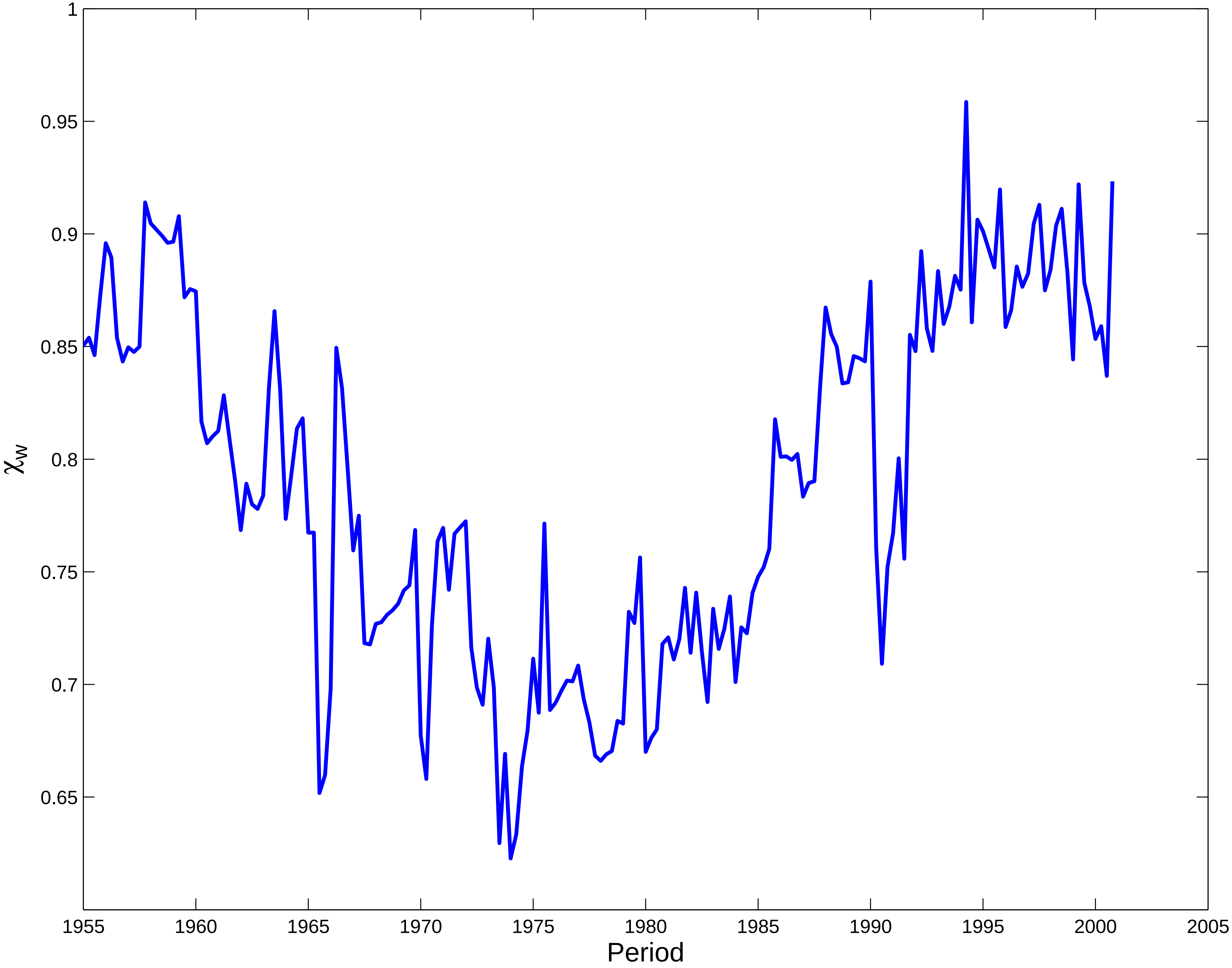


Figure 6.3.9: Wage Indexation



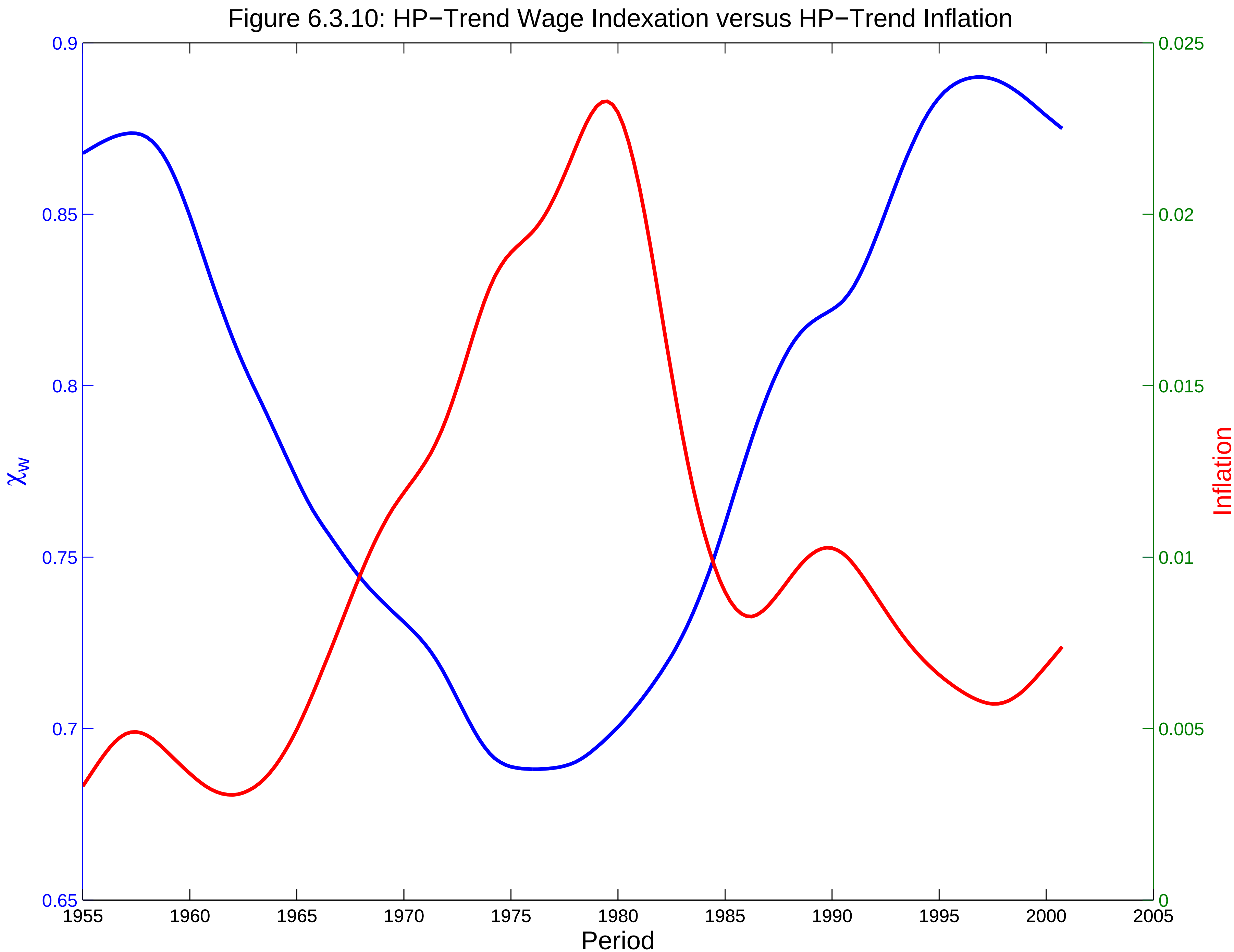


Figure 6.3.10: HP−Trend Wage Indexation versus HP−Trend Inflation
0.9
0.85
0.8
0.75
0.7
0.65
χ_w
0.025
0.02
0.015
0.01
0.005
0
Inflation
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period

Figure 6.3.11: Wage Rigidity versus Indexation

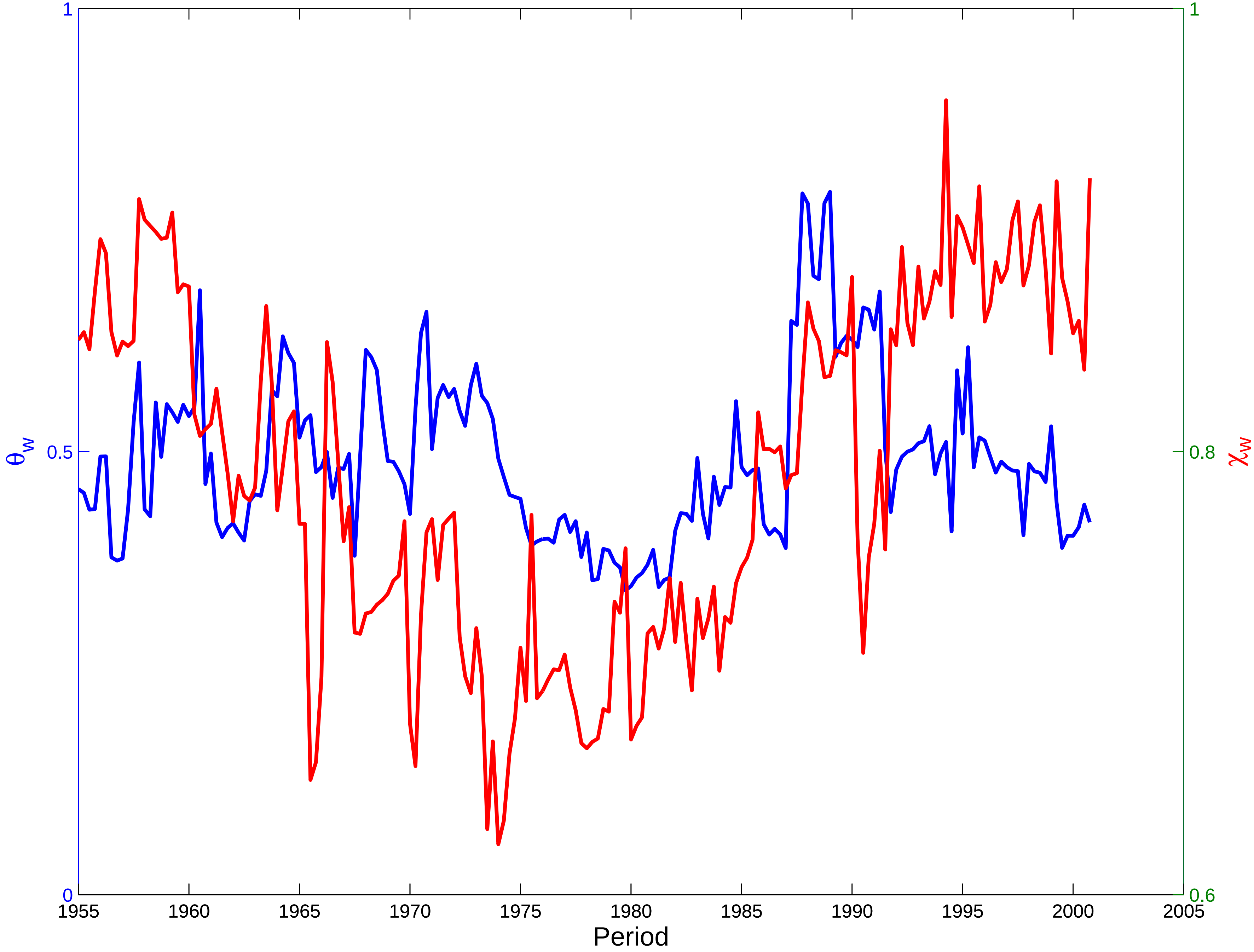
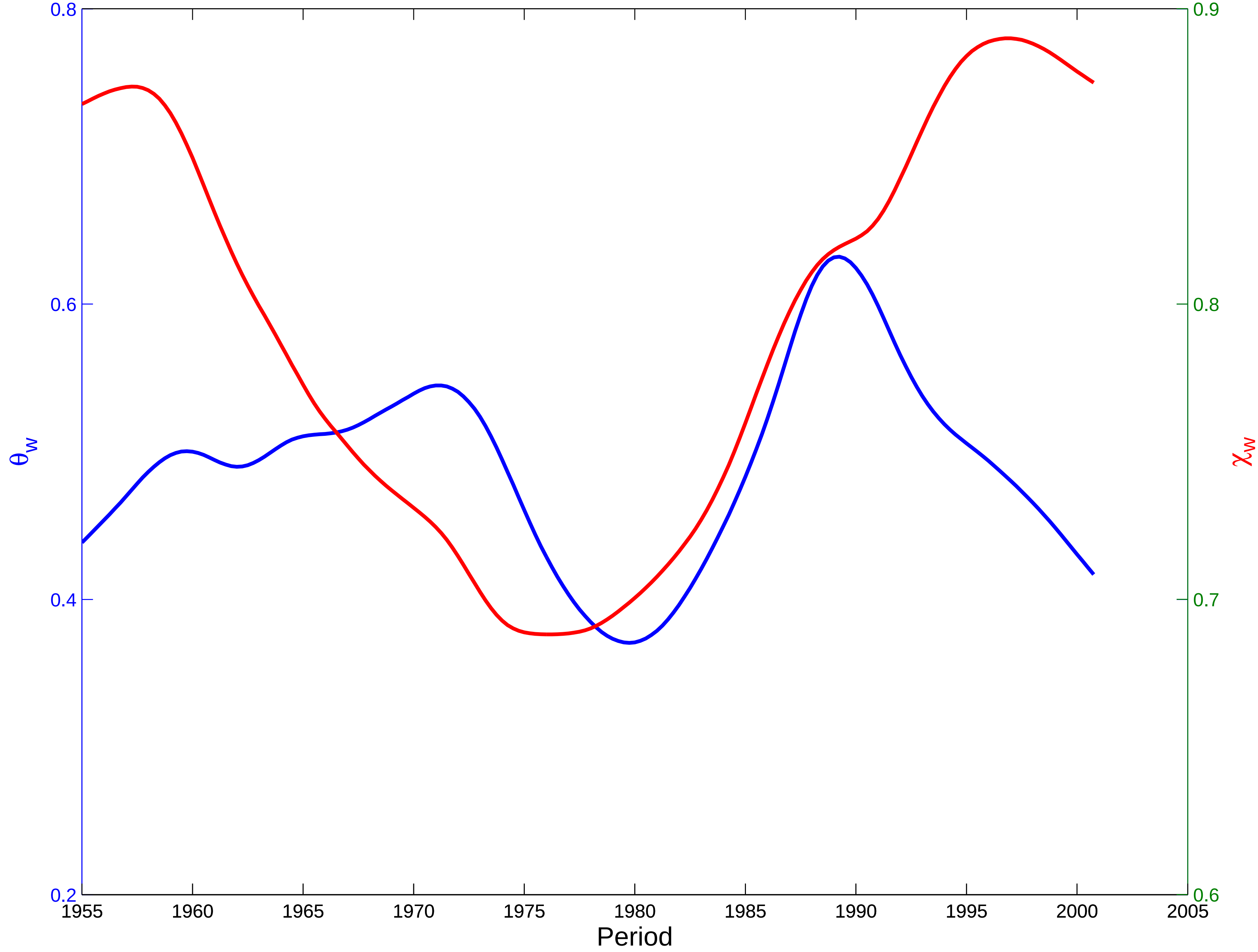


Figure 6.3.12: HP−Trend Wage Rigidity versus HP−Trend Indexation



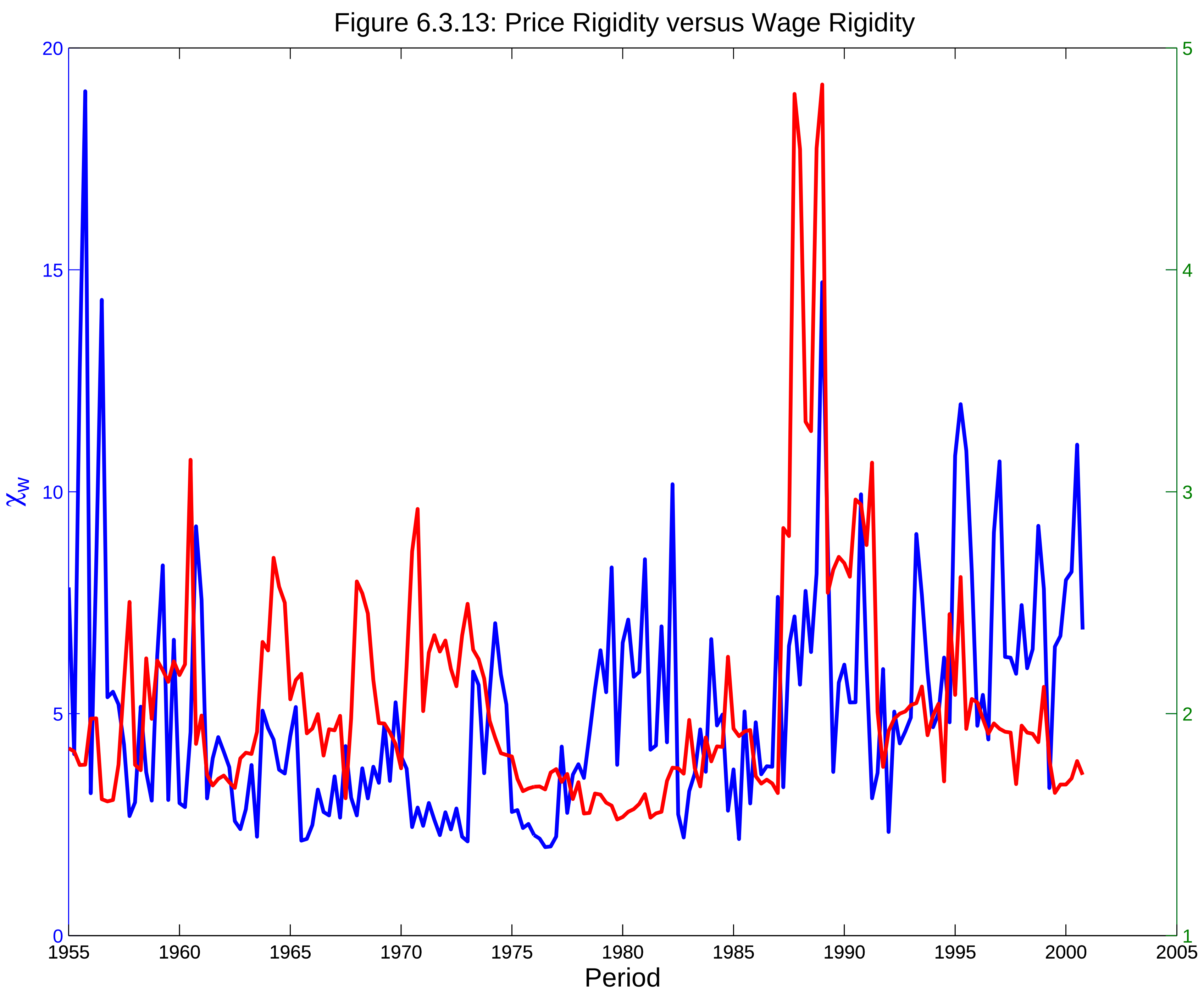


Figure 6.3.13: Price Rigidity versus Wage Rigidity
χw
0
5
10
15
20
1
2
3
4
5
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period

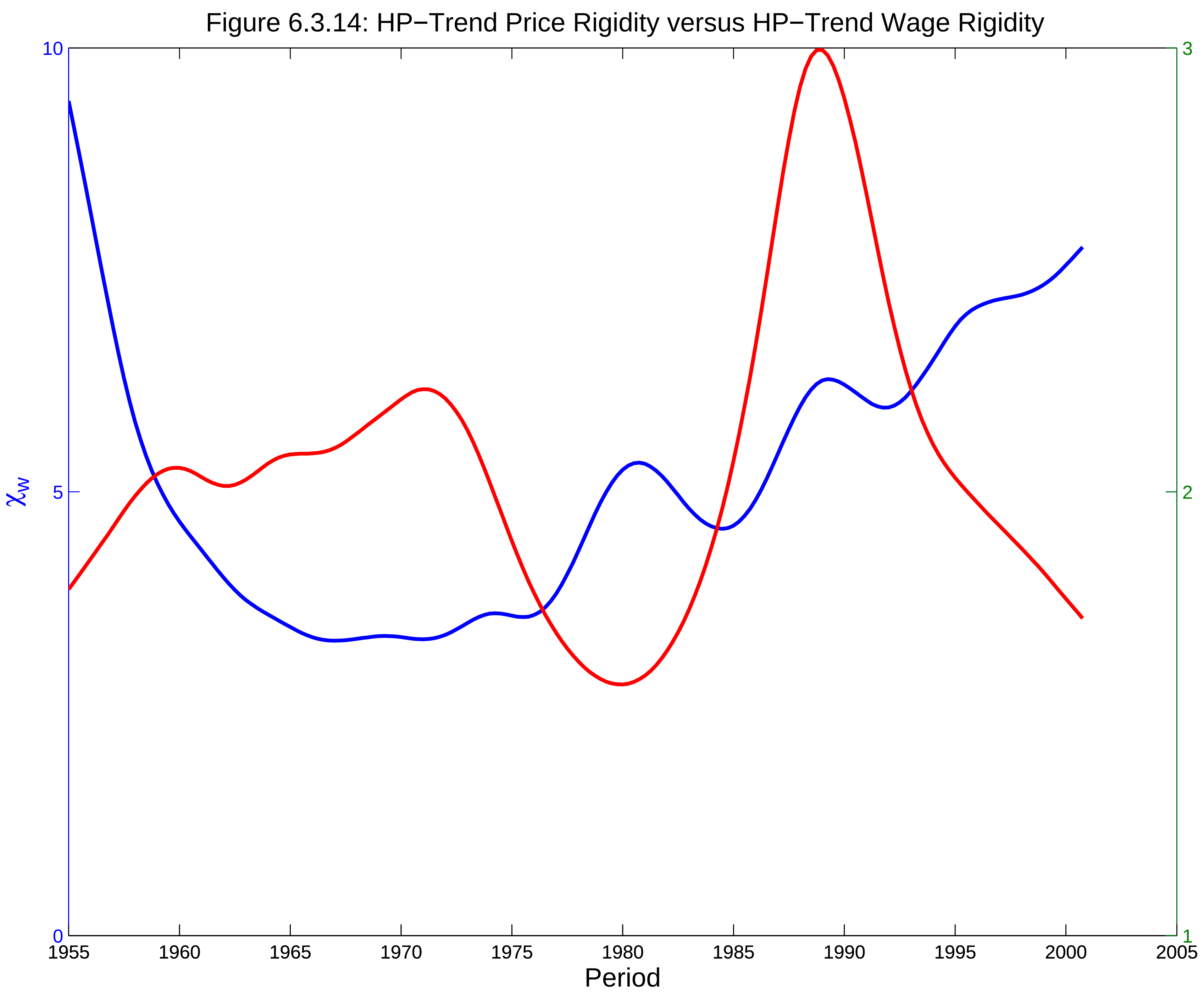
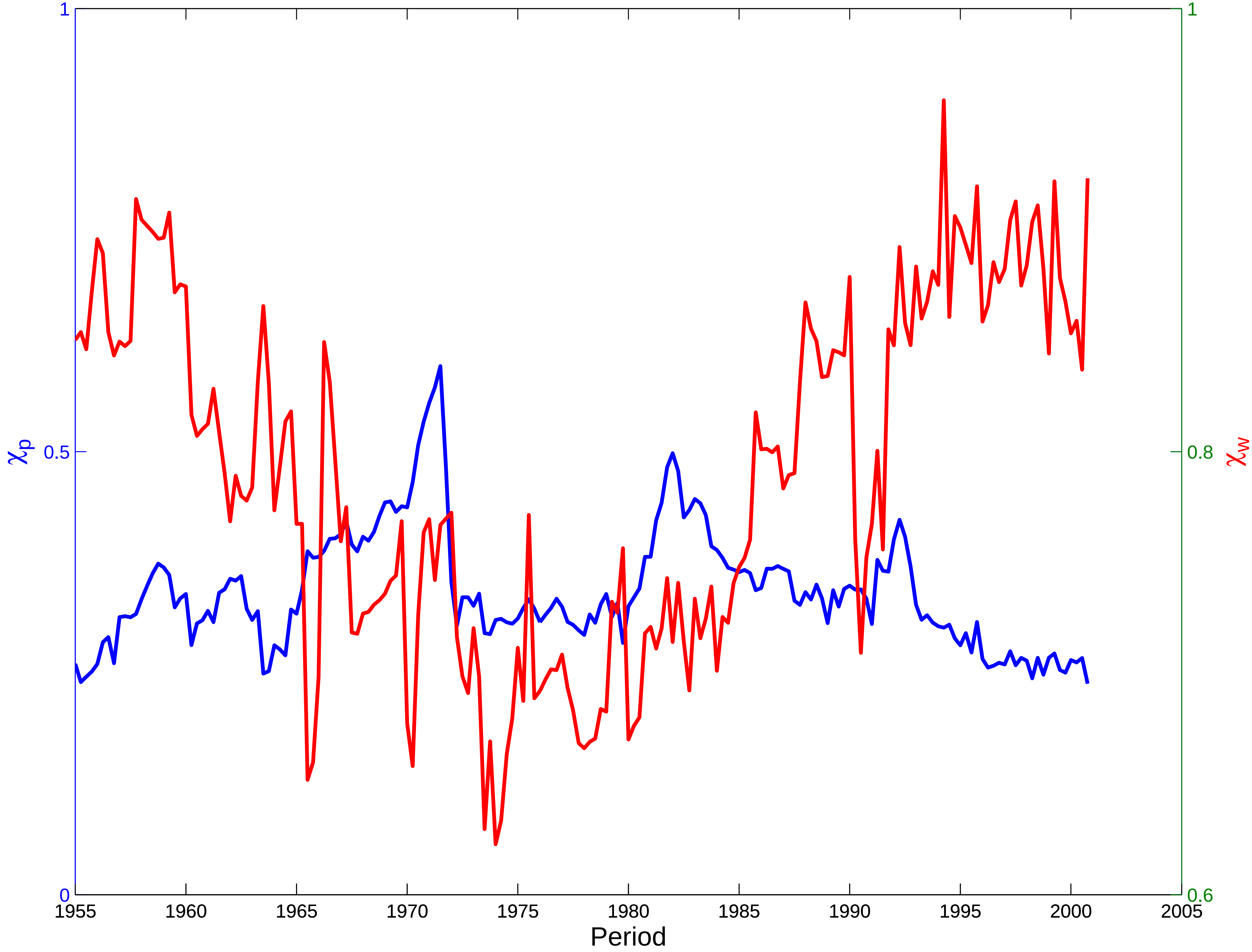


Figure 6.3.14: HP−Trend Price Rigidity versus HP−Trend Wage Rigidity
10
5
0
χ_w
3
2
1
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period

Figure 6.3.15: Price Indexation versus Wage Indexation



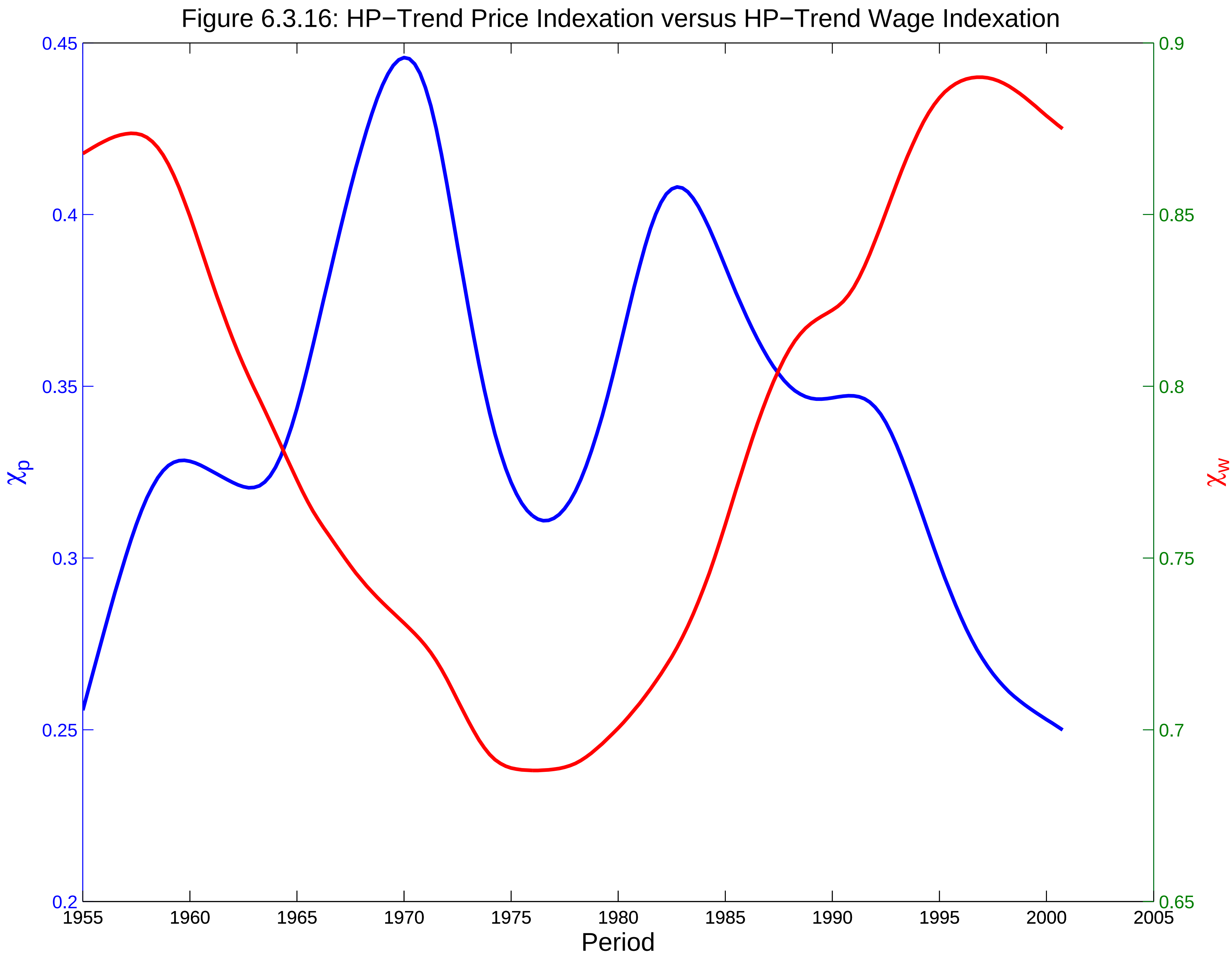


Figure 6.3.16: HP−Trend Price Indexation versus HP−Trend Wage Indexation
0.45
0.4
0.35
0.3
0.25
0.2
0.9
0.85
0.8
0.75
0.7
0.65
χp
χw
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
Period