Comments on "Neoclassical Models of Aggregate Economies" by Gary Hansen and Lee Ohanian

John H. Cochrane*

April 11 2015

1 Introduction

Gary and Lee's paper is really a review, like many of the papers in this volume. But rather than a long summary of who did what, it gives three specific and detailed examples. I will complain – that's my job – but my complaints are not really about this paper as much as suggestions for how the reviewed literature can do even better in the future.

2 Facts

This paper starts with an interesting fact, which I'll review here: There is lots of medium-term variation in GDP growth, in the postwar period as well as the well-known Great Depression and WWII boom. In Figure 1 I graphed their GDP per capita with some artfully drawn lines to draw your eye to the decade-long variations in growth, without bandpass filters.

Growth variations matter. 10 years of 1% vs. 2% growth adds up to 10% of GDP, which is one huge recession, and it lasts a lot longer.

This is a timely concern. Are we headed for a "great stagnation?" If so, do its cause and solutions lie in endlessly deficient "aggregate demand," per Larry Summers, in deficient TFP-raising innovation, per Bob Gordon, or in massive sand in the gears inhibiting the normal forces of recovery, as per Casey Mulligan? (And, I'll admit, John Cochrane.) Not just the prosperity of our population but whether the US can pay its debts, solve its entitlement problem, and avoid a chaotic end to pax Americana depend on growth.

*University of Chicago Booth School of Business, Hoover Institution, NBER.

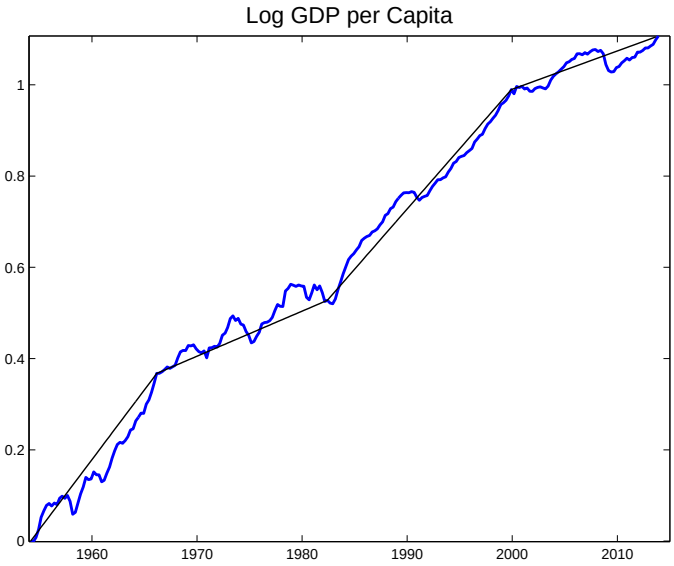

Figure 1: Log GDP per Capita. Source: Hansen and Ohanian

It's interesting though that our policy process seems to care so much about the high frequency fluctuations only. Maybe there is something to habit preferences!

So where do fluctuations in output come from?

Figure 2 illustrates the well known point that all long term growth comes from productivity. This graph is a simple accounting, but represents the conclusions of standard growth models as well.

That leads to the natural question, where does productivity come from, which is what the growth theory surveyed by Chad talks about. I'll return to the central question, where do shorter term fluctuations in productivity come from.

To look at the medium and higher frequencies, Figure 3 subtracts the average growth from each series. This is a simple high-pass filter.

The big growth spurt of the 1955-1965 era still corresponds primarily to productivity. But from 1970 on, lots of short and medium-term growth corresponds to variation in hours. The correlations in these medium-run fluctuations are the same as those in the traditional shorter-run fluctuations, which is a big point of the paper.

The data since 1970 in Figure 4 show this pattern more clearly. Medium and short-term fluctuations in output are now mostly due to hours.

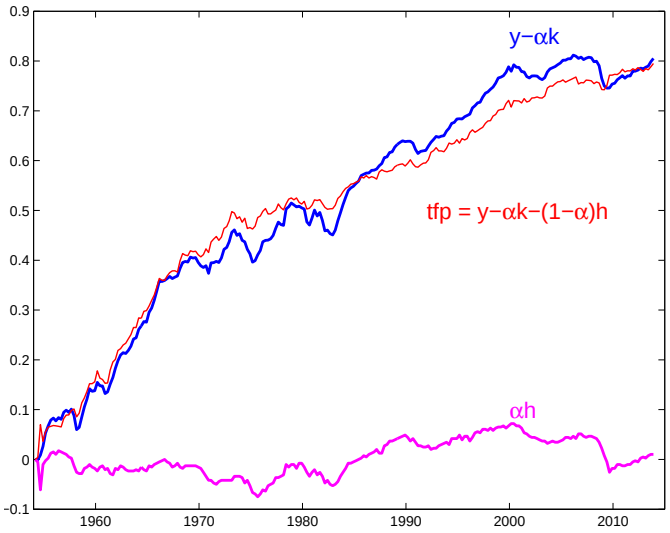

Figure 2: Output adjusted for capital, hours, and TFP

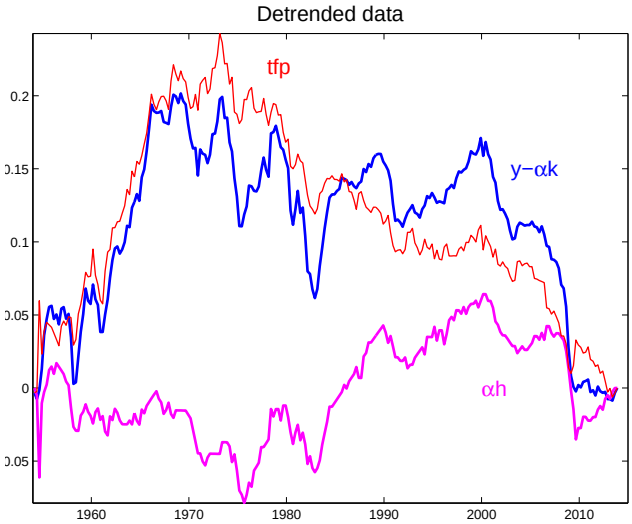

Figure 3: Detrended output, tfp and hours. Each series subtracts its average growth rate.

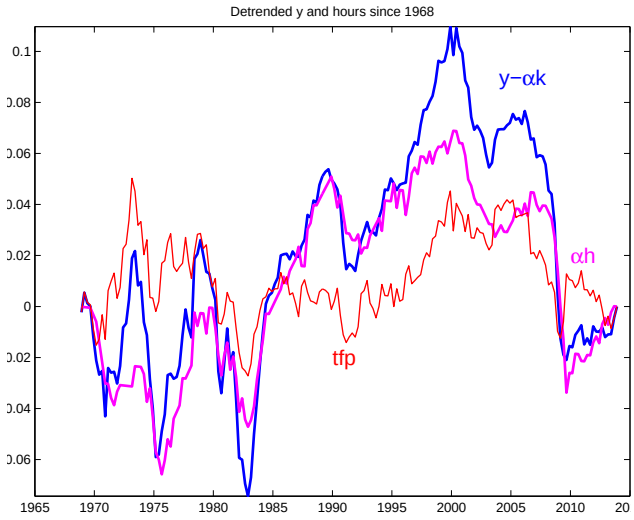

Figure 4: Detrended output, tfp and hours, shorter sample.

Again, this graph is accounting, not a model. Productivity can still be the root cause of fluctuations. But now any root cause must go through hours to produce fluctuations, at least with this simple $F(A, K, H)$ technology. That's what a "business cycle" *is*, and how a cycle differs from growth. The question is, how – what model produces this surely endogenous variation in hours.

3 This paper – data handling

This paper argues that "Neoclassical models" account for these medium and short-term fluctuations, including also the Great Depression slump and the WWII boom, based largely on exogenous technology shocks, with some twists to overcome the usual inadequate labor supply elasticity.

Now I will fulfill my job as discussant to whine – I mean to suggest directions for improvement.

Evaluating models with bandpass filtered data is common in this literature. But there are some dangers.

Economics does not separate by frequency. Think of the standard permanent income model.

$$c_t = rk_t + r\beta E_t \sum_{j=1}^{\infty} \beta^j y_{t+j};\ \Delta c_{t+1} = r\beta \left(E_{t+1} - E_t\right) \sum_{j=1}^{\infty} \beta^j y_{t+1+j}$$

Seasonal, nonseasonal, long run and short run movements in income all produce the same random walk consumption.

Filtering consumption growth would ruin the basic Hall prediction, that consumption growth is unpredictable from other variables.

(Credit where it's due: Lars Hansen pointed this out to me long ago after I wrote a paper using bandpass filters. I haven't touched one since!)

This is whining, not a fatal flaw. It is useful to filter both model simulations and data in the same way, and to find that models match data in some frequency ranges but not in others.

But do not expect that one can simply create a "long run" model and a "short run" model and then add the modeling elements that work in each domain. Expect actions in one frequency range to respond to ingredients useful in another.

Second, filtering out low frequencies obscures a beautiful fact. As Stock and Watson pointed out empirically, and as standard models predict theoretically, aggregates follow one common trend with stationary "great ratios." This is a much nicer, and inherently multivariate, way to separate trend and cycle.

This matters, because if a "cycle" means anything, it means that we can forecast higher growth in bad times and vice versa. The time series are not just correlated random walks. And forecasts using great ratios (possibly filtered – it's ok on the right hand side) are some of the best. (Technically these are error-correction representations.)

$$\Delta y_{t+1} = a + b(c_t - y_t) + \epsilon_{t+1}$$

Focusing on correlations of two-sided filtered data loses contact with predictability.

4 Too Many Models

So, it would be a lot better if both model and data handled trend and cycle in one breath. (Bob Hall made this comment on the first paper.)

And why not?

The great beauty of "neoclassical" models is that they extend the spectacularly successful growth model to fluctuations.

$$\max E \sum_{t=0}^{\infty} \beta^t u(C_t, H_t)$$
$$Y_t = C_t + I_t + G_t = F(K_t, H_t; A_t) \quad [-\phi(I_t, K_t)]$$
$$K_{t+1} = (1-\delta)K_t + I_t$$
$$\Delta \log(A_t) = \theta(L)\varepsilon_t$$

Why do we not do that? After extending this model, Start by verifying that we've preserved the growth implications, then go on to look at higher frequency measures of fit, and revel in the model's ability to capture both sets of facts.

I think that chopping up the frequencies is a vestigial habit. Keynesian ISLM models were developed before there *was* a growth theory, and they had no intellectual commonality with growth models when those came along. So naturally, ISLM modelers wanted to focus only on detrended data as if it were a separate phenomeon, imagine a separate economics of cycles and economics of growth, and wave hands about how in the long run we're all dead.

Kydland and Prescott filtered, and we tend to follow famous methodologies. But I think they wanted to connect to that Keynesian audience, to say "look, we can win a fight in your sandbox."

But Kydland and Prescott won. It's our sandbox now. We can do it right and emphasize the unity of neoclassical macroeconomics and growth theories, and make our project cumulative.

This paper exemplifies a deeper weakness of the current literature. It uses three utterly different models to address its three episodes! If insiders and outsiders held up wages in the great depression, surely those unions were still around in the 1950s. If equipment investment prices were important in the 1990s, surely they might have mattered in the 1930s or 1940s. If government-induced innovation was important in the propeller age, surely it was important in the space age.

Maybe not. Maybe each ingredient is only quantitatively important in its episode. But that's something that should be tried and evaluated quantitatively.

The promise of "neoclassical" economics is that there is *one* economic structure, common across frequencies, across shocks, and across episodes. We eschew saltwater one fact - one model economics. We emulate big hits like Friedman's permanent income model, which explain many disparate facts with one model. But empirically, "neoclassical" macroeco-

nomics is awfully noncumulative. Each paper does one variation on the stochastic growth model, nails one moment, and moves on.

Now, I don't want to encourage too much the opposite tendency. The empirical "DSGE" or "New Keynesian" literature, distinguishable only by the presence of nominal rigidities, and comprised of many (ex?) neoclassicists, does construct medium or large scale models, integrating tens or hundreds of economic epicycles into one big black box. Aside from the often rather ad-hoc micro-foundations, the real problem is the black box. The identification problem is too extreme – too many models can "fit the data," so we cannot distinguish models on that basis. Models must tell stories too. Economic models are quantitative parables, not black boxes. These models also have one ingredient per moment. They just put it all in the same paper.

5 Productivity shocks, and wedges

The models in this paper still rely on exogenous fluctuations in productivity, identified and measured as residuals, as the main driving variable for fluctuations. As long as "wedges" remain identified as residuals, they will suffer from some of the same doubts. This is an old complaint, but worth restating anyway.

One way to express the doubt: We have a parallel experience in monetary economics. Is $PY = MV$ or is $V = PY/M$? Monetarists thought that velocity was "stable," so money and velocity control nominal income. Now velocity is just whatever PY happens to be divided by whatever M happens to be.

Is the Solow residual any more "stable" or "exogenous" at business cycle frequencies than the Friedman residual? Is $Y = AK^{\alpha}H^{1-\alpha}$ or is $A = Y/\left(K^{\alpha}H^{1-\alpha}\right)$?

As Pete Klenow mentioned yesterday, there's a long literature suggesting that we work harder per reported hour, use machines more intensively, move in to higher TFP occupations, and so forth to produce this cyclical variation in Solow residuals.

We need to understand the economics of cyclical TFP variation, and measure it separately. This is my main take-away point today.

Gary and Lee see the need and give stories. For example, they write that the key WWII TFP increase by which they deny the usual multiplier, "likely" comes from technical innovation:

> "There is a considerable increase in TFP, which likely reflects a number of technological advances during the 1940s, including innovations directly or indirectly fostered by federal R & D expenditures. These include the development of modern airframes, radar, microwave technology, fertilizer, oxygen steel,

> synthetic rubber, nylon, sulfa drugs and chemotherapy, insecticides, and Teflon and related industrial coatings. ?'

That's not going to silence fiscal stimultators. The modern airframes ended up blown up in the skies of Germany and Japan, not raising productivity. Radar, rubber, nylon, sulfa drugs went to war, at non-market prices. These are not processes that raise GDP for given capital and labor. And this is just a story. We really need a separate, quantitative, measurement of tfp or other driving "wedges".

It's a more deeply worrying story. If this is true, to end the Great Depression, all the government had to do is to subsidize research that brought these products and raised TFP.

Conversely, even these items are not at all "exogenous." Why did industry and government rush to invent new products? Well, because of a massive government preference shift for more output, now. If the government had gone on a massive debt-financed military spending and R&D binge in the 1930s, output would have been a lot higher. Hmm. Sounds like a "multiplier" to me. (In the following discussion, Valerie Ramey brought up the important point that process innovation, bringing the assembly line to the production of aircraft and ships was a key part of the productivity increase.)

I've been pretty critical of "secular stagnation" models that posit a sharply negative nominal rate as a deus-ex-machina, without separately measuring its existence. Productivity shocks are for now in the same category, it-must-have-been-there-because-output-fell sorts of things. They should be separately measured from microeconomic investigation. (And, as I'll argue later, we should recognize all the things beyond inventions that go in to productivity, including policy distortions.)

6 Neoclassical models of everything?

The standard view of long slumps such as the great depression and our recent troubles is that something other than tfp caused the sharp drop – a financial panic, monetary policy, etc. The task of "neoclassical" "medium run" theory is to explain why some of those recessions turned in to slumps, and why some, like 1921, did not. Even Friedman and Schwartz took this view.

This paper takes a rather audacious view that no, they've got the whole enchilada.

> "models that generate fluctuations from temporarily inflexible prices and/or wages or that feature other transient impediments to mutually advantageous trades between agents may be of limited interest in understanding the bulk of U.S. Fluctuations"

This strikes me as a claim too far. Gary and Lee base their claim on the fact that output and hours comovement looks the same at their lower frequencies and at conventional "business cycle" frequencies. But not all series look the same in "growth slowdowns" as they do "recessions." Unemployment (AKA job search), financial variables, investment, etc. are all examples. There *is* a distinction between "Recessions " and "growth slowdowns."

Figure 5 adds unemployment to the graph to make this point. Unemployment practically defines conventional recessions. Perhaps our policy makers are more concerned with the pain of job search than they are about the level of GDP, justifying the focus on conventional recessions rather than growth slowdowns.

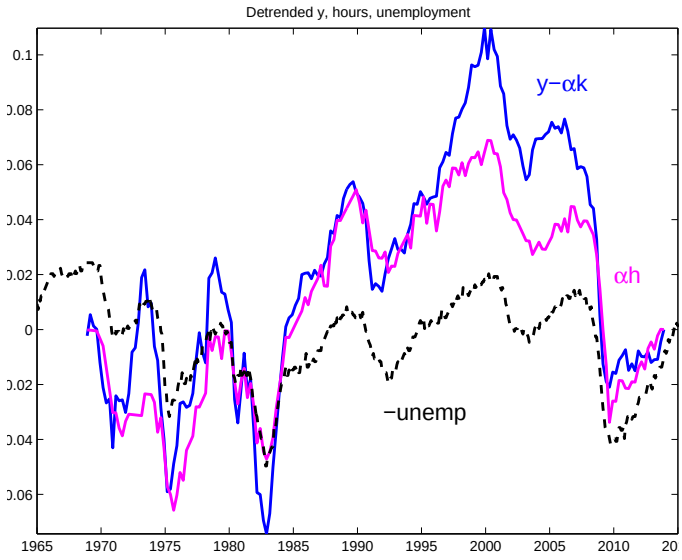

Figure 5: Detrended output, hours, productivity, and the inverse of unemployment

7 An alternative neoclassical agenda

As I think about our current malaise, then, this quote struck me

> "The failure of market economies to return to trend was the consequence of government policies that reduced the steady state allocation of time to market work. A sharply slow steady state level of market hours reduce the return to capital, which in turn depressed capital. "

I wanted to cheer – Yes!

Oh wait. Wrong quote. Lee and Gary were talking about the great depression there. Talking about now, they only offer the milquetoasty

> The model captures much of the Great Recession and the failure of the economy to recover.. partially because of a significant slowdown in the equipment-specific shock.

Let's get back to the first quote! Let's get back to Cole and Ohanian: Sure, money and bank runs can have the shock. But then Roosevelt's cartelizations, alphabet soup agencies, 70% marginal tax rate, unions, "war on capital" ad so forth kept the economy from recovering. WWII mostly represented and end to that war to fight a more important one.

Likewise, we had a crisis, having little to do with technology shocks. But then social program disincentives, taxes, regulation, Dodd-Frank, ACA, FDA, EEOC, NLRB, EPA, DOJ, our alphabet soup are impeding recovery.

These distortions reduce productivity and introduce other wedges. Such sand in the gears seems to me a much more compelling possibility than a reduction in the pace of exogenous scientific discoveries. Ideas need to be produced, refined, embodied in new products and processes and businesses, diffuse, then suffer a welter of policy-induced inefficiencies. Maybe the latter doesn't matter in the low-frequency growth picture. But it sure can matter for a decade.

Yes, this proposal abandons the Pareto-efficient cycles tradition in favor of an older tradition that microeconomic distortions have macroeconomic effects. But the innovation process already abandons Pareto-efficiency in growth theory, and studying tax and regulatory distortions is as "neoclassical" as you can get. And it's not the end of the world to recover a macroeconomics in which policy can help, though mostly by getting out of the way.

For now it's a story. Does it add up quantitatively to tfp and other macroeconomic wedges? Let's find out! Let's separately, quantitatively, measure the microeconomic distortions. That will also address the longstanding productivity as deus-ex-machina wedge complaint.

8 What's a Neoclassical model?

A last complaint about the paper. What's a "neoclassical" model, anyway? The introduction define one as having

> .. (1) well-defined optimization problems for consumers and firms who have rational expectations, (2) prices and wages are perfectly flexible, (3) .. an explicit definition of general equilibrium, and the equilibrium is unique, (4) there are (typically) complete markets, (5) technologies are convex production sets, and (6), the main drivers of shifts in economic activity are shocks to technologies and government policies that impact the incentives and opportunities to produce, consume, invest, and trade.

and the conclusion adds

> ... standard preferences and technologies, ...well-functioning markets, ... shifts in economic activity are driven primarily by technological change and shifts in government policies.

But this paper proceeds to construct a insider-outsider model of cartelized industries holding wages up in manufacturing. (The WWII and postwar models are similarly innovative) They just threw out the first three items!

- ~~Prices and wages are perfectly flexible~~
- ~~Well-functioning markets~~
- ~~Standard preferences and technologies~~

For example, in the the household objective: "The value of being an insider, ... is

$$V_t(n) = \max_{\bar{w}_t, \bar{n}_t} \left\{ min[1, \frac{\bar{n}}{n}]([\bar{w}_t - w_{ft}) + \pi \left(\frac{Q_{t+1}}{Q_t} \right) V_{t+1}(\pi \bar{n})] \right\}''$$

"Standard?"

The "neoclassical" literature has likewise shown great creativity in preferences and technology. There is no such thing as a "standard." If anything, we are guilty of the opposite, each paper has one fact and a different set of preferences and technologies.

- ~~Rational expectations~~

I'm not sure I want to kick Tom Sargent's work on learning out of the club. Or say, Martin Schneider and Monika Piazzesi's work thinking about how expectations adjusted to news in the 1980s. Sure, we don't do simple adaptive expectations, but the tent is pretty big.

- ~~Technologies are convex production sets,~~

Fixed costs, adjustment costs, and irreversibiltiies all seem like ingredients we don't want to throw out of the club. Kydland and Prescott had a pretty "nonstandard" time to build!

- ~~Complete markets~~

Uninsured shocks and heterogeneity, and precautionary savings (Krussell and Smith) seem like common and acceptable parts of the "neoclassical" program. Surely if we want to think about unemployment statistics, meaning job search, we need models in which search rather than Walrasian auctioneers clear markets.

- ~~Shifts in economic activity are driven primarily by technological change and shifts in government policies.~~

Leaving out what, preferences shocks? Even the saltiest new-Keynesian model attributes fluctuations to government policies and avoids preference shifts if possible.

Let me suggest an alternative definition.

- Well-posed.

Neoclassical models are well-posed. Behavior comes from optimization, subject to budget constraints, with well described equilibrium. I.e. they're not ISLM models.

- Intertemporal, explicit predictions for time series data.

Neoclassical models are intertemporal, they can can produce explicit predictions for time series, which one checks against actual data. In a way that, say, the three-period asymmetric information models of corporate finance cannot, and must remain suggestive parables. (That's not a criticism, really. I know such modelers would like to produce explicit dynamic models, but it's technically infeasible at the moment.)

We're doing well on these two items.

- Micro founded, or at least reasonable preference, technology, market structure

The preference, technology and market structure specifications are supposed to be well founded and documented in microeconomic analysis.

I gave us a C- here. There is a lot of making it up as we go along.

- Uniform. One model, many phenomena, frequencies

Neoclassical models are supposed to be uniform and integrative. Many phenomena and frequencies fall out of a few model assumptions. As above, we get another C-, we need to try harder.

- No *nominal* price stickiness?

By and large neoclassical models don't have nominal price stickiness. However, if monetary policy does not have a lot of shocks, you can explain the bulk of output variations in models that abstract from monetary policy. If you want to describe real effects of monetary policy, however, a super-neutral model isn't going to do the trick. Well micro-founded nominal stickinesses certainly fall in the neoclassical philosophy, so I'm reluctant to bar Mike Woodford or John Taylor at the door.

Perhaps

- Philosophy: Work harder.

Perhaps the right distinction is philosophical. When simple $Y = Af(K, L)$, $\sum \beta^t u(c_t)$ models don't work, neoclassical modelers try harder, adding detail of preferences technology and market structure, rather than quickly adding ad-hoc frictions to straw-man simple models.

- Result: Not One humor "Aggregate Demand."

Perhaps the right distinction is by result rather than ingredient. 18th century medicine had three humors, and diagnosed diseases by their imbalances. Paleo-Keynesian economics has one humor, "aggregate demand," and one undifferentiated remedy, "stimulus." Any kind of stimulus can cure any source of deficient aggregate demand. Neoclassical models do not try to remedy a (say) tax distortion by more spending. When they suggest policy interventions at all, those are targeted at the specific disease.

- Macro is Micro.

Or perhaps the right distinction is simply a victory of microeconomics over macroeconomics. A lot of Harbinger triangles *are* the Okun gap.

In sum, it seems to me there is really a very big tent, one club, and that is just the club of serious economics. A lot of policy analysis occurs outside that club. But within the club we can just have a productive debate contrasting models with data to figure out which ingredients are most important to match the data without too many labels.