There is by now a huge literature on the increase in the college premium and other dimensions of inequality in the United States and many other Western nations (see Acemoglu and Autor, 2011, for an overview of this literature). As I discuss below, the focal explanation in this literature is that technological changes of the last four decades have increased the demand for skills and have pushed up premia to different kinds of skills, college education among them (though other factors including globalization and changes in labor market institutions have also contributed to these trends).

The paper by Jaimovich, Rebelo, Wong and Zhang tackles an important topic and develops a relatively underresearched line of inquiry within this broad literature. The main idea is that a major contributor to the increase in the demand for skills has been “trading up” (the authors’ term) by households to higher-quality products as they have become richer. Higher-quality products are argued to be more intensive in skilled labor. As a result, this process has naturally brought a higher demand for skills as a byproduct of economic growth.

This is an important idea, and one I sympathize with a lot. The paper also has a noteworthy original contribution in providing compelling motivating evidence. It estimates product quality from a variety of sources, links these to establishment-level demand for skills from the microdata of the Occupational Employment Statistics (OES) dataset of the Bureau of Labor Statistics, and verifies that higher-quality products are more skill intensive. This empirical work alone is worth more than the price of admission.

But the paper does not fully deliver on this very promising research agenda. The reason why it fails to do that is interesting and instructive. It is because it follows a methodology I will call quantitative Friedmanite modeling. This approach combines Friedman’s (1953) famous methodological dictum that realism of assumptions does not matter (sometimes called the “as if” hypothesis) with an emphasis on developing quantitative evaluations of macro models calibrated with some plausible choice of parameters. This methodology has some obvious shortcomings at the best of times (replicating some moments in the data based on microeconomic parameter choices does not
reject alternative hypotheses, nor does it provide clear support for the proposed mechanisms. In the current context, however, it is even more problematic because it pushes the authors away from engaging with the key economic forces their own hypothesis and evidence bring to the table.

I am going to use the rest of this short essay to briefly discuss the paper’s contribution, what I mean by the quantitative Friedmanite modeling, why this methodology fails to shed light on the issues at hand, and what different approaches might have been more fruitful.

1 The Contribution

Most analyses of the demand for skills in labor and macro literatures do not distinguish between quality differences across goods and assume that the demand for the different types of goods (with different skill intensities) have the same income elasticity, so that changes in the income level (or even income distribution) of households does not directly change the demand for skills.

The current paper starts by relaxing these assumptions. First, it allows for quality differences between different goods. For example, Dunkin’ Donuts coffee and hand-crafted artisanal coffee are not perfect substitutes, and the latter has a higher income elasticity so that as consumers become richer, their demand will shift away from Dunkin’ Donuts towards the specialized coffee shops (see, e.g., Shaked and Sutton, 1982). Second, it posits that the production functions for lower and higher quality goods are different, and producing higher-quality goods requires more skills. In the context of the coffee example, Dunkin’ Donuts is assumed to need less skilled workers than the specialized coffee shop.

Put these two pieces together and you can conclude that as (some) households in the economy become richer, they will change the composition of their demand and in particular they will start demanding higher-quality vintages — more of the hand-crafted coffee and less of the Dunkin’ Donuts fare. All else equal, this will increase the demand for skills and put upward pressure on the skill premium.

This is a very plausible mechanism, but of course the devil is in the details. How important is it? Why is it that higher-quality vintages require more skilled workers?

Moreover, the mechanism, though underresearched and probably underappreciated, has featured in other papers already. Two lines of work are particularly noteworthy. The first is a series of papers by Marco Leonardi (2003, 2015), which argue for the same type of differential income elasticity and investigate how much of the increase in the demand for skills this mechanism can account for. In Leonardi’s work, for example, higher-skilled workers demand more of the goods produced by other higher-skilled workers, while low-skill workers consume less skill-intensive goods. The second line comprises a number of theoretical investigations of the inequality implications of non-homothetic preferences. These include: Zweimuller (2000), Foellmi and Zweimuller (2006), and
Foellmi, Wuergler and Zweimuller (2014), as well as related work by Matsuyama (2002).

Despite this earlier literature, it is fair to say that there was no fully convincing evidence that the production of higher-quality products is more skill intensive. The most important contribution of the current paper is to establish this fact.

The authors first estimate the quality of different products using data from Yelp! and the Nielsen home scan data. For example, from Yelp!, they use the information on the price categories (low, middle, high and very high) provided by users. They then match the Yelp! establishments to the OES establishments, where they can measure various proxies for skill intensity. They focus on measures of skills based on wages (rather than education), so skilled workers are those with relatively high wages (either on average or compared to other workers in the same industry). This raises a nontrivial concern — rent-sharing will show up as greater demand for skills.

Setting aside this concern (which should be investigated more systematically), the results are fairly consistent and very interesting: establishments selling higher-quality products employ a greater share of high-skill workers. Another result, which is even more telling (as I will argue below), is that using data on occupations from OES, they demonstrate that a smaller fraction of workers in high-quality establishments perform routine tasks. So on the basis of this correlation, it appears that high-quality production necessitates more non-routine tasks to be completed.¹

These results are an important contribution to the literature and are telling about the relationship between quality upgrading and the nature of work. Notably, this relationship probably goes beyond the demand for skills; as the authors’ own findings illustrate, there is an important change in the types of tasks that are being performed (for example, the shift from routine to non-routine tasks).

Even if it needs to be probed and investigated more, and especially separated from rent-sharing effects, this intriguing empirical pattern suggests the need to understand how the production process of higher-quality products differs from that of lower-quality products, what dimensions of skills are more important for producing high-quality, and how this interacts with technology. The types of models and approaches that would permit such an investigation are available in the economic growth and labor economics literatures. But the current paper takes another path, not uncommon in modern macro, but as I will argue ultimately unsatisfactory for analyzing the issues at hand.

2 The Quantitative Friedmanite Modeling Methodology

Milton Friedman’s famous (1953) essay applied a (simplified) version of Karl Popper’s approach to economics and proposed a simple economic methodology. It can be summarized by stating that

¹See Acemoglu and Autor (2011) and Autor and Dorn (2013) on routine and non-routine tasks and their relationship to the organization of production.
a theory should be judged solely on the basis of its “predictions”, with no regards to whether its assumptions are accurate or descriptively realistic. Friedman boldly stated

“Viewed as a body of substantive hypotheses, theory is to be judged by its predictive power for the class of phenomena which it is intended to ‘explain’. Only factual evidence can show whether it is ‘right’ or ‘wrong’ or, better tentatively ‘accepted’ as valid or ‘rejected’ (p. 149).

Friedman had a harsh assessment of efforts to judge a theory by its descriptive realism, calling such attempts

“... fundamentally wrong and productive of much mischief. Far from providing an easy means of sifting valid from invalid hypotheses, it only confuses the issue, promotes misunderstanding about the significance of empirical evidence for economic theory, produces a misdirection of much intellectual effort devoted to the development of positive economics, and impedes the attainment of consensus on tentative hypotheses in positive economics” (p. 153).

When describing the behavior of an expert billiards player, we can make much progress by modeling their behavior as if they are undertaking the full mathematical calculations of the trajectory of the ball once it is hit by the cue. We can, Friedman argued, make progress in economics by similarly imposing various as if assumptions, even if these are patently false. All that matters is that our theory, building on these assumptions, provides valid predictions for the problem at hand. Applying this reasoning, for example, he concluded that much of the work on monopolistic competition was misguided because it was motivated by the desire to provide a better approximation to markets in which firms were neither pure monopolists nor one of many perfectly competitive businesses (p. 153). So far as Friedman’s methodology was concerned perfect competition was just fine because its predictions about the effect of changes in demand were not falsified, and hence the theory could be “tentatively accepted”.

Although many philosophers and economists raised myriad valid concerns about Friedman’s economic methodology, it has had a curiously enduring influence on economic research. In the words of the philosopher Daniel Hausman, “Methodologits have had a few kind words for Milton Friedman’s [methodology], yet its influence persists” (1992, p. 183). Hausman instead advocated “looking under the hood”, that is, studying how different components of the theory generate the relevant predictions and whether they are realistic and receive empirical support. As such, he identified the most fundamental weakness of Friedman’s approach: reliability of empirical predictions have to be evaluated recognizing that any model, and particularly so in social sciences, is useful primarily
as an aid to better understand the problem being studied. The wrong mechanisms, even if they lead to the right empirical predictions within some context, are worse than useless because they propagate the wrong kind of understanding. Looking under the hood and striving for some sort of congruence between features of what we include in our models and the reality we are studying are some of the ways in which we can attempt to achieve this. Looking under the hood does not mean shying away from simplifying assumptions. But it does require that we are clear about the core mechanisms for the phenomena we are studying, and we judiciously use simplifying assumptions for abstracting from other aspects, while striving to represent and systematically investigate these core mechanisms.

Friedman’s methodology has influenced modern macroeconomics too. The ideal espoused by Friedman was to subject economic hypotheses derived from various as if assumptions (and preferably in Friedman’s assessment starting with perfect competition and similar settings where the market worked well) to a battery of rigorous empirical tests. One branch of modern macro has combined Friedman’s methodology with quantitative evaluation/calibration. In standard statistical theory, a null hypothesis is compared to an alternative. In its modern versions, there is an effort to undertake “causal inference”, for example, using randomized control trials, regression discontinuity type strategies or instrumental variables estimation (e.g., Angrist and Pischke, 2011). What I am calling the quantitative Friedmanite modeling methodology instead starts with similar as if assumptions and then compares the magnitudes implied by the model under some parametric assumptions (sometimes chosen on the basis of standard parameter choices in the literature and sometimes on the basis of estimates from micro data) to some selected moments in the macro data.

This methodology can be a powerful approach for evaluating whether a particular mechanism can be “quantitatively important”. One of its most celebrated applications was to argue that productivity-shock driven business cycles could account for the magnitude of fluctuations in the US data (Kydland in Prescott, 1982). But this methodology may sometimes discourage efforts to look “under the hood”. This, I argue, is what has held back the current paper.

3 Jaimovich et al.’s Model of Higher-Quality Production

Jaimovich et al. make two major simplifying assumptions. First, they assume (in their main model) that households are homogeneous and thus the demand for quality is uniform across the entire economy. Second, they model the production of higher quality goods with a small variation on the canonical approach to the demand for skills in labor economics, which builds on Katz and Murphy’s (1992) specification derived from a constant elasticity of substitution aggregate production function with factor-augmenting technologies (see also Tinbergen, 1975, and Goldin and Katz, 2007). Let me start with the latter choice.
The aggregate production function the authors specify takes the form

\[ Y = A[\alpha(S^H H)^\rho + q^{-\gamma \rho}(1 - \alpha)(S^L L)^\rho]^{1/\rho}. \] (1)

Here, \( L \) denotes the supply of unskilled labor and \( H \) is the supply of skilled labor, while \( 1/(1 - \rho) \) is the elasticity of substitution between skilled and unskilled labor (and is taken to be greater than 1), \( \alpha \) is a distribution parameter designating the importance of skilled labor relative to unskilled labor, and \( S^H \) represents any technology that increases the relative (physical) productivity of skilled labor (which is equivalent to generic skill-biased technological change under the assumption that \( \rho > 0 \)). I have also added a symmetric term \( S^L \) to their specification for later discussion. Crucially, all technological change takes a factor-augmenting form as in the canonical approach. The new term is \( q^{-\gamma \rho} \), and captures the effects of quality. Higher-quality, corresponding to higher \( q \), directly reduces the productivity of lower skilled workers. (The model is equivalent to Katz and Murphy’s formulation when \( \gamma = 0 \)).

The authors then choose some parameter values motivated by the previous literature and their own descriptive work and derive the quantitative implications of their model. They conclude that when the “trading up” mechanism is included, the implications are much more plausible. Comparing their model to the baseline without \( q \) (or with \( \gamma = 0 \)) where \( S^H \) would need to increase by about 5.5% annually to account for the rise of the college premium, they report that with the trading up mechanism an annual increase of only 1.05% in \( S^H \) is necessary to account for the data.

Jaimovich et al.’s production function, (1), is a strange one. Higher quality directly makes low-skill workers less productive (recall that \( \gamma > 0 \)). Implicitly drawing on the quantitative Friedmanite methodology, the authors do not defend the realism of this production function. The predictions (or the quantitative implications) are derived as if there is such an aggregate production function, since this matches their own empirical work that higher-quality goods are more skill intensive. The lack of descriptive realism is not viewed as a road block.

Yet this assumption is problematic, and its lack of realism is a telltale sign of these problems. To understand these issues, let us first review some of the recent developments in the labor economics literature on the demand for skills. (With the full admission that this is my own take of these developments, partly based on my own work).

4 Problems with the Canonical Approach to the Demand for Skills

Acemoglu and Autor (2011, 2012) point out three problems with the canonical approach on which Jaimovich et al. build. First, the empirical fit of the Katz and Murphy’s (1992) approach deteriorates considerably after their sample ends. Second, in contrast to the prediction of a model in which the demand for more skilled activities is growing, the US data paint a picture in which firms are
expanding employment more in low-skill occupations than in higher-skill occupations. In fact, there is a notable pattern of employment polarization where middle-skill occupations are disappearing and being replaced mostly by lower-skill occupations (see also Acemoglu, 1999, and Autor and Dorn, 2013).

Third and most importantly, in the standard model, skill-biased technological change increases the demand for skills and skill premia, but unless there is technological regress and new technologies reduce the productivity of some types of workers, the model cannot generate declines in real wages. Mathematically, in equation (1) above, any combination of increases in $S^H$ and $S^L$ will always raise the real wage of low-skill workers (see Acemoglu, 2002). To generate a decline in the real wages of low-skill workers it is not sufficient to have skill-biased technological change — we need a decline in $S^L$, meaning technological regress. Secular deteriorations in technology are implausible to say the least. But in the data, the real wages of low-skill workers have declined precipitously since the late 1970s, especially when we focus on men.

Acemoglu and Autor (2011, 2012) interpret these as fundamental failures of the canonical approach and propose an alternative based on tasks. In this approach, production requires the performance of a range of tasks. Technology and factor prices determine the allocation of tasks to factors. Technology is no longer just factor-augmenting. Technological changes that reduce the range of tasks allocated to a factor can lead to a decline in the real wage of that factor — even if these technological changes have nothing to do with technological regress. These papers, as well as Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018, 2019a,b), propose models of automation where new technologies embedded in machines, such as computerized control or robots, enable the substitution of capital for tasks previously performed by labor, especially low-skilled labor (see also Zeira, 1998). Such automation will increase the demand for skills, but more importantly, may also reduce the real wages of low-skill workers. The evidence in Acemoglu and Restrepo (2019a), for example, shows that the introduction of industrial robots is associated with significant wage declines for workers with less than college education.

Let me give a brief overview of how this would work, drawing on the model from Acemoglu and Restrepo (2019b), which assumes there is a single type of labor (introducing workers with different levels of skills is straightforward).\(^2\) Suppose that the unique final good in the economy, $Y$, is produced by combining a set of tasks, with measure normalized 1, with production function given by

$$Y = \left( \int_0^1 Y(z)^{\frac{1}{\sigma}} dz \right)^{\frac{\sigma}{\sigma-1}},$$  

(2)

where $Y(z)$ denotes the output of task $z$ for $z \in [0,1]$ and $\sigma \geq 0$ is the elasticity of substitution

\(^2\)Models with different types of labor and automation are considered in Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018, 2020).
Tasks can be produced using capital or labor according to the production function
\[ Y(z) = \begin{cases} A^L \gamma^L(z) l(z) + A^K \gamma^K(z) k(z) & \text{if } z \in [0, I] \\ A^L \gamma^L(z) l(z) & \text{for all } z. \end{cases} \]
Tasks \( z \leq I \) are (technologically) automated and can be produced with capital, while those \( z > I \) are not automated and can only be produced with labor. In addition, \( l(z) \) and \( k(z) \) denote the total labor and capital allocated to producing task \( z \). The framework also allows for standard factor-augmenting technology terms, \( A^L \) and \( A^K \). The terms \( \gamma^L(z) \) and \( \gamma^K(z) \) represent the productivity of labor and capital in different tasks. Let us assume that \( \gamma^L(z) / \gamma^K(z) \) is increasing in \( z \), so that labor has a comparative advantage in higher-indexed tasks. In this framework, an increase in \( I \) corresponds to automation, expanding the set of tasks that can be produced with capital. Under the assumption that capital is cheap so that firms are happy to produce (technologically) automated tasks with capital, the equilibrium of this model can be equivalently represented as the equilibrium of an economy with an aggregate production function — but in this instance derived from the microstructure of the model at the task level. In particular, this derived aggregate production function takes the form
\[
Y = \left( \left( \int_0^I \gamma^K(z)^{\sigma-1} dz \right)^{\frac{1}{\sigma}} (A^K)^{\frac{\sigma-1}{\sigma}} + \left( \int_I^1 \gamma^L(z)^{\sigma-1} dz \right)^{\frac{1}{\sigma}} (A^L)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{3}
\]
Holding the level of automation, \( I \), constant, the results from this framework are identical to those from the canonical approach (the elasticity of substitution between capital and labor is now given by the elasticity of substitution between tasks, \( \sigma \)). But critically, automation (an increase in \( I \)) changes things. Most notably, an increase in \( I \) allocates tasks from labor to capital, and always reduces the term in front of labor \( \left( \left( \int_I^1 \gamma^L(z)^{\sigma-1} dz \right)^{1/\sigma} \right) \) and increases the term in front of capital \( \left( \left( \int_0^I \gamma^K(z)^{\sigma-1} dz \right)^{1/\sigma} \right) \). Because of this reallocation of tasks from labor to capital, automation may reduce the value of the marginal product of labor and real wages. This happens despite the fact that there is no technological regress (Acemoglu and Restrepo, 2018, 2019b).

Here we come back full circle to the quantitative Friedmanite methodology. Suppose the task-based approach described in the previous paragraph is on target; it is indeed the case that new technologies enable automation and as a result reduce real wages (even though they increase productivity). The Friedmanite methodology might try to capture the same phenomenon by imposing a reduced-form aggregate production function (like the one in (1)), and then forcing this production function and its implications on the data, it would conclude that this is happening because \( A^L \) is decreasing (technological regress). If the “prediction” we care about is matching the decline in real wages, this fix works. But it would lack complete descriptive realism for at least two reasons. First, the idea that there is actual technological regress, though imposed in this approach, would make
no empirical sense. Second, the Friedmanite approach would also eschew any engagement with the key economic mechanisms at work, in this instance the reallocation of tasks across factors (which is in fact responsible for the phenomenon we are trying to understand, the decline in real wages).

The situation is actually worse than this, because once we approach the problem at the right level (in this instance, at the level of the allocation of tasks to factors), we understand that automation does not always reduce real wages. As Acemoglu and Restrepo (2018, 2019b) show, the impact of automation on real wages depends on the balance between a displacement effect (what I have emphasized so far) and a productivity effect (resulting from the fact that the substitution of cheaper capital for labor increases effective productivity and thus may increase the demand for labor). This implies that the effects of different waves of automation could be quite different. One wave of automation may have a smaller productivity effect, reducing real wages. Yet another wave with a more substantive productivity effect may raise real wages. In consequence, the Friedmanite approach would be forced to maintain that in the case of the first wave there is significant technological regress, but not so in the second wave. In short, it would have to get into lots of twists and turns to try to “get the right predictions”, and at the root of these problems is exactly its insistence of not calibrating its assumptions to the micro structure of the problem being studied.

The parallel of the Friedmanite solution to this problem and Jaimovich et al.’s modeling approach are evident. Instead of assuming that there is a decline in $A^L$ (or in $S^L$ in (1)), the authors introduce another parameter, $q$, which does the same and effectively reduces the productivity of low-skill workers. They are not deterred by their assumptions not matching the micro structure. In particular, despite their own very interesting empirical finding that quality upgrading is associated with a change in the task structure of establishments, they prefer the reduced-form modeling approach that ignores what is happening at the task level.

On the basis of this, it is not far-fetched to conjecture that their modeling would be subject to similar problems. For example, it may well be that different waves of quality upgrading are associated with different ways of reorganizing tasks and differently-sized productivity effects. Hence directly assuming that the production of higher-quality goods reduces the productivity of low-skill workers may fit the facts for one wave and be a terrible approximation for another.

None of this increases one’s confidence in the quantitative exercise the authors perform. Even if the parameter values they use for this may be justified, the quantitative estimates they generate still depend heavily on the specific model imposed on the data (in the form of their equation (1)), so unless we have confidence in this model, we cannot place much stock in its quantitative findings.

This reasoning underpins my conclusion that the descriptive work they present is intriguing and

---

3 It also depends on the simultaneous creation of new tasks which I am suppressing here because this is not central to the methodological issue at hand.

4 However, regardless of the elasticity of substitution, automation always reduces the labor share in value added (see Acemoglu and Restrepo, 2018, 2019b).
in fact supports the idea that there are systematic differences in the skill intensity of the production process of low- and high-quality goods, but their model does not further enlighten the mechanisms at work, and the quantitative exercise does not generate numbers we can really trust.

5 An Alternative Approach

On the basis of what I have discussed so far, my preferred approach to this problem should be evident. This approach would start by specifying how the set of tasks that need to be performed for the production of a high-quality product is different from the set of tasks for a low-quality product. For example, perhaps customers paying more for a hand-crafted artisanal coffee also want the barista to talk to them about movies or politics, so the set of tasks associated with the production of such a cup of coffee is different than the production of a cup of Dunkin’ Donuts coffee. It is this change in the set of tasks that then necessitates the establishment to have a different composition of skills.

The data the authors use can start giving us some clues about this. For example, in addition to the change in the composition of routine and non-routine occupations in the aggregate, the authors could look at how this differs depending on the nature of the product the establishment is supplying. Is it more pronounced for establishments that are more service-intensive? For those that directly interact with customers more? For those that are more technology-intensive? For products that are more customized?

Another interesting question concerns whether greater skill intensity of the production of higher-quality products is a consequence of prevailing factor prices. For example, if college graduates earned twice as much, would specialized coffee shops still hire them as baristas? Put differently, the attributes of higher-quality products themselves may be endogenous, and when skill premia are sufficiently high, the attributes that require these skills may be cut back.

One could also investigate whether the composition of occupations changes significantly and whether the occupations are added as goods become higher quality (Acemoglu and Restrepo, 2019b, show that the diversity of occupations in an industry is associated with the introduction of new tasks).

6 Modeling Demand for Quality

The modeling of the demand for quality is secondary for the approach of the paper, and this is the reason the authors start with a representative agent model, and then consider heterogeneity only in the Appendix. However, the secondary nature of this aspect of the model is itself a consequence of another *as if* assumption — this time eschewing the competition between products of different qualities.
It is natural to presume, once again from the micro structure of the problem, that different quality variants of the same good are going to be much closer substitute than two distinct products. If so, the location of different variants on the quality ladder will affect the market power of producers. This was the starting point of Shaked and Sutton’s classic (1982) paper I mentioned above, showing how the distribution of income determines both the quality levels of products and markups. The linkage between quality distributions and markups emerges through a related but distinct channel in other growth and industrial organization models, for example, Aghion, Harris, Howitt and Vickers (2001) and Acemoglu and Akcigit (2012).

These micro interactions may matter as well when it comes to the demand for skills, since a high level of markup for an establishment reduces the demand for the type of labor it employs, and via this channel may impact the skill premium. This too is an important and interesting area for future research.

7 Conclusion

In sum, this paper is on an important and exciting topic and starts by documenting a novel and fascinating fact — establishments producing higher-quality products use more skilled workers. The paper is most likely correct that quality-upgrading (what the authors call “trading up”) contributes to the demand for skills. But how important is this effect? And even more critically, what are the micro mechanisms via which it operates?

The modeling approach the paper takes, which is to assume that quality enters as a shifter of an otherwise canonical constant elasticity of substitution production function for a unique final good, does not ultimately help us appreciate these mechanisms any better, and partly as a result of this, is not a good basis for answering questions related to the quantitative importance of this channel.

All the same, this paper has taken an important step in drawing our attention to the role of quality upgrading and its labor market implications. We have every reason to expect that others will follow and will build on the interesting facts that this paper has already started documenting.

References


Katz, Lawrence F., and Kevin M. Murphy (1992) “Changes in Relative Wages, 1963-


