FEDERAL RESERVE BANK *of* ATLANTA

Capital-Skill Complementarity and Inequality: A Sensitivity Analysis

Linnea Polgreen and Pedro Silos

Working Paper 2005-20
August 2005

WORKING PAPER SERIES

FEDERAL RESERVE BANK *of* ATLANTA | WORKING PAPER SERIES

Capital-Skill Complementarity and Inequality: A Sensitivity Analysis

Linnea Polgreen and Pedro Silos

Working Paper 2005-20
August 2005

Abstract: In "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," Krusell et al. (2000) analyzed the capital-skill complementarity hypothesis as an explanation for the behavior of the U.S. skill premium. This paper shows that their model's fit and the values of the estimated parameters are very sensitive to the data used: Alternative measures of the capital series predict skill premia that bear little resemblance to the data. We also include ten additional years of data to address the claim made by other authors that the evolution of the skill premium changed during the 1990s, but we find little evidence of this change.

JEL classification: C11, C82, E24, J31

Key words: capital-skill complementarity, Bayesian estimation

---

Beth Ingram, George Neumann, and Charles Whiteman provided invaluable help. In addition, we have received useful comments from Siddhartha Chib, John Geweke, B. Ravikumar, Gianluca Violante, and seminar participants at the University of Iowa and the Midwest Macroeconomics Meetings 2004. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Linnea Polgreen, Department of Economics, University of Iowa, Iowa City, Iowa 52245, 319-335-3797, 319-335-1956 (fax), linnea-polgreen@uiowa.edu, or Pedro Silos, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, Georgia 30309-4470, 404-498-8630, 404-498-8956 (fax), Pedro.Silos@atl.frb.org.

Capital-Skill Complementarity and Inequality: A Sensitivity Analysis

1 Introduction

One of the most salient observations concerning the U.S. labor market over the past thirty years is the simultaneous rise in the wage paid to college-educated workers and the increase in the supply of such workers. Theoretical explanations of this phenomenon have increasingly focused on capital-skill complementarity due to the concurrent rise in the stock of capital, and have been spurred in part by the estimates obtained in "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis" (2000) by Krusell, Ohanian, Ríos-Rull and Violante (KORV). The underlying contribution of KORV was to provide an empirical foundation for the theoretical notion of capital-skill complementarity.

KORV estimate the parameters of a constant-elasticity-of-substitution (CES) production function, using data on the prices and quantities of four factors: skilled labor, unskilled labor, structures and equipment. The functional form combines a CES aggregation of unskilled labor with an aggregation of equipment and skilled labor (also CES). This entire aggregation is combined in a Cobb-Douglas function with structures. The resulting production function is general enough to accommodate a broad pattern of substitutability and complementarity among the four factors; when KORV use this model to estimate the elasticities of substitution between skilled labor and equipment as well as that of unskilled labor and equipment, they find strong evidence of capital-skill complementarity: unskilled labor is more substitutable for equipment than skilled labor. Confidence in this result is encouraged by the ability of the model to reproduce the major changes in the skill premium over time.

KORV's paper has been widely cited, and the data and parameters estimated in this model have been used by many authors to calibrate their own models. Examples of this include Blankenau and Ingram (2002), Crifo-Tillet and Lehmann (2004), and Hendricks (2004), who use the KORV estimate of the elasticity between skilled labor and capital. Caselli and Coleman (2002) use KORV's labor series, and Lindquist (2002) uses many parameters as well as the data from KORV.

There are two data-related issues associated with the estimates in KORV. First, KORV estimate their model using data from 1963 to 1992. Ten more years of data are now available, and there is evidence that the skill premium changed during this period. Card and DiNardo (2002) find that, although the skill premium rose dramatically during the 1980s, the rate of increase slowed in the 1990s. This may reflect a different degree of capital-skill complementarity, and ten more years of data could have a substantial effect on estimated elasticities of substitution between equipment and skilled and unskilled labor respectively.

Second, KORV use a price series for capital that implies *very* rapid growth in the stock of capital equipment, starting around 1975. If the capital stock is not rising as quickly as they assume, it may not be responsible for the rise in the skill premium.

In this paper, we re-evaluate the estimates derived in KORV by studying the sensitivity of their estimates to alternative measures of the capital stock and the labor input, and by updating the estimates using the newly available data.

In addition, we estimate the model with a more stable estimation methodology. The estimation is done using a Markov Chain Monte Carlo (MCMC) algorithm following Geweke and Tanizaki (2000). This algorithm samples from the joint posterior distribution of the parameters and latent variables in a non-linear state-space model. One advantage of our methodology is that by sampling from a distribution in an efficient manner, we avoid the instabilities that can arise in maximizing a multidimensional function numerically. KORV's model provides a suitable framework for illustrating this method that is of independent interest to macroeconomists not familiar with ways of dealing with non-linear state-space models.[1]

We obtain three important results. First, using an alternative capital price index suggested by Greenwood, Hercowitz and Krusell (GHK, 1987), we find the model predicts a skill premium that bears little resemblance to the data: the results from the model significantly over-predict the skill premium at the beginning of the sample and under-predict the skill premium for most of the rest of it. That is, upon substituting one reasonable estimate of the (admittedly elusive) price of capital for another, the model's ability to fit the time path of the measured skill premium falls dramatically. Moreover, all of our estimates of the substitutability between unskilled labor and capital equipment – the measurement most-often used by other authors – are larger than KORV's point estimate of 1.67, some substantially so. We also find that the choice of labor series does not change the parameter estimates, implying that the elaborate methods used by KORV to construct a measure of the labor input may be replaced by simpler ones in future research.

2 KORV's Model

The theoretical model to be estimated is derived from a profit-maximizing firm's first-order conditions for choosing four factors of production: skilled labor (s_t), unskilled labor (u_t), structures (k_{st}) and equipment (k_{et}). The production function form combines a CES aggregation of unskilled labor and an aggregation of equipment and skilled labor in a Cobb-Douglas function with structures:

$$G(k_{st}, k_{et}, u_t, s_t) = k_{st}^{\alpha}[\mu u_t^{\sigma} + (1-\mu)(\lambda k_{et}^{\rho} + (1-\lambda)s_t^{\rho})^{\sigma/\rho}]^{(1-\alpha)/\sigma}, \qquad (1)$$

where μ and λ are parameters that govern income shares, $1/(1-\sigma)$ is the elasticity of substitution between equipment and unskilled labor, and $1/(1-\rho)$ is the elasticity of

[1] It is important to note that we obtain roughly the same estimated values for the parameters when KORV's dataset is used.

substitution between equipment and skilled labor ($\sigma > \rho$ implying capital-skill complementarity). The skilled and unskilled labor inputs, s_t and u_t are functions of hours (h_s and h_u) and efficiency indices (ψ_s and ψ_u): $s_t = \psi_{st} h_{st}$ and $u_t = \psi_{ut} h_{ut}$. In this model, the elasticity of substitution between unskilled labor and skilled labor equals the elasticity of substitution between unskilled labor and equipment.

To gain some intuition for the implications of the model, KORV derive an expression for the ratio of the marginal products of the skilled and unskilled labor inputs for the skill premium, then log-linearize and differentiate this ratio with respect to time to obtain an expression for the growth rate of the skill premium, $g_{\pi t}$. It can be decomposed into three components:

$$g_{\pi t} = (1-\sigma)(g_{h_{ut}} - g_{h_{st}}) + \sigma(g_{\psi_{st}} - g_{\psi_{ut}}) + (\sigma - \rho)\lambda \left(\frac{k_{et}}{\psi_{st} h_{st}} \right)^{\rho} (g_{k_{et}} - g_{h_{st}} - g_{\psi_{st}}). \quad (2)$$

The first component, $(1-\sigma)(g_{h_{ut}} - g_{h_{st}})$ is the relative quantity effect. Since $\sigma < 1$, the skill premium rises if the growth rate of skilled labor is less than that of unskilled labor.

The second term is the relative efficiency effect, which depends on the sign of σ; if $\sigma > 0$ and the efficiency of skilled labor is rising with respect to that of unskilled labor, the skill premium rises. The opposite happens if $\sigma < 0$.

The third term is the capital-skill complementarity effect: if there is capital-skill complementarity ($\sigma > \rho$) and the growth rate of equipment rises, the growth rate of the skill premium will also rise.

To estimate the model, the first-order conditions are simplified into three equations. The first two equations (3 & 4) are obtained by rearranging the first-order conditions for skilled and unskilled labor. The first equation,

$$\frac{w_{st} h_{st} + w_{ut} h_{ut}}{y_t} = (1-\alpha)\{\mu(h_{ut}\psi_{ut})^{\sigma} + (1-\mu)[\lambda k_{et}^{\rho} + (1-\lambda)(h_{st}\psi_{st})^{\rho}]^{\sigma/\rho}\}^{-1}$$

$$\{\mu(h_{ut}\psi_{ut})^{\sigma} + (1-\mu)[\lambda k_{et}^{\rho} + (1-\lambda)(h_{st}\psi_{st})^{\rho}]^{\frac{\sigma}{\rho}-1}(1-\lambda)(h_{st}\psi_{st})^{\rho}\}, \tag{3}$$

sets the share of labor in aggregate income ($\frac{w_{st}h_{st}+w_{ut}h_{ut}}{y_t}$) in the data equal to the analogue from the production function. The labor share, shown in figure 1, is obtained in a manner similar to that explained in Cooley and Prescott (1995), taking the ratio of compensation of workers to personal income.

The second equation,

$$\frac{w_{st}h_{st}}{w_{ut}h_{ut}} = \frac{1-\mu}{\mu}\sigma(1-\lambda)[\lambda k_{et}^{\rho} + (1-\lambda)(h_{st}\psi_{st})^{\rho}]^{\frac{\sigma}{\rho}-1}\frac{(h_{st}\psi_{st})^{\rho}}{(h_{ut}\psi_{ut})^{\sigma}}, \tag{4}$$

involves the ratio of the wage bill for skilled workers to that of unskilled workers. Equation (4) sets this ratio ($\frac{w_{st}h_{st}}{w_{ut}h_{ut}}$) in the data equal to the same ratio using the corresponding marginal products.

The third equation,

$$\frac{q_{t-1}}{q_t} = \frac{1}{(1-\delta_e)}\{(1-\delta_s) - G_{k_{st}} - q_{t-1}G_{k_{et}}\} + \epsilon_t, \tag{5}$$

is obtained from the marginal products of equipment and structures. It sets the expected return on equipment equal to the expected return on structures.

Two sources of estimation error are given by the workers' abilities, which are observed by the firm owner, but not the analyst. Also, the relative price of equipment (q_t) is not observed by the firm owner because production involves a one-period time-to-build feature: investment in equipment occurs in one period and the equipment is used in production during the following period. This uncertainty will be reflected by a forecasting error the firm owner makes when predicting prices.

Finally, KORV needs to specify a stochastic process for the vector (ψ_{st}, ψ_{ut}), the latent abilities of workers. Since KORV only want to concern themselves with changes in the skill premium due to observables, they constrain the ability factors to be without trend:

$$\phi_t = \phi_0 + \nu_t, \tag{6}$$

where $\phi_t = [log(\psi_{st}), log(\psi_{ut})]'$ and $\nu_t \sim N(0, \Sigma)$.

3 Data

Our data cover the period 1963–2001. Following KORV(2000), we have obtained almost all series from the Bureau of Economic Analysis' (BEA) National Income and Product Accounts (NIPA) and the Current Population Survey (CPS) March outgoing-rotation files. Appendix A provides detailed sources and definitions of all the series described in this section. We used various measures of capital on the one hand, and the labor input on the other. We turn first to the capital data.

To summarize the evolution of the real value of the stock of capital equipment, we need first to find an appropriate deflator. In its NIPA price series, the BEA provides an implicit deflator for the stock of equipment, but this series has been criticized for overestimating the increase in prices: it does not take quality adjustment into account. Still, NIPA prices during the last part of our sample capture some quality changes that have taken place, particularly for computers and software. Nonetheless, KORV avoid this series and use instead the alternative series provided by Gordon (1990) that incorporates quality changes by means of hedonic regressions. The data cover the period 1963–1983. Since they want to estimate the model through 1992, they construct an estimate of Gordon's price series for 1984–1992 using a regression method. In particular, they estimate the "near" vector autoregression $P_t^G = \beta_0 + \beta_1 P_{t-1}^G + \beta_2 P_{t-1}^N + \epsilon_t$, where P_t^G is a 3×1 vector of prices in the "General Industrial Equipment", the "Transportation" and "Others" sectors, obtained from Gordon's dataset, and P_t^N is the official NIPA "capital equipment price index". Using the 21 annual observations available, they construct the forecast $\hat{P}_t^G$ for $t = 1984, \ldots, 1993$ using $\hat{P}_{t-1}^G$ and actual NIPA prices P_t^N.[2] To examine the robustness

[2] Any changes in the quality adjustment procedure introduced in NIPA in the beginning of the 1980s would cause a large bias in this forecast.

of KORV's results to alternative measurements, we also consider Greenwood, Hercowitz and Krusell's (1997) quality-adjusted price series.[3] The GHK price series is the average between the Producer Price Index for capital equipment and the NIPA deflator for capital equipment. This series and Gordon's series are very different from 1963–1973 but very similar from 1974–1983. However, the forecast of Gordon's prices implies negative growth rates on average since 1984, while the GHK series, which does not require forecasting, implies a little growth (0.78% per year). Figures 2 and 3 show graphs with the levels and growth rates for these series respectively. Examining the price levels shown in figure 2, one problem with KORV's forecast of Gordon's series is evident. The series ended in 1984, just as the price level was beginning to fall. Because of this, for the remaining years, the subsequent one-lag-VAR forecasted price series falls substantially more than the other series.

These different ways of measuring prices imply *very* different capital stocks for equipment. Figure 4 shows the stock of equipment calculated using two different price series: one based on a KORV-type forecast of Gordon's prices and the second based on the GHK series. The average annual growth rates over the period were 7.6% for the Gordon-based stock of equipment and only 3.3% for the GHK-based measure, while the stock of non-residential structures has averaged only 0.82% growth. The growth rates do show a large degree of volatility, though. Still, the two series give very different pictures of the evolution of the stock of equipment: according to the GHK measure, the stock of equipment has risen by a factor of 4 since the early 1960s; according to the Gordon-VAR estimates the stock has risen by a factor of 17!

The steep fall in equipment prices and subsequent steep rise in equipment stock implied by the Gordon-based data may be driving the results obtained by KORV. Examining

[3]GHK briefly discuss this series in a footnote. They mention that Gordon suggested this series as an extension of his own.

equation (2), if capital-skill complementarity exists (*i.e.* $\sigma > \rho$), an increase in the growth rate of equipment, $g_{k_{et}}$, all else equal, will cause an increase in the growth rate of the skill premium, $g_{\pi t}$. If, however, the growth rate of equipment were not so high, capital-skill complementarity could not have been as responsible for the observed increase in the skill premium, as KORV maintain. We return to this point presently.

We next turn to the labor data. We measured labor input and wages following KORV as well as using two other methods. Details are given in Appendix A1. The CPS data are used in order to obtain wage and hours data for the skilled and unskilled separately. As KORV do, we define the skilled as college graduates, and the unskilled as those without a college degree.

As is well known, the CPS labor data is problematic: the survey asks respondents what their income was *last year*, and how many hours they worked *last week*. During any particular week, many people are on vacation or otherwise not at work. The result of this is that there are observations with income last year and no hours worked last week, so hourly wages are difficult to compute. To preserve these observations, KORV impute the missing hours by sorting non-missing observations into 264 groups based on education, gender, race and age, calculating the mean hours for each group and assigning these mean hours to the missing observations in the same group.

We created labor data according to their method. We also developed a set of labor data using linear regression to impute hours as well as a set of data where the observations with missing hours were deleted. As it turned out, none of the methods – whether imputing or deleting missing data – differed substantially in the resulting labor series they produced, suggesting that simpler methods suffice.

The most significant feature of the data is the increase in the ratio of skilled labor input to unskilled labor input. This ratio is shown in Figure 5, for both the KORV-type data and the data with missing hours deleted. Both are monotonically increasing. The average annual growth rates for the KORV-type data and the data with missing hours deleted are 2.80% and 2.78%, respectively. Unlike the capital data, minimal differences are observed in the labor data. Indeed we have conducted all of the analyses below with the three alternative labor data series, and found no qualitative differences in the results.

From these data we also calculate the skill premium, which is shown in Figure 6. There are three distinct periods: a slight increase through the sixties, a slight decrease during the seventies and a substantial increase beginning in the 1980s. During the last period, the skill premium has increased at an average rate of 1.5% annually.

4 Methodology

The three measurement equations and the stochastic specification for abilities form a non-linear state-space model. Linear state-space models have been estimated using classical approaches by Anderson, Hansen, McGrattan and Sargent (1996), and Ireland (1999), for example. Bayesian approaches have been taken by DeJong, Ingram and Whiteman (2000a, 2000b) and Otrok (2001). Recently, there has been an increasing interest in the estimation of non-linear cases, with examples such as KORV and Fernandez-Villaverde and Rubio (2002).

There are several estimation methods available. KORV use a pseudo-maximum-likelihood approach. Given parameter values and data on hours and capital stocks, they simulate the right hand side of the measurement equations (3) – (5); averaging across these simulations yields 21 annual observations for the mean vector and covariance matrix. Parameters are chosen to maximize the normal likelihood for the observed data on the left-hand-side of the measurement equations– labor's share, the wage bill ratio, and the inverse of the inflation rate in the relative price of equipment.

We adopt an alternative procedure based on the explicit assumption of measurement error (KORV did this implicitly) and the use of Markov Chain Monte Carlo methods, which are described in Appendix B. This method is related to work done in the Bayesian statistics literature by Gordon, Salmon and Smith (1993), Carlin, Polson and Stoffer (1992), and Geweke and Tanizaki (2000). Although computationally more expensive by some measures, Bayesian methods provide several advantages. The high dimensionality of the parameter space is less of a problem than in pseudo-maximum-likelihood methods, since we avoid any derivative-based numerical optimization. In addition, constraints that naturally arise from economic theory are generally easier to impose in a Bayesian framework than in maximum-likelihood estimation. Regarding the assumption of measurement errors in the first two measurement equations, we think this assumption (made exclusively for technical reasons) is innocuous, given that variances of these errors turn out to be very small.

In our application, Bayesian inference involves specifying a prior distribution for the vector of parameters $\theta = \{\sigma, \rho, \mu, \lambda, \Omega, \Psi_0, \gamma, \Sigma\}$, and coupling it with the normal measurement-error likelihood function for the data, conditional on the parameters. By Bayes' theorem the posterior distribution will be given by:

$$p(\theta|Y, X) \propto p(\theta)L(Y|\theta, X),$$

where $p(\theta)$ denotes the prior distribution, $L(Y|\theta, X)$ refers to the likelihood of the model, and Y and X are endogenous and exogenous variables respectively. The goal is a complete characterization of the moments of $p(\theta|Y, X)$, which are usually obtained by random sampling. Details of our MCMC method are provided in Appendix B.

4.1 Priors

In specifying the prior distribution for the vector of parameters, we have drawn from previous studies, especially for the two most important parameters, ρ and σ. For computational simplicity, we have restricted ourselves to use only (truncated) Normal and Gamma distributions for all parameters.

One of the reasons KORV's paper is so widely cited is that there are few alternative estimates for σ and ρ available. All of the estimates we used are taken from Hamermesh's 1993 survey of labor demand and are consistent with what little literature there is. For σ, we specified a Normal $(0.57, 0.25^2)$ truncated to the $(-\infty, 1]$ region.[4] The parameter driving the elasticity of substitution between capital equipment and skilled labor, ρ, was endowed with a normal prior distribution with the mean set at -0.76.[5] The standard deviation and the truncated region of ρ are the same as those of σ.

The shares, λ and μ, were given prior Normal distributions truncated to the $[0, 1]$ range, with a mean of 0.5 and a standard deviation of 0.2. The standard deviation of α was set at 0.005, with a mean of 0.11, which is the value estimated by KORV, which in turn is close to the value of 0.13 used in previous calibration studies (*e.g.* Greenwood, Hercowitz and Krusell (1997)). In all cases, the prior standard deviations were chosen so that a two-standard deviation band around the mean provided a reasonable range of estimates.

[4]The mean was chosen to match the estimate obtained by Clark and Freeman (1977) in their annual (unpublished) sample of the manufacturing sector.

[5]This value is halfway between the value of 0.08 from Berndt and White's unpublished 1978 paper on the demand for energy, and the value of -1.6 estimated in Dennis and Smith's 1978 study of the demand for real cash balances. Both of these studies focus on the manufacturing sector, roughly covering the period 1950-1973. Although both of these studies are rather dated, and energy and cash balances do not seem to be related to the topic at hand, these were the best estimates we could find.

5 Results

We present the results for five sets of data. The first set is KORV's own data and the other sets are ours, which include two sets of capital series, each using data from 1963-2001 as well as 1963-1992 (the period examined by KORV). All models were estimated with a labor series generated using KORV's method.[6] The results are given in Table 1. For the capital series used, Gordon refers to the estimates using the Gordon-based VAR, and GHK refers to the series from GHK (1997). KORV1 are KORV's own results, and KORV2 are the results using our estimation method and KORV's own data.

Table 1: Results

		σ		ρ	
Capital Data	Years	Mean	Std. Dev.	Mean	Std. Dev.
KORV1	1963-1992	0.401	0.234	-0.495	0.048
KORV2	1963-1992	0.446	0.012	-0.384	0.025
Gordon	1963-1992	0.520	0.019	-0.167	0.031
GHK	1963-1992	0.557	0.003	-0.569	0.012
Gordon	1963-2001	0.503	0.022	-0.239	0.031
GHK	1963-2001	0.842	0.063	-0.657	0.076

Like KORV, all of our estimates imply capital-skill complementarity ($\sigma > \rho$), although many differ quantitatively from theirs. Specifically, all of our estimates for σ are larger than theirs. The estimate using the longer set of GHK data is substantially larger. Our values for ρ range from -0.167 to -0.657, and all are significantly different from the estimate of -0.495 provided by KORV. The posterior distribution of the difference between σ and ρ is shown in Figure 7. The dotted curve represents the GHK data and the solid curve represents the VAR data. The differences created by the two data sets are clear: the GHK data implies a much larger difference between the two parameters.

Adding the extra years of data did not substantially alter many of the estimates. With

[6]Results using other labor series are not presented because they are similar to those using KORV's method.

respect to σ, the only substantial difference occurs when using the GHK data. The longer data set (1963–2001) produces a σ that is much larger than that of the shorter data set (1963–1992). However, for both capital series, the estimates for ρ for the larger data set are significantly smaller than those for the shorter data set.

The equipment-skilled labor and equipment-unskilled labor elasticities are given in Table 2. The estimates for the elasticity of substitution between unskilled labor and equipment are between 2 and 9. Compared to KORV's estimate of 1.67, our estimates are much higher: we estimate unskilled labor to be more substitutable for equipment than they do. However, considering the large standard error on KORV's estimate of σ, only the estimate using the longer data set and the GHK capital series differs significantly. The posterior distributions of these elasticities are given in Figure 8. Again, the dotted curve represents the GHK data and the solid curve represents the VAR data. The variance of the posterior using the KORV data is less than 3% of the variance of the posterior using the GHK data.

Considering the elasticity between equipment and skilled labor, all of our estimates for ρ are significantly different from those of KORV. Figure 9 gives the posterior distributions of estimates of the elasticity between equipment and skilled labor using the Gordon-type VAR data (solid curve) and the GHK data (dotted curve). The elasticities produced using the Gordon-based VAR are much larger. These elasticities imply that skilled labor is also more substitutable for equipment than KORV estimate.

Table 2: Elasticities

Capital Data	$1/(1-\sigma)$ Equipment-Unskilled Labor 1963-1992	1963-2001	$1/(1-\rho)$ Equipment-Skilled Labor 1963-1992	1963-2001
KORV	1.669		0.669	
Gordon	2.087	2.017	0.857	0.8084
GHK	2.255	9.052	0.637	0.602

Most importantly, these estimates generate skill premia that differ from the data. Figures 10 – 13 show the skill premia calculated from the data and those implied by the model. The dotted lines show the results of the models, and the solid lines represent the data. Figures 10 and 11 use the GHK capital data; figure 10 shows the shorter data set, and figure 11 the longer. The skill premia generated by the model do not follow those of the data. Both show a falling skill premium until 1982, a period in which the actual skill premium was relatively flat. In the longer series, the skill premium generated by the model rises after 1982, but its rate of growth increases after 1995, unlike the data.

Figures 12 and 13 show the skill premium using the VAR-generated data, with 12 showing the shorter data set. Although the model follows the general trends in the data, the skill premium from the model falls when the skill premium from the data rises and vice versa. That is, our "KORV-style" data results are qualitatively similar to KORV's own results: they only claim that their model captures the rising skill premium in the 1960s, the falling skill premium in the 1970s, and the steeply rising skill premium in the 1980s. This can be seen in figure 14, where we use KORV's data. The key result is the distinction between Figures 10 and 11 on the one hand, and 12 and 13 on the other. Because our "KORV-style" results in Figures 12 and 13 are so similar to KORV's conclusion (our version of which is in Figure 14; their Figure 8 is quite similar), we can be confident tat had KORV used the GHK data correction procedure in place of what they used, they would have produced a very different (and not very convincing) picture of the evolution of the skill premium – i.e., the poor-fitting Figure 10. Evidently, the results from KORV's model do not fit the model data well, and when an alternative and quite reasonable time series for equipment prices is used, the fit worsens substantially.

6 Conclusion

We re-estimated the model in "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis" (2000) by Krusell, Ohanian, Ríos-Rull and Violante using Bayesian methods, alternatively-specified data series and ten extra years of data. We began by estimating the changes in capital-skill complementarity and the skill premium by adding the most recent years of data. However, because KORV generated their data series in a particular way, we examined how these changes affected their results. No matter what series we used, like KORV, we find some evidence of capital-skill complementary. However, KORV's model is extremely sensitive to the data used, especially the capital series. Using the capital equipment data generated by the Grenwood, Hercowitz and Krusell (1997) (GHK) method, we find unskilled labor to be more substitutable for equipment than KORV do. In fact, our estimates of the unskilled labor-equipment elasticity are up to five times larger than KORV's. Using the GHK capital data also produced estimates of the skilled labor-equipment elasticity that are noticeably smaller than KORV's, and our VAR-generated series produced estimates that were larger than KORV's. Although these differences are interesting in and of themselves, what ultimately matters is if they lead to different conclusions when used in other contexts. In our case, when used to generate skill premia, the results are dramatically different. In fact, the GHK data series produced a skill premium that counterfactually falls for most of the shorter series (1963–1992), and half of the longer data series (1963-2001).

Unlike Card and DiNardo (2002), in no data series used do we find evidence that the growth of the skill premium has slowed. On the contrary, using the GHK data produces a rapidly accelerating skill premium during the 1990s.

KORV also use a tedious method to impute missing hours in their labor data. We find that simpler methods, specifically imputing missing hours using linear regression or simply deleting the observations with missing hours, do not substantially affect the labor series or the results of this model. In this case, simpler methods for compiling labor data are sufficient.

We conclude that one should use caution when using KORV's estimated parameters to calibrate models. Although the model was not sensitive to the labor series used, it was sensitive to the price series for capital equipment and the resulting capital equipment series: alternative, reasonable equipment price series lead to very different estimates of the elasticity of substitution between unskilled labor and equipment, and to very different views of the evolution of the skill premium.

A Data Appendix

A.1 Labor Input and Wages

Following the method of KORV, we use the CPS data set. We recorded variables for age, gender, race, education, weeks worked last year, hours worked last week, income from wage and salary, and the CPS sampling weights. We eliminate those younger than 16 or older than 70. We exclude anyone who does not have income from wages or salary (including the self-employed). We eliminate observations that are missing weeks worked or have unreasonable hourly wages.[7]

To show how the data set is constructed, we use 1988 as an example. The raw CPS data set for that year includes 144,687 observations. We eliminated 72,801 observations for missing income, 620 for missing the number of weeks worked last year, 3 for unreplaceable missing hours (discussed later), and 5722 for unreasonable wages. The resulting number of observations was 65,541.

Since the CPS asks what one's income was last year and how many hours one worked last week, often observations that include income have missing hours: interviewees who were on vacation or on any other type of leave last week would have income from last year but no hours worked. The number of observations in this category is often substantial. KORV did not want to eliminate observations with reported income, so they imputed missing hours. The first thing we do is follow their method. Secondly, we impute hours using a linear regression, and thirdly, we eliminate all observations with missing hours.

A.1.1 Method 1

Following the method presented in the appendix of KORV, the sample is divided into 264 groups consisting of 11 age groups (5 years / group), 3 race groups (white, black and

[7]Following Card and DiNardo (2002), we consider unreasonable wages to be less than \$1 or greater than \$100 in 1979 dollars.

other), 2 genders, and 4 education groups (no high school diploma, high school diploma, some college, college graduate). For those with missing hours or hours equal to zero, hours are set to the weighted average hours of the other members of the observation's group that year. For a few cases each year (one in 1989) there were observations with missing hours for which no imputation could be made: there were no observations in that group for that year that reported hours.[8] Groups are then categorized as either skilled or unskilled. In skilled groups, all members have 16 or more years of education, and in unskilled groups, all members have fewer than 16 years of education.

The measure of skilled labor input is the sum of the total annual hours of all the members of all skilled groups weighted, not only by the CPS weights (so that the sample more accurately represents the U.S. population), but also by the average wage of the group in 1980. This is a common practice.[9] In our final model, we weight the labor input by the average group wage in 1996.[10] The unskilled labor input follows analogously.

To calculate the labor input series, we used their formula

$$N_{j,t} = \sum_{g \epsilon G_{j,t}} l_{g,t} w_{g,96} \mu_{g,t},$$

where l are hours, w are wages, μ are the CPS weights, j is an indicator of skill, t is time, and g is the demographic group. In our final model, we divide the ability index (the wage in 1996) by the average wage in 1996 for each skill level.[11] In this way, during periods of wage inflation, the labor input measure does not inflate as well. This also creates a labor

[8]Some groups had few or no members. For example, consider the group of non-white, non-black females aged 16-20 with college educations.

[9]In "Accounting for Slower Economic Growth: The United States in the 1970s", Denison (1979) writes that in calculating an index of hours worked by demographic group, it is appropriate to use hourly wages as weights. If one's wage represents one's marginal product, those with higher wages represent a larger amount of labor input per hour.

[10]Three labor series using ability indices based in years 1980, 1996, and 2001 were considered. There were no substantial differences among them, so we chose the 1996 series to be consistent with the capital data.

[11]KORV probably did a similar data transformation but did not report it. If we do not divide by the average wage, our indices do not match theirs.

input series in units of hours.

We also calculated the wage series using the method given in KORV's appendix, where

$$W_{j,t} = \frac{\sum_{g \epsilon G_{j,t}} w_{g,t} l_{g,t} \mu_{g,t}}{N_{j,t}}.$$

The numerator is the total wage bill: the average wage in the group times the average labor input in the group, weighted by the CPS weights. This is divided by their measure of labor input, N, to get the wage series for both the skilled and the unskilled. With this specification, the wage series is essentially a ratio between the wage in the current year and the wage in 1980 (1996 in our case). Because these values are nominal, and prices are rising during the period, what should result is a series where the wage is below 1.0 for most of the years before 1980, above 1.0 for most of the years after 1980, and equal to 1.0 in 1980. For both the skilled and the unskilled, the wage should equal unity in 1980, resulting in a skill premium also equal to unity in 1980. This is not what they report. KORV reports the skill premium with respect to 1963, and for many years in their series, the skill premium actually falls below one, implying that for 1980, for example, the skill premium was less than it was in 1963. This has not been reported by other authors,[12] and by adjusting the ability index in the labor input (as discussed earlier), we have eliminated this problem. If we set the skill premium equal to one in 1963, our skill premium series resembles that of KORV but does not fall below one.

A.1.2 Method 2

The method KORV used to generate missing hours is equivalent to running a regression with 264 dummy variables. By putting all observations into groups, all variation within the groups is lost. A simpler method, and one that uses all available information, would be to impute missing hours by estimating an equation for hours using the same variables

[12]See Autor and Katz, 1999 for a review of the literature.

KORV did. So, for the second method, for each year we estimated an hours equation using age, age^2, female, black, white, and education as covariates. We then predicted hours for each missing observation. The new hours variable is multiplied by the weeks worked last year variable to obtain annual hours for each observation. The annual hours are weighted by the CPS weights and the wage index and then summed over the skilled and unskilled categories to obtain the hours series.

The wages are calculated by dividing income from wage and salary by the new annual hours variable for each observation. A wage series is created by taking the CPS-weighted mean wage for both skill categories.

A.1.3 Method 3

Since both methods one and two use existing data to impute missing hours, keeping the observations with missing hours may not add any extra information. So, for the third method, we eliminated all observations with hours missing. The labor input and wage series were aggregated in the same way as in method two.

A.2 Capital Stocks and Prices

Obtaining prices for capital equipment that adequately reflect quality adjustments in capital equipment has been a concern for data-collecting agencies and institutions. Prices constructed by the Bureau of Economic Analysis (National Income and Product Accounts, NIPA) underestimate changes in quality, leading to mismeasurement when these series are used to deflate nominal quantities. In this paper, we have used capital equipment price series adjusted for quality, obtained in two different ways.

The first method follows KORV using Gordon's (1990) price series for capital equipment. This series ends in 1983, hence we need to project prices until 2001. This was done by fitting a VAR to Gordon's data using some covariates (*e.g.* prices provided by NIPA) to help us forecast. Two different models were fitted. The first was just a bivariate VAR (in growth rates) with the aggregate capital equipment price index provided by NIPA and Gordon's capital price series. Projected prices were obtained by using the information in NIPA prices which are available until the end of the sample. The second model fitted was a "disaggregated" VAR, which included price series for different sectors (industrial, transportation and other equipment) in NIPA, to forecast prices. Results from the estimation are shown in Tables 3 and 4. The resulting series were aggregated with the NIPA price series for software and communications equipment, which takes into account quality changes using the methodology described in Cole, *et al.* (1986).

Table 3: 2-Variable VAR results
(Equation 1is Gordon price index and Equation 2 is NIPA investment price index)

Variable	Coefficient	Standard Error
Constant Eq. 1	-0.0214	0.009
Gordon Price$\{t-1\}$	-0.28410	0.1798
NIPA price$\{t-1\}$	1.10	0.254
Constant Eq. 2	0.0104	0.007
Gordon price$\{t-1\}$	-0.2136	0.1325
NIPA Price$\{t-1\}$	0.739	0.187

Table 4: 4-Variable VAR results
(Order of equations: Gordon price index, Industrial Equipment (IE), Transportation Equip. (TE), Other Equip. (OE))

Variable	Coefficient	Standard Error
Constant Eq. 1	-0.0247	0.009
Gordon Price$\{t-1\}$	-0.436	0.230
IE price$\{t-1\}$	-0.469	0.413
TE price$\{t-1\}$	0.763	0.271
OE price$\{t-1\}$	0.900	0.52
Constant Eq. 2	0.016	0.010
Gordon price$\{t-1\}$	-0.370	0.256
IE price$\{t-1\}$	0.253	0.460
TE price$\{t-1\}$	0.246	0.301
OE price$\{t-1\}$	0.370	0.582
Constant Eq. 3	0.0004	0.008
Gordon price$\{t-1\}$	-0.248	0.201
IE price$\{t-1\}$	0.759	0.360
TE price$\{t-1\}$	-0.305	0.237
OE price$\{t-1\}$	0.249	0.459
Constant Eq. 4	0.008	0.009
Gordon price$\{t-1\}$	-0.448	0.22
IE price$\{t-1\}$	-0.305	0.397
TE price$\{t-1\}$	0.387	0.269
OE price$\{t-1\}$	0.967	0.503

The second method follows the suggestion of Greenwood, Hercowitz and Krusell (GHK,1997) of combining NIPA prices with the Producer Price Index (PPI) for capital equipment constructed by the Bureau of Labor Statistics. PPIs try to capture some quality adjustment using various techniques, such as hedonic regressions or cost estimates of changes in characteristics provided by the producer, as explained in the BLS Handbook of Methods, Chapter 14. The final price index suggested by Greenwood, *et al.* (1997) is the average of the PPI and the NIPA price indices for capital equipment.

Once one has an appropriate price index for capital equipment, it is used to deflate the nominal series. To obtain the capital stock, we used investment data with the fixed rate of depreciation used in the model, 0.125. The initial stock, (for 1962) comes from Table 2.1 in the Fixed Assets Tables (BEA) and investment numbers from Table 5.4 of the NIPA Tables. The series is constructed using the perpetual inventory method and is deflated using the price index obtained by one of the methods described in the previous paragraph. The final price of capital in terms of the numeraire in the model (q_t) will be the ratio of this price index to the implicit price deflator for personal consumption expenditures of non-durables and services (Table 2.3.4, NIPA). Figure 4 shows the stock of equipment in our sample.

Obtaining the series for capital structures is simpler. The assumption in the model is that investment in structures has the same price as the numeraire consumption good. Hence, the price series used to deflate the nominal series is simply the implicit price deflator for non-durables and service consumption. The nominal stock series was constructed analogously to the equipment series by combining an initial capital stock for 1962 (Table 2.1, Fixed Assets Tables) and data on investment in non-residential private structures (Table 5.4, NIPA). The depreciation rate used was 5%.

B Methodology Appendix: Metropolis Within Gibbs algorithm

From Section 4 we obtained the following nonlinear state space form:

$$y_t = f(X_t, \theta, \Psi_t) + w_t \quad (measurement\ \ equation) \tag{7}$$

$$\Psi_t = \Psi_0 + \gamma t + v_t \quad (transition\ \ equation) \tag{8}$$

In the above equations y_t is a (3x1) vector of observables, X_t is a (4x1) vector of covariates, and Ψ_t is a (2x1) vector representing the underlying state (abilities). The stochastic component of the model is represented by w_t and v_t, which are i.i.d. multivariate normal processes, jointly independent, with zero mean and covariance matrices Ω and Σ respectively.

The approach taken here is Bayesian. We couple prior distributions over the unknown parameters with the likelihood function of the state space model to derive a posterior distribution for those parameters (and the unobserved state). The final goal is then to apply Markov Chain Monte Carlo (MCMC) methods to sample from that posterior. The method chosen comes from the Bayesian statistics literature and is due to Geweke and Tanizaki, although it is related to other papers such as Gordon, Salmon and Smith (1993), Carlin, Polson and Stoffer (1992) in the nonlinear setups, and to Carter and Kohn (1994) and others in linear state-space models.

The specific prior distributions were described in Section 4, and for the purpose of this appendix, it suffices to acknowledge that the joint prior can be written as a product of independent priors for the individual parameters.

$$p(\theta) = p(\sigma)p(\rho)p(\mu)p(\lambda)p(\Psi_0)p(\gamma)p(\Sigma)p(\Omega). \tag{9}$$

The next step is to choose the Gibbs sampling blocks and partition the vector of interest θ into N blocks, $\theta = (\theta_1, \ldots \theta_N)$.[13] There are 4 blocks in our estimation, $\{\sigma, \rho, \mu, \lambda\}$, $\{\Omega\}$,$\{\Psi_0, \gamma, \Sigma\}$, and a block for the unobserved state, $\{\{\Psi_t\}_{t=1}^{T}\}$, where T is the number of observations. The algorithm consists of sampling from each of the conditional posteriors for the four Gibbs sampling blocks.

[13] Good introductions to Gibbs sampling and Metropolis Hastings algorithms are Casella and George (1992) and Chib and Greenberg (1995).

B.1 Unobserved State

To sample from $p(\{\Psi_t\}_{t=1}^{T}|\theta, Y^T, X^T)$, we start by defining the vector $z^t = \{z_1, \ldots, z_t\}$ for any variable z_t, the densities $p(\Psi^t, Y^t|\theta, X^t)$, $p(\Psi^t|\theta, X^t)$, and $p(Y^t|\Psi_t, \theta, X_t)$, the joint density for Y^t and Ψ^t, the density for Ψ^t and the conditional density for Y^t given Ψ^t. All are conditional on the model parameters and exogenous variables. Instead of sampling jointly the entire vector Ψ^T, we will sample for $t = 1, \ldots, T$ from the distribution $p(\Psi_t|\Psi^{t-1}, \Psi^{t+1,*}, \theta, Y^T, X^T)$, where $\Psi^{t+1,*} = \{\Psi_{t+1}, \ldots, \Psi_T\}$. Rewrite the joint distribution for Ψ^t and Y^t as the product:

$$p(\Psi^t, Y^t|\theta, X^t) = p(\Psi^t|\theta, X^t)p(Y^t|\Psi^t, \theta, X^t). \tag{10}$$

Moreover, by the independence assumption between w_t and v_t these densities can be further simplified:

$$p(\Psi^t, Y^t|\theta, X^t) = p(\Psi^t|\Psi_0, \gamma, \Sigma)p(Y^t|\Psi^t, \sigma, \rho, \mu, \lambda, \Omega, X^t). \tag{11}$$

Since we assume that the error terms are i.i.d., we rewrite $p(\Psi^t|\Psi_0, \gamma, \Sigma)$ as:

$$p(\Psi^t|\Psi_0, \gamma, \Sigma) = p(\alpha_0) \prod_{t=1}^{T} p(\Psi_t|\Psi_0, \gamma, \Sigma), \tag{12}$$

and $p(Y^t|\Psi^t, \sigma, \rho, \mu, \lambda, \Omega, X^t)$ as:

$$p(Y^t|\Psi^t, \sigma, \rho, \mu, \lambda, \Omega, X^t) = \prod_{t=1}^{T} p(y_t|\Psi_t, \sigma, \rho, \mu, \lambda, \Omega, X_t). \tag{13}$$

An expression for the conditional density of Ψ^T given the data is given by Bayes' Theorem:

$$p(\Psi^T|Y^T, \theta, X^T) = \frac{p(\Psi^T, Y^T|\theta, X^T)}{\int p(\Psi^T, Y^T|\theta, X^T)d\Psi^T}. \tag{14}$$

To obtain the desired density $p(\Psi_t|\Psi^{t-1}, \Psi^{t+1,*}, Y^T, \theta, X^T)$, we begin by rewriting it:

$$p(\Psi_t|\Psi^{t-1}, \Psi^{t+1,*}, Y^T, \theta, X^T) = \frac{p(\Psi^T|Y^T, \theta, X^T)}{\int p(\Psi^T|Y^T, \theta, X^T)d\Psi_t}$$

$$= \frac{p(Y^T|\Psi^T, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi^T|\Psi_0, \gamma, \Sigma)}{\int p(Y^T|\Psi^T, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi^T|\Psi_0, \gamma, \Sigma)d\Psi_t}. \tag{15}$$

We can now substitute expressions (6) and (7) in both the numerator and the denominator, simplifying to:

$$p(\Psi_t|\Psi^{t-1}, \Psi^{t+1,*}, Y^T, \theta, X^T) = \frac{p(y_t|Y^{t-1}, \Psi_t, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi_t|\Psi_0, \gamma, \Sigma)}{\int p(y_t|Y^{t-1}, \Psi_t, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi_t|\Psi_0, \gamma, \Sigma)d\Psi_t}. \tag{16}$$

Once we have specified a candidate proposal density $q(\Psi_t^{(j,*)}|\Psi_t^{(j-1)}, \Sigma)$, we apply a Metropolis step for each time period $t = 1, \ldots, T$. We used univariate normals centered at the previously accepted value (random walk chain) with a variance proportional to Σ. Therefore, given a candidate $\Psi_t^{j,*}$, accept it with probability:

$$\phi_j = min\left(1, \frac{p(y_t|Y^{t-1}, \Psi_t^{(j,*)}, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi_t^{(j,*)}|\Psi_0, \gamma, \Sigma)}{p(y_t|Y^{t-1}, \Psi_t^{(j-1)}, \sigma, \rho, \mu, \lambda, \Omega, X^t)p(\Psi_t^{(j-1)}|\Psi_0, \gamma, \Sigma)} \right). \tag{17}$$

All remaining parameters in the previous expression are understood to be accepted values at iteration $j-1$.

B.2 Measurement Equation Parameters

Given values for the state variable $\{\Psi_t\}_{t=1}^T$, the joint distribution for the observables is a product of T normals with mean $f(X_t, \theta, \Psi_t)$ and covariance matrix Ω. Each of these normal densities is denoted by $p(y_t|\Psi_t, \sigma, \rho, \mu, \lambda, \Omega, X_t)$. Multiplying the observables' joint density by the priors for all parameters gives the posterior kernel:

$$p(\sigma, \rho, \mu, \lambda, \Omega|Y^T, \{\Psi_t\}_{t=1}^T, X^T) \propto p(\sigma)p(\rho)p(\mu)p(\lambda)p(\Omega) \prod_{t=1}^T p(y_t|\Psi_t, \sigma, \rho, \mu, \lambda, \Omega, X_t). \tag{18}$$

The sampling is done in two blocks. The first block is the 4×1 vector $(\sigma, \rho, \mu, \lambda)'$. The second block is the covariance matrix Ω. In both cases, sampling was done through a Metropolis Hastings step (equivalent to (17)) using independent univariate random walk chains for each of the parameters, *i.e.* the covariance matrix in the candidate proposal density is diagonal. Drawing candidates for Ω was done in a separate step to facilitate the gauging of the variance matrix in the candidate proposal density.

B.3 Transition Equation Parameters

Finally, we need to draw Ψ_0, γ and Σ. Given knowledge of the values for the unobserved state, we have a system of two equations in which the variance of the innovation term is the same across equations. The underlying state acts as an observable. Since the covariance matrix (Σ) is constrained to have equal diagonal elements, it is convenient to sample from this system with a Metropolis-Hastings step, with just one blocking, $\{\Psi_0, \gamma, \Sigma\}$.

References

[1] E.W. Anderson, L.P. Hansen, E.R. McGrattan, and T.J. Sargent. On the mechanics of forming and estimating dynamic linear economies. In Hans M. Amman, David A. Kendrick, and John Rust, editors, *Handbook of Computational Economics*. Elsevier, 1996.

[2] David Autor and Lawrence Katz. Changes in the wage structure and earnings inequality. In Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 3A, pages 1169–1214. Elsevier, Amsterdam, 1999.

[3] Ernst Berndt and Catherine White. Technology, prices, and the derived demand for energy. Unpublished Paper, University of British Columbia, 1978.

[4] William F. Blankenau and Beth F. Ingram. Welfare implications of factor taxation with rising wage inequality. *Macroeconomic Dynamics*, 6:408–28, 2000.

[5] David Card and John E. DiNardo. Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, 20(4):733–81, 2002.

[6] B.P. Carlin, N.G. Polson, and D.S. Stoffer. A Monte-Carlo approach to nonnormal and nonlinear state-space modeling. *JASA*, 87(418):493–500, 1992.

[7] J. Casella and E. George. Explaining the Gibbs sampler. *The American Statistician*, 46:167–174, 1992.

[8] F. Caselli and W.J. Coleman. The U.S. technology frontier. *American Economic Review*, 92(2):148–152, May 2002.

[9] S. Chib and E. Greenberg. Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49:327–335, 1995.

[10] Kim Clark and Richard Freeman. Time-series models of the elasticity of demand for labor in manufacturing. Unpublished Paper, Harvard University, 1977.

[11] R. Cole, Y. Cehn, J. Barquin-Stolleman, E. Dulberger, N. Helvacian, and J. Hodge. Quality-adjusted price indexes for computer processors and selected peripherals equipment. *Survey of Current Business*, January 1986.

[12] Thomas F. Cooley and Edward C. Prescott. *Frontiers of Business Cycle Research*. Princeton University Press, Princeton, NJ, 1995.

[13] P. Crifo-Tillet and E. Lehmann. Why will technical change not be permanently skill-biased? *Review of Economic Dynamics*, 7(1):157–180, January 2004.

[14] D. DeJong, B. Ingram, and C. Whiteman. A bayesian approach to dynamic macroeconomics. *Journal of Econometrics*, 15:311–320, May-June 2000a.

[15] D. DeJong, B. Ingram, and C. Whiteman. Keynes vs. Prescott and Solow: Identifying sources of business cycle fluctuations. *Journal of Applied Econometrics*, 98:203–223, 2000b.

[16] Edward F. Denison. *Accounting for Slower Economic Growth: The United States in the 1970s*. The Brookings Institution, Washington D.C., 1979.

[17] Enid Dennis and V. Kerry Smith. A neoclassical analysis of the demand for real cash balances by firms. *Journal of Political Economy*, 86:793–814, 1978.

[18] J. Fernandez-Villaverde and J. Rubio-Ramirez. Estimating nonlinear dynamic equilibrium economies: A likelihood approach. Working Paper, University of Pennsylvania, 2002.

[19] J. Geweke and H. Tanizaki. Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. Working Paper, Kobe University, 2000.

[20] N.J. Gordon, D.J. Salmon, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE-Proceedings-F*, 140:107–113, 1993.

[21] Jeremy Greenwood, Zvi Hercowitz, and Per Krusell. Long-run implications of investment-specific technological change. *The American Economic Review*, 87(3):342–62, June 1997.

[22] Daniel S. Hamermesh. *Labor Demand*. Princeton University Press, Princeton, NJ, 1993.

[23] Lutz Hendricks. Why does educational attainment differ across U.S. states? Working Paper, Iowa State University, February 2004.

[24] Peter Ireland. A method for taking models to the data. Working Paper, Boston College, 1999.

[25] Per Krusell, Lee E. Ohanian, José Víctor Ríos-Rull, and Giovanni L. Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–53, September 2000.

[26] Matthew J. Lindquist. Capital-skill complementarity and inequality over the business cycle. Working Papers in Economics, Stockholm University, 2002.

[27] Lee Ohanian, Giovanni L. Violante, Per Krusell, and José Víctor Ríos-Rull. Simulation-based estimation of a non-linear, latent factor aggregate production function. In Robert S. Mariano, Tel Schuelrmann, and Melvyn Weeks, editors, *Simulation-Based Inference in Econometrics: Methods and Applications*, pages 359–99. Cambridge University Press, Cambridge, U.K., 2000.

[28] Christopher Otrok. On measuring the welfare cost of business cycles. *Journal of Monetary Economics*, 47:61–92, February 2001.

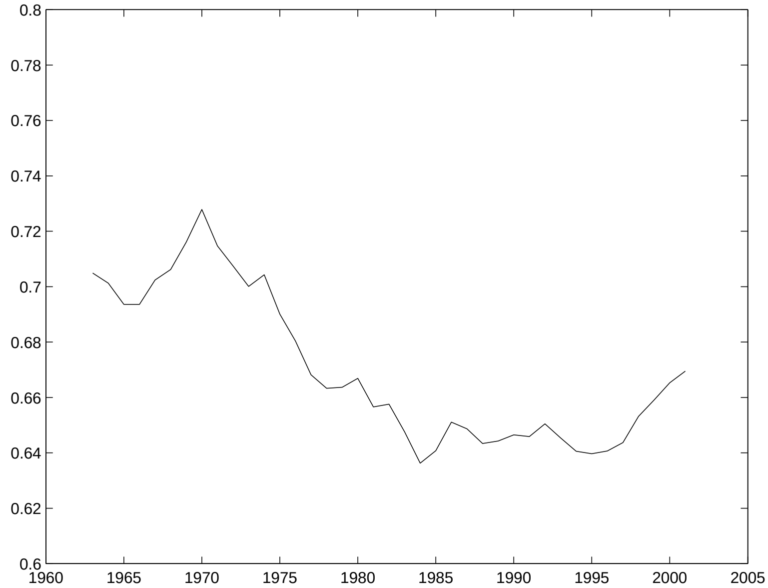

Figure 1: Labor Share

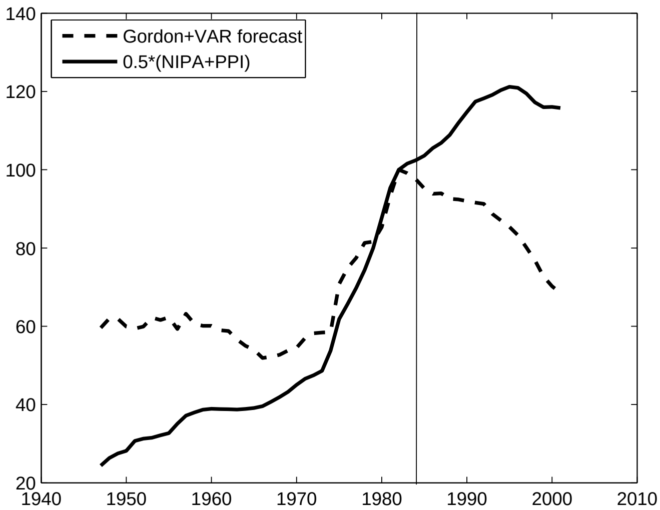

Figure 2: Gordon and GHK Price Levels

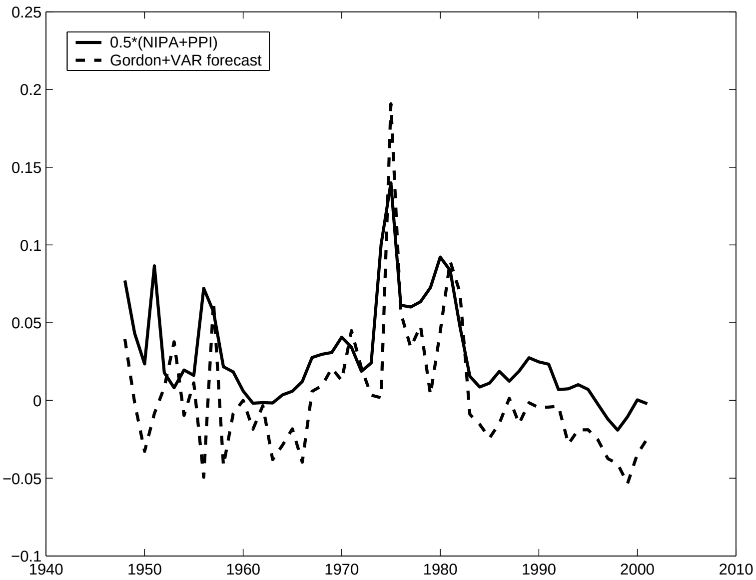

Figure 3: Growth Rates of Gordon and GHK Prices

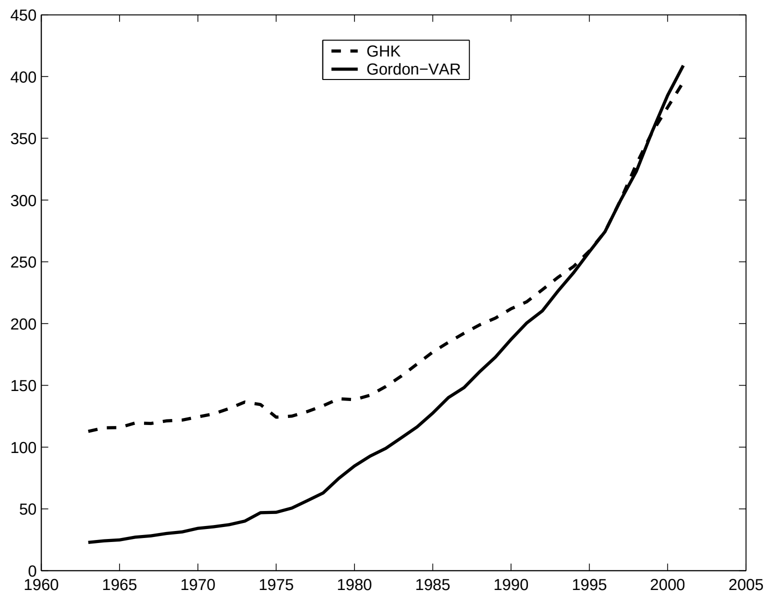

Figure 4: Stock of Equipment

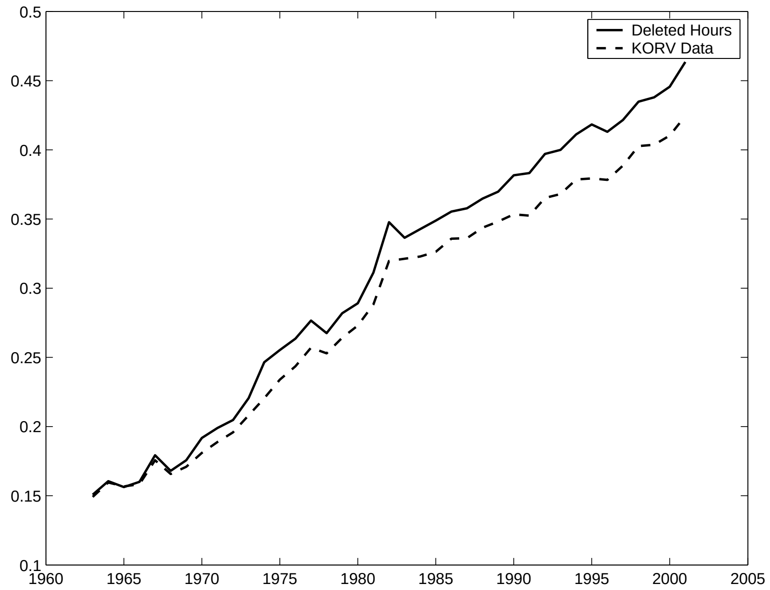

Figure 5: Labor Input Ratio

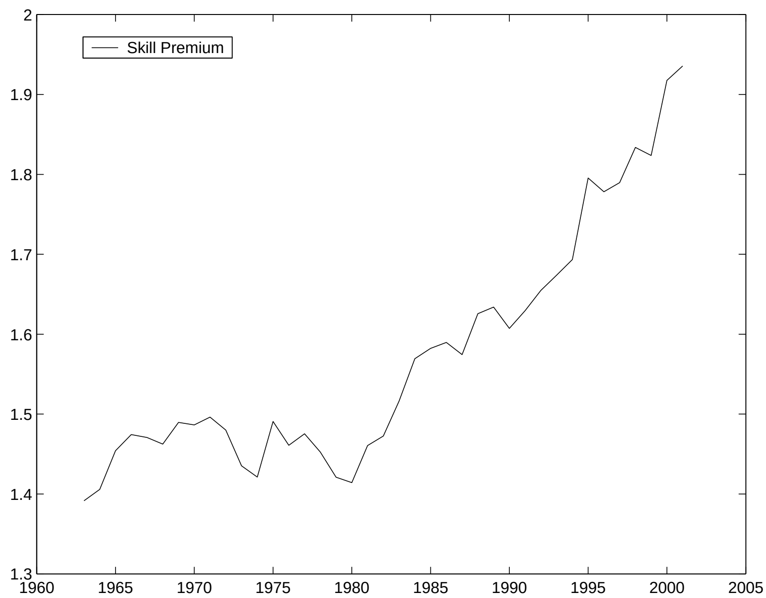

Figure 6: Skill Premium

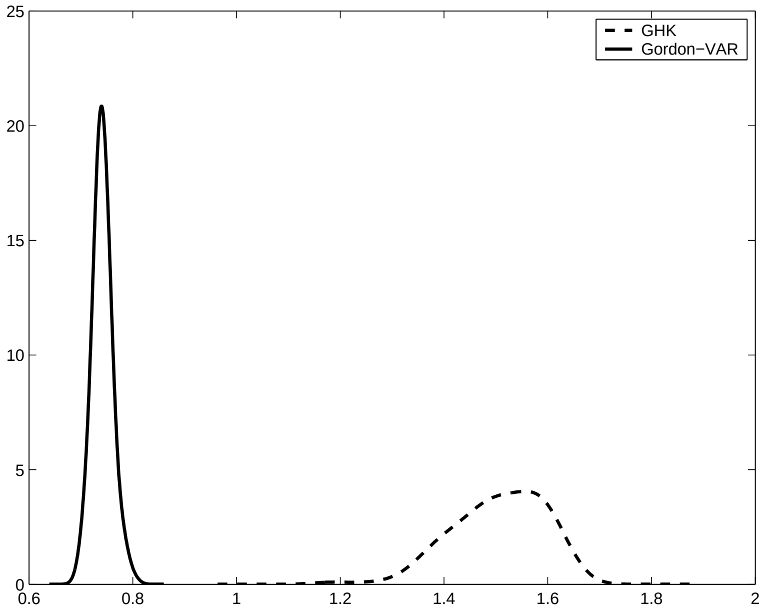

Figure 7: $\sigma - \rho$

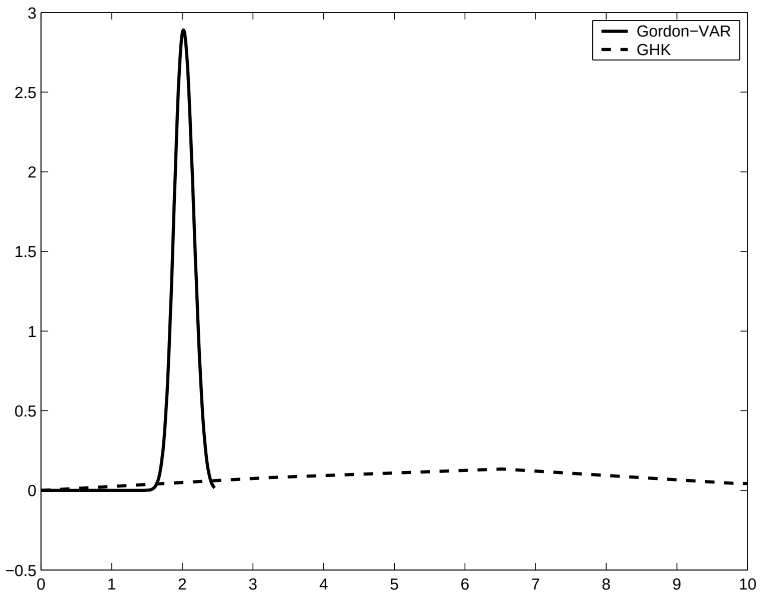

Figure 8: $1/(1-\sigma)$

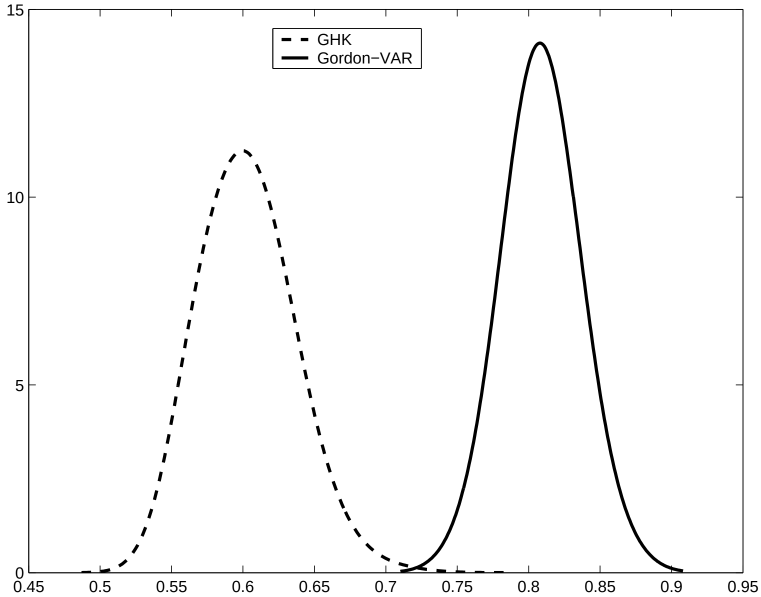

Figure 9: $1/(1-\rho)$

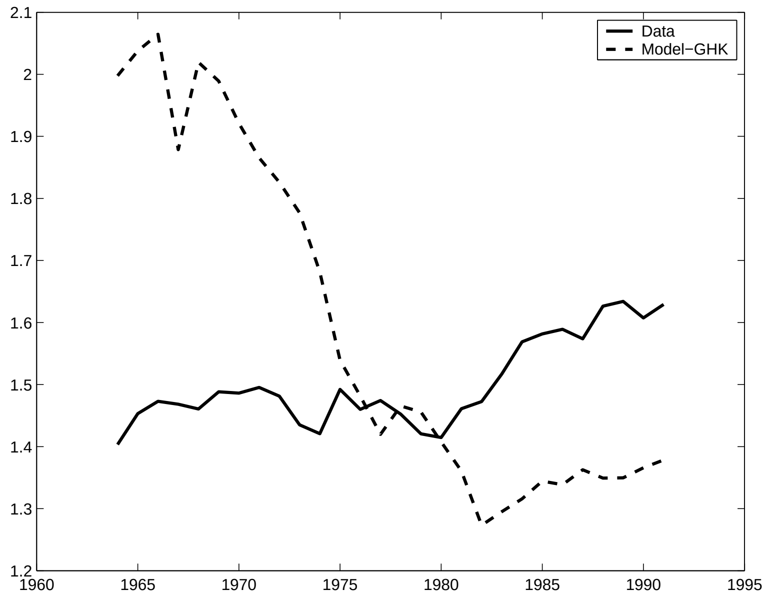

Figure 10: Skill Premium – GHK Data, Short Series

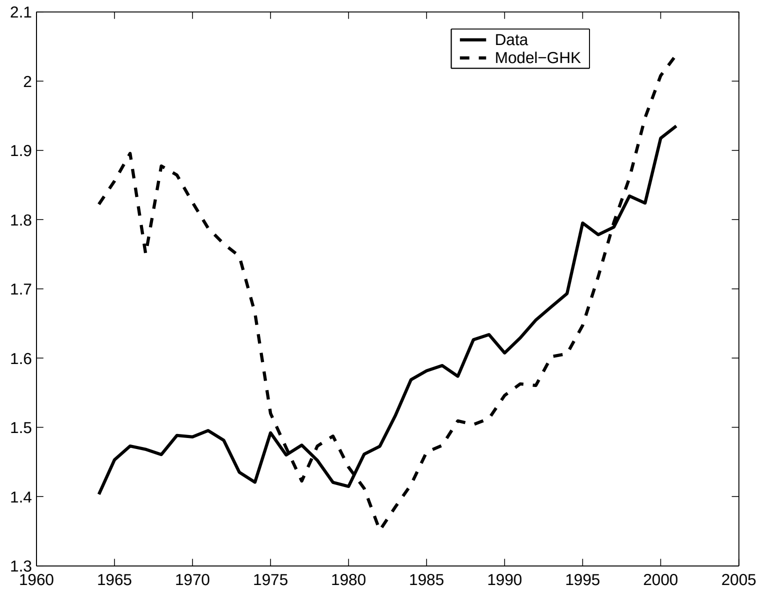

Figure 11: Skill Premium – GHK Data, Long Series

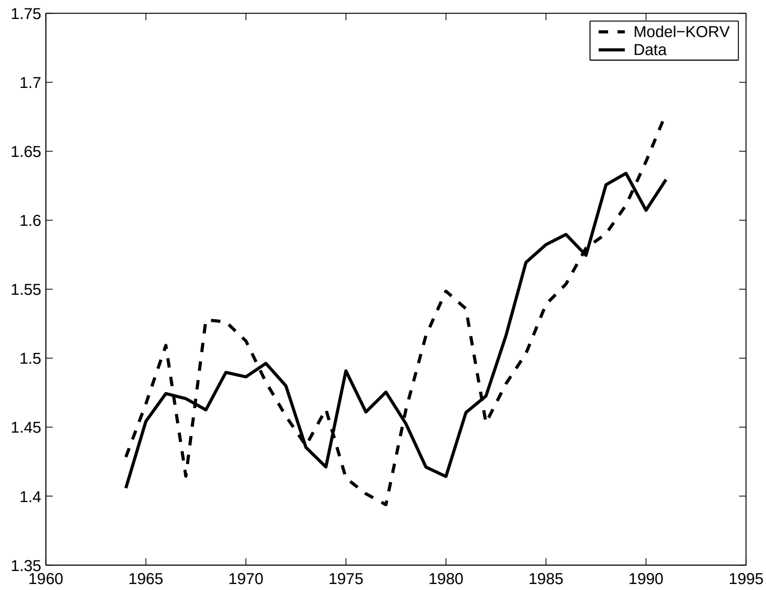

Figure 12: Skill Premium – KORV-Style Data, Short Series

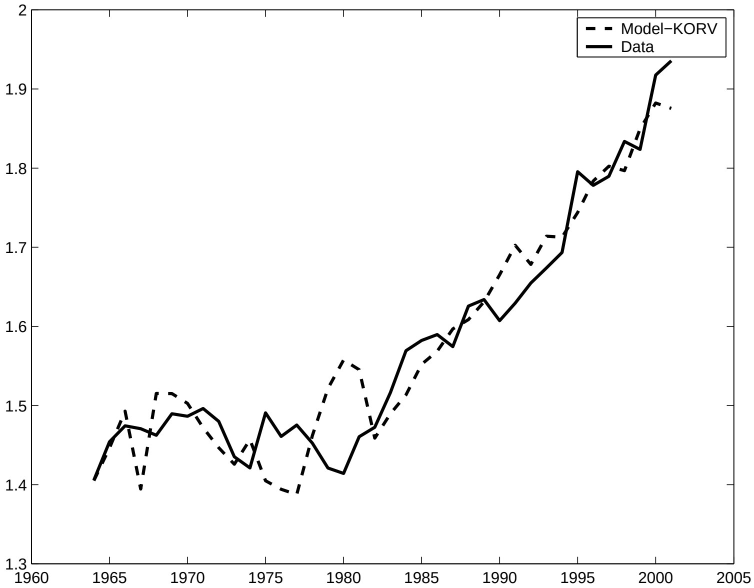

Figure 13: Skill Premium – KORV-Style Data, Long Series

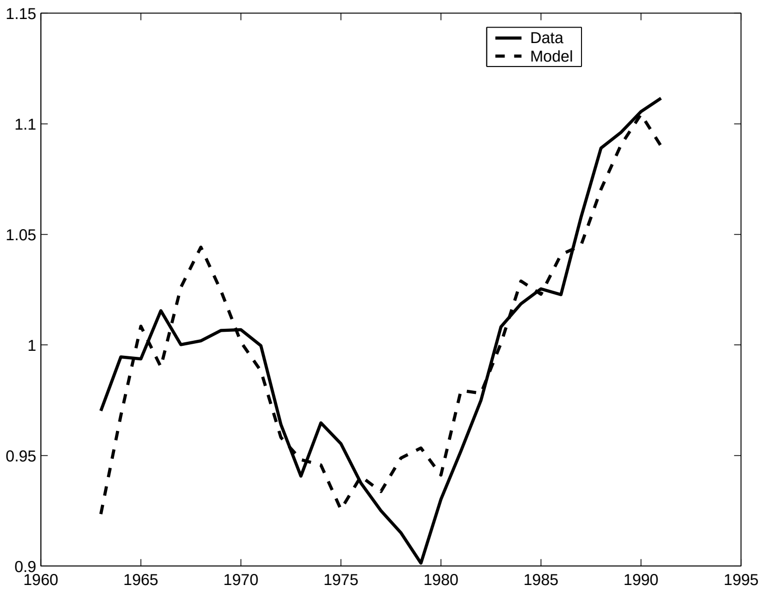

Figure 14: Skill Premium Using KORV's Own Data