Federal Reserve Bank of St. Louis

REVIEW

JULY/AUGUST 2011 VOLUME 93, NUMBER 4

A Message from the Editor

Measuring Inflation: The Core Is Rotten
James Bullard

Challenges for Monetary Policy in the European Monetary Union
Axel A. Weber

Decomposing the Education Wage Gap: Everything but the Kitchen Sink
Julie L. Hotchkiss and Menbere Shiferaw

"Frictions in Financial and Labor Markets": A Summary of the 35th Annual Economic Policy Conference
Rodolfo Manuelli and Adrian Peralta-Alva

Readership Survey

REVIEW

Director of Research
Christopher J. Waller

Policy Adviser
Cletus C. Coughlin

Deputy Director of Research
David. C. Wheelock

Review *Editor-in-Chief*
William T. Gavin

Research Economists
Richard G. Anderson
David Andolfatto
Alejandro Badel
Subhayu Bandyopadhyay
Maria E. Canon
Silvio Contessi
Riccardo DiCecio
Thomas A. Garrett
Carlos Garriga
Rubén Hernández-Murillo
Luciana Juvenal
Natalia A. Kolesnikova
Michael W. McCracken
Christopher J. Neely
Michael T. Owyang
Adrian Peralta-Alva
B. Ravikumar
Juan M. Sánchez
Rajdeep Sengupta
Daniel L. Thornton
Yi Wen
Christian M. Zimmermann

Managing Editor
George E. Fortier

Editors
Judith A. Ahlers
Lydia H. Johnson

Graphic Designer
Donna M. Stiller

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

A Message from the Editor

223
Measuring Inflation: The Core Is Rotten
James Bullard

235
Challenges for Monetary Policy in the European Monetary Union
Axel A. Weber

243
Decomposing the Education Wage Gap: Everything but the Kitchen Sink
Julie L. Hotchkiss and Menbere Shiferaw

273
"Frictions in Financial and Labor Markets": A Summary of the 35th Annual Economic Policy Conference
Rodolfo Manuelli and Adrian Peralta-Alva

Readership Survey

Review is published six times per year by the Research Division of the Federal Reserve Bank of St. Louis and may be accessed through our website: **research.stlouisfed.org/publications/review**. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* also are available to our readers on this website. These data and programs are also available through Inter-university Consortium for Political and Social Research (ICPSR) via their FTP site: www.icpsr.umich.edu/pra/index.html. Or contact the ICPSR at P.O. Box 1248, Ann Arbor, MI 48106-1248; 734-647-5000; **netmail@icpsr.umich.edu**.

Single-copy subscriptions are available free of charge. Send requests to: Federal Reserve Bank of St. Louis, Public Affairs Department, P.O. Box 442, St. Louis, MO 63166-0442, or call (314) 444-8809.

General data can be obtained through FRED (Federal Reserve Economic Data), a database providing U.S. economic and financial data and regional data for the Eighth Federal Reserve District. You may access FRED through our website: **research.stlouisfed.org/fred**.

Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Please send a copy of any reprinted, published, or displayed materials to George Fortier, Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442; **george.e.fortier@stls.frb.org**. Please note: Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis. Please contact the Research Division at the above address to request permission.

© 2011, Federal Reserve Bank of St. Louis.

ISSN 0014-9187

A Message from the Editor

Dear Readers,

The *Review* publishes research written by St. Louis Fed economists, along with the work of our guest lecturers, conference participants, and visiting scholars.

Our recent editorial policy had been to also publish the research of economists affiliated with the Federal Reserve System, but in a supplementary series—*Review II*—with a separate submission and referee system and its own editorial board.

We continue to invite economists affiliated with the Federal Reserve System to submit research for publication in the *Review*. However, their work will no longer be published in a separate series, but will be incorporated into the pages of each issue as it is accepted. In fact, this issue of the *Review* contains an article from Julie Hotchkiss and Menbere Shiferaw from the Federal Reserve Bank of Atlanta.

Please see our website for more information:

http://research.stlouisfed.org/publications/review.

As always, we thank you for your readership.

William T. Gavin
Review Editor-in-Chief

Measuring Inflation: The Core Is Rotten

James Bullard

An earlier version of this article was delivered as a speech to the Money Marketeers of New York University, New York, New York, May 18, 2011.*

Federal Reserve Bank of St. Louis *Review*, July/August 2011, *93*(4), pp. 223-33.

Controlling overall inflation is a goal of monetary policy. Measures of overall, or headline, inflation attempt to include changes in the prices paid for a wide variety of goods—that is, what households actually have to pay for their daily purchases. This is a sensible notion of precisely what the central bank can and should control over the medium term.

Many discussions of monetary policy, even within the central banking community, discuss movements of subsets of prices instead of the overall or headline measure of price changes. The most famous subset is the "core"—all prices except those relating to food or energy. *Core inflation* is the measured rate of increase of these prices.[1] Control of core inflation is not the goal of monetary policy, although it sometimes seems to be, given the amount of emphasis put on this concept in the United States.

Many of the old arguments in favor of a focus on core inflation have become rotten over the years. It is time to drop the emphasis on core inflation as a meaningful way to interpret the inflation process in the United States.

One immediate benefit of dropping the emphasis on core inflation would be to reconnect the Federal Reserve with households and businesses who know price changes when they see them. With trips to the gas station and the grocery store being some of the most frequent shopping experiences for many Americans, it is hardly helpful for Fed credibility to appear to exclude all those prices from consideration in the formation of monetary policy.

There are several key arguments that are commonly used to favor a focus on core inflation in monetary policy discussions.[2] I will argue that all of them are essentially misguided. Because of this, the best the central bank can do is to focus on *headline* measures of inflation. The headline measures were designed to be the best measures of inflation available—the Fed should respect that construction and accept the policy problem it poses. Many other central banks have solidified their position on this question by adopting explicit, numerical inflation targets for headline inflation, thus keeping faith with their citizens

[1] Figures 1 and 2 show core and headline inflation for the consumer price index (CPI) and the personal consumption expenditures (PCE) chain price index.

[2] See, for instance, Mishkin (2007).

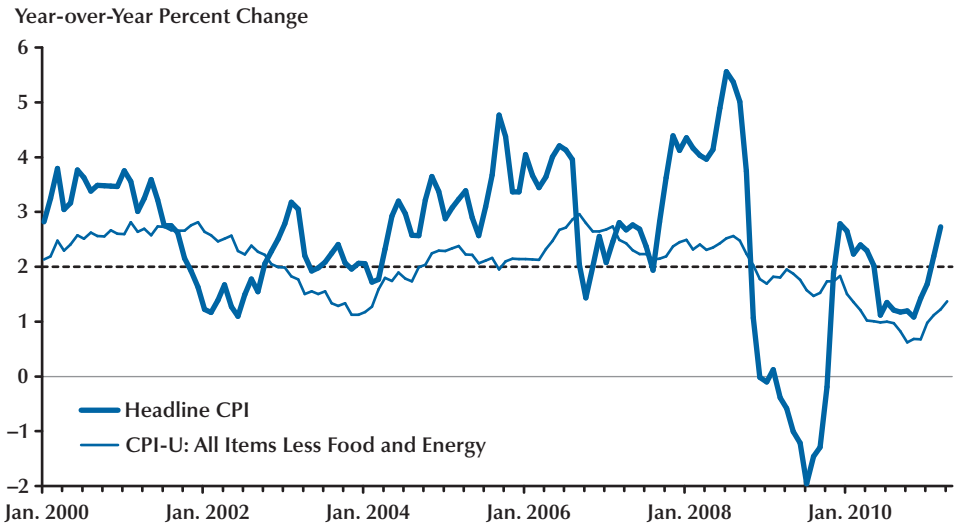
James Bullard is president and CEO of the Federal Reserve Bank of St. Louis. Any views expressed are his own and do not necessarily reflect the views of other Federal Open Market Committee members. The author benefited from review and comments by Cletus Coughlin, Riccardo DiCecio, William Gavin, Michael McCracken, Andrea Tambalotti, Daniel Thornton, Christopher Waller, and Jonathan Wright.

*The speech to the Money Marketeers of New York University (Bullard, 2011) can be accessed at http://research.stlouisfed.org/econ/bullard/pdf/Measuring_Inflation_May_18_2011_FINAL.pdf.

© 2011, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

Figure 1

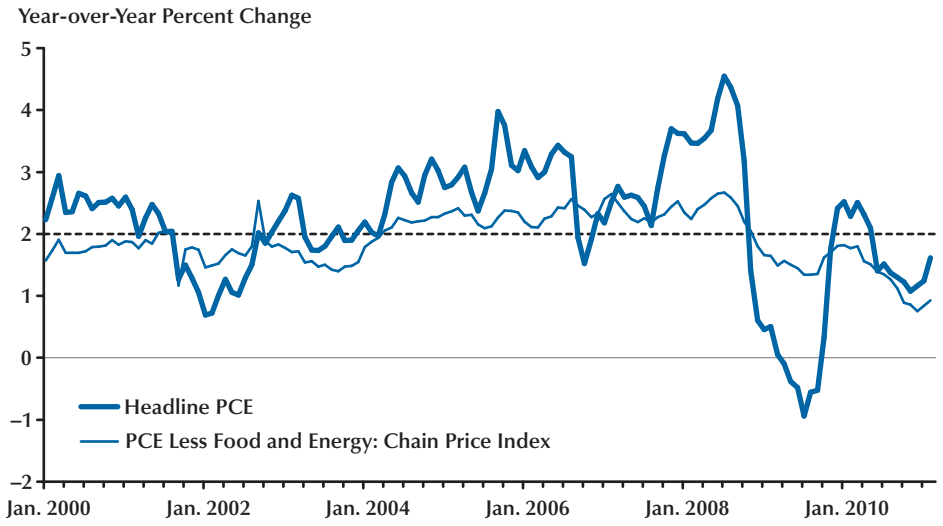
CPI Inflation Measures: Headline and Core



SOURCE: Bureau of Labor Statistics. Last observation April 2011.

Figure 2

PCE Chain Price Index Inflation Measures: Headline and Core



SOURCE: Bureau of Economic Analysis. Last observation April 2011.

that they will work to keep headline inflation low and stable. The Fed should do the same.

THE "VOLATILITY" ARGUMENT

I will start with an easy one, the argument that headline inflation is more volatile than core inflation and that, therefore, if monetary policy reacts systematically to headline inflation the economy itself would become more volatile. This could also be termed the "all hell would break loose" argument.

Yes, it is true that headline inflation tends to be more volatile than subset inflation measures that exclude or downweight the most volatile components. However, I do not think this says anything about how policy should or should not react to movements in headline inflation. Any policy response can of course be adjusted appropriately to take into consideration that the price index contains a certain level of volatility. In other words, the policy response can be optimized given the inflation index being targeted. Some monetary policy simulations that I have seen in this area simply take an existing policy rule that has been designed for core inflation and use the same rule with headline inflation—resulting in increased volatility in goal variables. That type of experiment is just saying that an inappropriate policy rule will produce less-than-satisfactory results, which is hardly surprising.

One might very legitimately turn the headline volatility question on its head. With core inflation as the preferred index for monetary policy analysis, the policymaker will tend to react to relatively small movements in measured core inflation. In that case, arguably, any policy response has to be larger—possibly substantially larger—when even small changes in measured core inflation are observed in order to execute the optimal policy. This may be ill-advised to the extent that small movements in core inflation are, in fact, simply noise.

Recent experience offers something to ponder in this regard. While many think that the recent financial crisis provides an illustration of the merits of the focus on core inflation, I do not see it that way at all. During the second half of 2008 and into 2009, headline inflation measured from one year earlier fell dramatically and in fact moved into negative territory. This was a signal—one among many, to be sure—that a dramatic shock was impacting the U.S. economy. Inflation was not immune to this shock. The Federal Open Market Committee (FOMC) reacted appropriately with an aggressive easing of monetary policy. Yet the movements in core inflation during this chilling period were far more muted and sent much less of a signal that action was required.

There is also the question of the wisdom of an intermediate target strategy with respect to inflation. Since headline inflation is the goal for monetary policy, the introduction of the core inflation concept as an intermediate target introduces some slippage between the variable the Committee is reacting to and the ultimate value of the goal variable. The intermediate target strategy works as follows: The Committee makes a policy tool adjustment (such as the policy interest rate), which is designed to target core inflation, which subsequently impacts headline inflation. It is not clear that this intermediate target strategy actually maximizes policy performance with respect to the overall price index. For that, much depends on the statistical properties of the relationship between core inflation and headline inflation, and that relationship tends to change over time.

And finally on the topic of the volatility of headline inflation, the headline index can be smoothed in any number of other ways that stop short of ignoring a wide class of important prices in the economy. One simple way is to consider headline inflation measured from one year earlier, but there are many others. To the extent that the volatility of headline inflation is a problem, there are better methods of addressing that than to simply dismiss troublesome prices.

THE "CORE PREDICTS HEADLINE" ARGUMENT

One popular argument for focusing on core inflation is that core inflation is a good and con-

sistent predictor of future headline inflation. I think this is wrongheaded, as well as wrong. Let's begin with the "wrongheaded" part. The idea that core predicts future headline is often based on univariate models of the inflation process—that is, ones that try to predict future headline inflation using only a single variable or a simple function of a single variable.[3] I do not think this is a good metric for understanding whether core or headline is the right inflation measure on which to gauge monetary policy decisions, regardless of whether it holds up in the data or not. Presumably, we would want to use a fully specified model to try to predict headline inflation, the goal variable with respect to prices, in the United States. The full model would include measures of expected inflation, developments in the real economy, the stance of monetary policy (including the implicit inflation target), and other variables to help to predict future headline inflation outcomes. One could throw all of these variables out in favor of a single variable—core inflation—when trying to predict future headline inflation, but presumably then one would have a misspecified model of the inflation process in which a simple function of core inflation is acting as a proxy for all the variables that are supposed to be important for predicting future headline inflation. In this misspecified model, a simple function of core inflation may or may not have been a good predictor of future headline inflation over a particular time period, but so what? I do not think this really tells us anything about whether it is a good idea to key policy off of core inflation or not.

There is a version of this argument that might make more sense. That version works like this. Suppose we have a full model of the inflation process, one that includes expected inflation, measures of real activity, and measures of the stance of monetary policy. We then add core inflation as a variable to this model and assess the marginal predictive value of core inflation given all other variables. If the marginal value of adding core inflation in this context is positive, one might then have a claim that core inflation contains some "special" information over and above information coming from the rest of the economy concerning the future course of inflation.[4] I have not seen convincing evidence of this type.

But let's go ahead and consider the merely "wrong" part of this argument. Let's examine whether today's core inflation is a good predictor of future headline inflation in the context of a univariate forecasting model. I do not think the evidence is very clear on this question. A number of choices have to be made to even proceed in this area: the horizon over which to forecast—let's say, one year; the function of core inflation to use, such as a distributed lag, and the length of that lag; the data sample over which to test the hypothesis; and the measures of core and headline inflation to use. One recent research paper in this area investigates personal consumption expenditures (PCE) inflation and standard core PCE inflation (that is, PCE inflation less food and energy components) alongside other types of inflation measures over U.S. data from 1982 to 2005.[5] In that paper, standard core PCE inflation performs relatively poorly as a predictor of future headline PCE inflation in most of the models discussed.[6] Alternative measures of inflation do better, such as the trimmed mean measure produced by the Federal Reserve Bank of Dallas for PCE, which is shown in Figure 3.[7] Analysis like this demonstrates that the idea that standard core inflation forecasts future headline inflation is far from an obvious conclusion given the U.S. data.[8]

Many intuitive discussions on the issue of core versus headline inflation contain the idea that noise should somehow be stripped out of the headline inflation measure in order to leave only the signal component. The energy price component of the headline price index often (depending

[3] For a bivariate approach, see Kiley (2008).

[4] Rich and Steindel (2007) do include simple measures of slack in conjunction with various measures of core inflation in their statistical model. They find that no one measure of core inflation consistently outperforms others in out-of-sample tests.

[5] See Smith (2010).

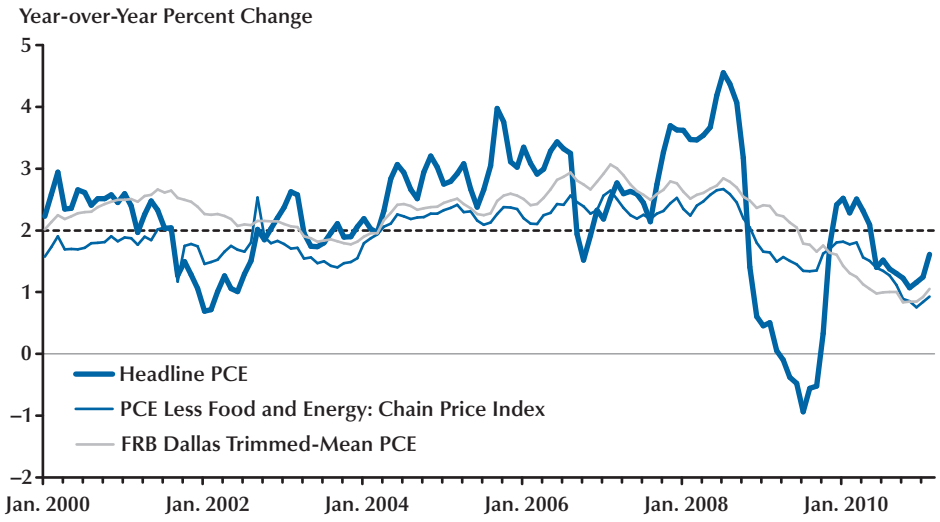
[6] Similar results are discussed in Crone et al. (2008).

[7] See Dolmas (2005). The Federal Reserve Bank of Cleveland constructs an analogous measure for CPI (Bryan and Cecchetti, 1994); see Figure 4.

[8] See also the discussion in Faust and Wright (forthcoming), Thornton (2007), and DiCecio (2007).

Figure 3

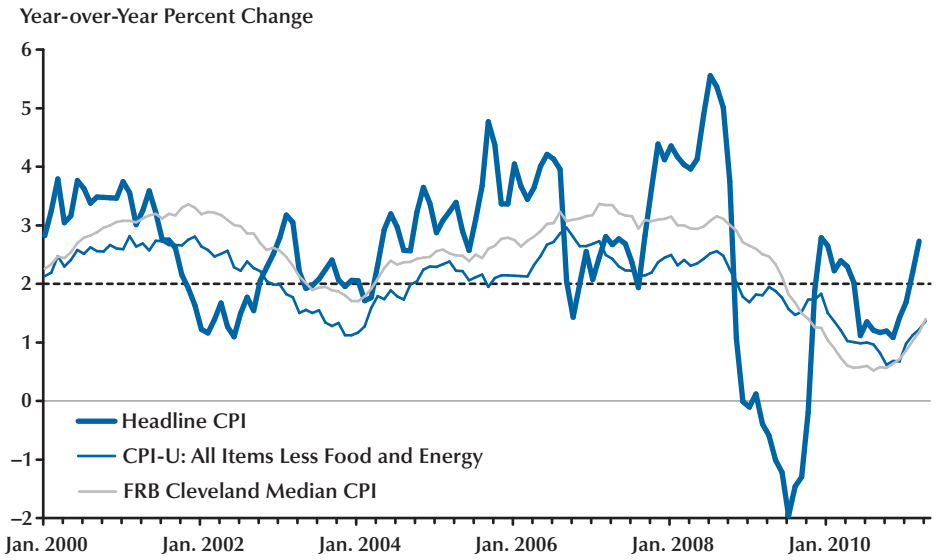
PCE Inflation Measures: Headline, Core, and Trimmed Mean



SOURCE: Bureau of Economic Analysis and Federal Reserve Bank of Dallas. Last observation April 2011.

Figure 4

CPI Inflation Measures: Headline, Core, and Median



SOURCE: Bureau of Labor Statistics and Federal Reserve Bank of Cleveland. Last observation April 2011.

on the time period) has the lowest signal-to-noise ratio, but the food component often does not have the second-lowest signal-to-noise ratio after 1984; for this reason, it is not clear that the food component should be routinely excluded on this basis. Also, the concept of a signal-to-noise ratio contains a notion that the noise component is stationary, whereas much of the contemporary worry about commodity prices is that relative price changes may be much more persistent going forward than they have been in the past. That brings us to the "relative price" argument for confining attention to core inflation.

THE "RELATIVE PRICE" ARGUMENT

The U.S. economy has many thousands of prices, and these prices are adjusting frequently. This is appropriate relative price adjustment. The prices are sending signals to households about what has become more expensive and less expensive at each moment in time. Think of the U.S. household sector as one unified household with a fixed budget constraint. Then if a particular price goes up, generally speaking another price has to fall or, alternatively, the expenditure shares on the two goods have to adjust so that the household can still meet its budget constraint. Either way, the household will adjust by changing behavior in response to the changing prices. The price indexes that are constructed by the Bureau of Labor Statistics (for the consumer price index [CPI]) and by the Bureau of Economic Analysis (for the PCE price index) already make adjustments for this behavior in various ways, especially in chain-weighted indexes that adjust expenditure weights more continuously. In this sense, relative price movements are already accounted for in the construction of existing indexes. Therefore, when the entire price index rises, it really does mean that there is inflation in the economy. Appealing to the idea of relative price change to explain increases in a price index is inappropriate in most circumstances, at least up to our ability to measure behavior induced by relative price changes correctly.

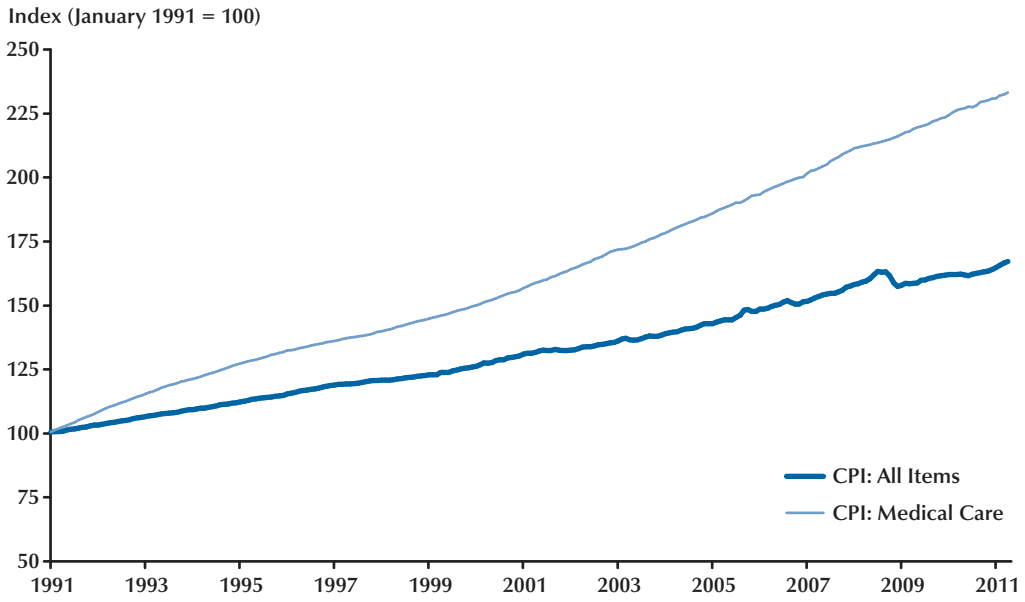
It is often asserted in discussions of this type that the Fed cannot help the fact that events occur around the world each day that affect various prices. Some prices are "out of our control." This is certainly true, but this is also true for every single price in the price index. Each one is affected by real supply and demand factors each day, none of which is susceptible to direct influence by the Federal Reserve. The only element of control the Fed has is over general movements in the entire price index, and only imprecisely over a period of quarters and years.

The key relative price changes in today's global economy are for energy and other commodities. Crude oil prices, in particular, are substantially higher in real terms than they were a decade ago and constitute a significant fraction of global expenditure. It is often asserted that energy prices cannot increase indefinitely—that a one-time rise in energy prices only temporarily contributes to inflation—and therefore that it makes sense to ignore such changes. However, the logic of relative prices suggests that if households are forced to spend more on energy consumption, then they have to spend less on the consumption of all other goods, thereby putting downward pressure on all other prices (and all other expenditure shares) in the economy. Ignoring energy prices would then understate the true inflation rate, as one would be focusing only on the prices facing downward pressure because of changing relative prices.

One might also reasonably question the "temporary" characterization of the shift in energy and other global commodity prices. It is certainly true that we should not expect energy prices to increase faster than the general price level without limit. But it is also true that there are well-known examples of long-term secular trends in certain prices. One example is medical care prices, which for decades have generally increased faster than the headline CPI index (Figure 5). Another example is computing technology, where prices have more or less continuously declined per unit of computing power (Figure 6), even as other prices have continued to rise. So it is possible—and indeed it does happen—that whole sectors of the economy experience relative price change

Figure 5

Price of Medical Care (CPI: Medical Care) and Overall Price Level (CPI: All Items)



SOURCE: Bureau of Labor Statistics and author's calculations. Last observation April 2011.

over very long periods of time. From this perspective, it is at least a reasonable hypothesis that global demand for energy will outstrip increased supply over the coming decades as the giant economies of Asia, particularly India and China, reach Western levels of real income per capita. If that scenario unfolds, then ignoring energy prices in a price index will systematically understate inflation for many years.[9]

SHOULD THE CENTRAL BANK TARGET A SUBSET OF PRICES?

The last set of arguments in favor of a notion of core inflation is far more sophisticated, but also far less established. Up to now we have taken it for granted that the prices that households care about include all the prices that households actually have to pay. This suggests that our existing headline price indexes are the right ones to look at when considering what is best for households. Yet there is some interesting literature that asks the following question: Can we think of a theoretical world in which the central bank would want to target a subset of the prices faced by households, instead of all the prices, on the grounds that this policy would be preferred by the households themselves? We could then call changes in this subset of prices "core inflation."

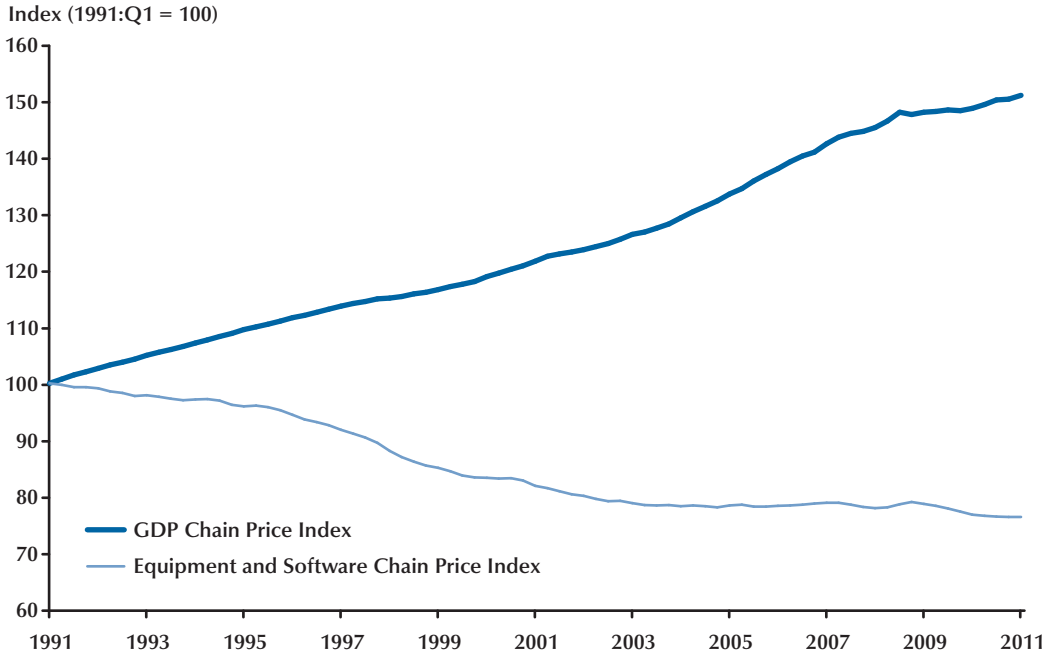
The general answer is that this is indeed possible, and I believe future research in this area has to proceed in this direction. But these models, while interesting, are not ready for prime time, and so I think for now the best we central bankers can do is focus on the best measures of overall inflation we have and attempt to stabilize those.

The key feature of the literature in this area is that some prices are considered "sticky" (in a

[9] The literature trying to disentangle the recessionary consequences of temporary oil-price shocks themselves from the consequences of the monetary policy response to those shocks is interesting but inconclusive. The results depend on many modeling details. See Kormilitsina (2011), Leduc and Sill (2004), and Dhawan and Jeske (2007).

Figure 6

Price of Computing Technology (Equipment and Software Chain Price Index) and Overall Price Level (GDP Chain Price Index)



NOTE: GDP, gross domestic product.

sense made precise in the research), while other prices are fully flexible.[10] For instance, one could posit that the energy price sector is characterized by fully flexible prices, while the rest of the economy is characterized by prices that do not adjust as readily to supply and demand disturbances and therefore are considered sticky.[11] A typical result from the literature is that it is the sticky prices that matter more from the perspective of the households in the model, since those prices are not clearing markets as effectively as they could if prices were fully flexible. For this reason, the central bank might want to focus on a subset of prices, namely the sticky ones. One could think of changes in these sticky prices as core inflation.[12]

This idea has a long way to go to gain general acceptability, and it is certainly not widely endorsed even within macroeconomics. But at least it is one way to think about why it might be better to focus on a subset of prices instead of the entire price index.

There is an international version of this argument as well.[13] In one area of research, there would be a sticky price sector in each country, and each central bank would provide the optimal monetary policy by focusing on the sticky price sector in its own country and ignoring import prices. This would divide up the prices in yet a

[10] There is plenty of argument about how realistic it is to think that sticky price assumptions provide an appropriate foundation for monetary policy intervention. See, for instance, Kehoe and Midrigan (2007).

[11] For a discussion of some of the issues, see Bodenstein, Erceg, and Guerrieri (2008).

[12] See Eusepi, Hobijn, and Tambalotti (forthcoming) for a detailed discussion of an optimal "cost-of-nominal-distortions index" built with this idea in mind.

[13] See, for instance, Clarida, Galí, and Gertler (2002).

different way, and perhaps one that makes a little bit of sense: Import prices would be excluded from domestic policy concerns because the foreign central bank would already be responding to the prices of its exports in setting its own policy (and exchange rates are flexible). Still, results like this depend on a lot of particular assumptions.

At this point in time, ideas like these are not widely entertained outside academic circles. I bring this literature up only to illustrate that there is interesting research about why it may be optimal to focus monetary policy on a subset of prices instead of a headline price index. But the existing literature tends to draw distinctions that are somewhat different from the way practitioners wish to view this issue. Most practitioners do not have in mind trying to divide up prices between those that are more "sticky" and those that are less so, or between domestic prices and import prices.[14]

WHAT SHOULD WE DO?

The theme of my remarks has been that U.S. monetary policy needs to de-emphasize core inflation. Core inflation is not the ultimate goal of monetary policy. I have considered four classes of arguments for a focus on core inflation and found all of them wanting. For this reason, I think the best the FOMC can do is to use headline inflation when looking at the price side of the dual mandate.

Core versus headline inflation has been a long-standing issue for the FOMC. The focus on core inflation in the United States seems to be more entrenched than in many other countries. I have argued that the older ideas justifying this focus have rotted over time—indeed, they probably made little sense from the start. The FOMC needs to get a better playbook on this question so that the Committee can reconnect with American households, who see price changes daily in many of the items the Committee seems to exclude from consideration in making monetary policy.

The headline measures of inflation were designed to be the best measures of inflation available. It is difficult to get around this fact with simple transformations of the price indexes. The Fed should respect the construction of the price indexes as they are and accept the policy problem it poses. To do otherwise may create the appearance of avoiding responsibility for inflation.

There is widespread agreement that headline inflation is the goal variable of monetary policy with respect to prices. Normally one would want to operate directly in terms of the goal variable whenever possible. The concept of core inflation suggests that somehow an intermediate target strategy with respect to price inflation is optimal for U.S. monetary policy. As I have outlined in this article, I do not think this has ever been convincingly demonstrated. In addition, the U.S. focus on core inflation tends to damage Fed credibility. As I noted in the introduction, many other central banks have solidified their position on this question by adopting explicit, numerical inflation targets in terms of headline inflation, thus keeping faith with their citizens that they will work to keep headline inflation low and stable. The Fed should do the same.

[14] For some actual data that attempt to distinguish sticky and flexible price inflation, see the Inflation Project at the Atlanta Fed (www.frbatlanta.org/research/inflationproject/data.cfm) and Bryan and Meyer (2010).

REFERENCES

Bodenstein, Martin; Erceg, Christopher J. and Guerrieri, Luca. "Optimal Monetary Policy with Distinct Core and Headline Inflation Rates." *Journal of Monetary Economics*, October 2008, *55*(Suppl. 1), pp. S18-S33.

Bryan, Michael F. and Cecchetti, Stephen G. "Measuring Core Inflation," in N. Greg Mankiw, ed., *Monetary Policy*. Chicago: University of Chicago Press, 1994, pp. 195-215.

Bryan, Michael F. and Meyer, Brent. "Are Some Prices in the CPI More Forward Looking than Others? We Think So." Federal Reserve Bank of Cleveland *Economic Commentary* No. 2010-2, May 19, 2010; www.clevelandfed.org/Research/commentary/2010/2010-2.pdf.

Bullard, James. "Measuring Inflation: The Core Is Rotten." Speech to the Money Marketeers of New York University, New York, NY, May 18, 2011; http://research.stlouisfed.org/econ/bullard/pdf/Measuring_Inflation_May_18_2011_FINAL.pdf.

Clarida, Richard H.; Galí, Jordi and Gertler, Mark L. "A Simple Framework for International Monetary Policy Analysis." *Journal of Monetary Economics*, July 2002, *49*(5), pp. 879-904.

Crone, Theodore M.; Khettry, N. Neil K.; Mester, Loretta J. and Novak, Jason A. "Core Measures of Inflation as Predictors of Total Inflation." Working Paper No. 08-9, Federal Reserve Bank of Philadelphia, May 2008; www.philadelphiafed.org/research-and-data/publications/working-papers/2008/wp08-9.pdf.

Dhawan, Rajeev and Jeske, Karsten. "Taylor Rules with Headline Inflation: A Bad Idea." Working Paper No. 2007-14, Federal Reserve Bank of Atlanta, July 2007; www.frbatlanta.org/filelegacydocs/wp0714.pdf.

DiCecio, Riccardo. "Inflation Disconnect?" Federal Reserve Bank of St. Louis *Monetary Trends*, July 2007; http://research.stlouisfed.org/publications/mt/20070701/cover.pdf.

Dolmas, Jim. "A Fitter, Trimmer Core Inflation Measure." Federal Reserve Bank of Dallas *Southwest Economy*, Issue 3, May/June 2005; http://dallasfed.org/research/swe/2005/swe0503b.html.

Eusepi, Stefano; Hobijn, Bart and Tambalotti, Andrea. "CONDI: A Cost-of-Nominal-Distortions Index." (Forthcoming in *American Economic Journal: Macroeconomics*.)

Faust, Jon and Wright, Jonathan H. "Forecasting Inflation." (Forthcoming in Graham Elliott and Allan Timmermann, eds., *Handbook of Economic Forecasting*. Volume 2. New York: Elsevier.)

Kehoe, Patrick J. and Midrigan, Virgiliu. "Sales and the Real Effects of Monetary Policy." Working Paper No. 652, Federal Reserve Bank of Minneapolis, April 2007; www.minneapolisfed.org/research/WP/WP652.pdf.

Kiley, Michael T. "Estimating the Common Trend Rate of Inflation for Consumer Prices and Consumer Prices Excluding Food and Energy." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series No. 2008-38, July 2008; www.federalreserve.gov/pubs/feds/2008/200838/200838pap.pdf.

Kormilitsina, Anna. "Oil Price Shocks and the Optimality of Monetary Policy." *Review of Economic Dynamics*, January 2011, *14*(1), pp. 199-223.

Leduc, Sylvain and Sill, Keith. "A Quantitative Analysis of Oil-Price Shocks, Systematic Monetary Policy, and Economic Downturns." *Journal of Monetary Economics*, May 2004, *51*(4), pp. 781-808.

Mishkin, Frederic S. "Headline Versus Core Inflation in the Conduct of Monetary Policy." Speech at the Business Cycles, International Transmission and Macroeconomic Policies conference, HEC Montreal, Montreal, Canada, October 20, 2007; www.federalreserve.gov/newsevents/speech/mishkin20071020a.htm.

Rich, Robert and Steindel, Charles. "A Comparison of Measures of Core Inflation." Federal Reserve Bank of New York *Economic Policy Review*, December 2007, *13*(3), pp. 19-38; www.ny.frb.org/research/epr/07v13n3/0712rich.pdf.

Smith, Julie K. "PCE Inflation and Core Inflation." Unpublished manuscript, Department of Economics, Lafayette College, January 2010.

Thornton, Daniel L. "Measure for Measure: Headline versus Core Inflation." Federal Reserve Bank of St. Louis *Monetary Trends*, September 2007; http://research.stlouisfed.org/publications/mt/20070901/cover.pdf.

Challenges for Monetary Policy in the European Monetary Union

Axel A. Weber

This article was originally presented as the Homer Jones Memorial Lecture, organized by the Federal Reserve Bank of St. Louis, St. Louis, Missouri, April 13, 2011.

Federal Reserve Bank of St. Louis *Review*, July/August 2011, *93*(4), pp. 235-42.

Four years ago, I was invited to give a speech in Paris. Its title—"From Academic to Policy Maker"—referred to the fact that I started out as an academic (Weber, 2007). I began that speech by mentioning a number of other academics who went on to become central bankers: Mervyn King, Ben Bernanke, Janet Yellen, Bill Poole, and Otmar Issing, to name but a few.

I continued by analyzing why there are so many academics in monetary policy. James Bullard, by the way, is another case in point, whom I did not mention at the time because he was not yet in his current position. One of the main reasons why, over the past years, academic researchers have been taking up leading positions at central banks is that monetary policy itself has been heavily influenced by the findings of academic research. During the financial crisis, monetary policy and economies across the world have benefited significantly from these insights since they have helped us to swiftly apply the appropriate policy responses to contain the crisis. Conversely, the crisis has also raised important issues for academic research in monetary economics as well as in other fields.

As of next month [May 2011], after seven years as a policymaker, I shall be taking up the position of a faculty member of the University of Chicago Booth School of Business. In spite of this upcoming transition "from policymaker to academic," my remarks on the challenges for monetary policy in the European Monetary Union (EMU) are from the policymaker's perspective.

THE FINANCIAL CRISIS AND ITS LESSONS FOR MONETARY POLICY

The financial crisis has brought the "monetary policy consensus" formed in the years prior to the crisis under scrutiny (Bean et al., 2010). The framework of monetary policy differed significantly from one central bank to another. Nevertheless, across the board their primary objective was price stability—defined as a stabilization of the inflation rate at around 2 percent across a horizon of approximately two years. Steering short-term interest rates was considered a sufficient means of achieving this target. Central bank forecasts played a key role in monetary policy decisionmaking, with monetary aggregates increasingly taking a backseat in many forecast models.

Furthermore, capital markets were mostly assumed to be efficient, meaning that financial imperfections and their potential macroeconomic

Axel A. Weber is the former president of the Deutsche Bundesbank and a visiting professor of economics at the University of Chicago Booth School of Business. This paper is based on his Homer Jones Memorial Lecture at the Federal Reserve Bank of St. Louis, April 13, 2011.

© 2011, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

effects were not taken into account. Temporary inefficiencies, such as asset price bubbles, were considered possible, but the majority view was that monetary policy could do little to counteract such developments.

Microprudential supervision was regarded as a sufficient means of containing risks in the financial sector. Monetary policymakers should intervene only *after* a financial crisis had occurred, minimizing the macroeconomic damage through resolute interest rate cuts.

Even though monetary policy proved indispensable and highly successful in containing the crisis and preventing a meltdown of the financial system, events have cast doubt on this consensus. The question now is whether and to what extent monetary policy should take account of financial market developments *before* a crisis occurs (Deutsche Bundesbank, 2011c; also see the following section). Let me elaborate on some aspects of this in greater depth.

A Stronger Role for Financial Markets in Monetary Policy Analysis...

Given the genesis of the crisis, it is undeniable that monetary policy with too short a policy horizon can fail to take account of financial imbalances that eventually spill over to the real economy, thus jeopardizing price stability. So, how should monetary policy incorporate the experience of the crisis into its decisionmaking process?

In the pre-crisis phase, monetary policy decisions were often based on models in which the financial sector played only a minor or no role at all. Therefore, an obvious and important lesson from the crisis is that the theoretical and empirical foundations of monetary policy must place a greater emphasis on both the banking sector and financial imperfections.

As regards the Eurosystem's monetary policy strategy, the monetary pillar already contains major elements of such an approach. In the more recent past, the Eurosystem has stepped up its efforts to continually enhance its monetary analysis (Papademos and Stark, 2010). The aim is to identify irregularities in the patterns of a number of variables, since an unusual pattern in loan developments and monetary aggregates can provide valuable indications of excessive credit growth. This requires, among other things, an extension of the usual decisionmaking horizon, as financial distortions often build over a fairly long period. As a result, monetary policy should become more symmetrical over the financial cycle and can thus make a key contribution to financial stability (Weber, 2010a).

...A Separate Toolkit for Financial Stability...

However, this alone is insufficient to ensure financial stability. Until the crisis, the majority view had been that asset price bubbles are difficult to identify in a timely manner and that interest rates are too blunt a tool to burst such bubbles at an early stage. These reservations have not been invalidated by the crisis and, therefore, the debate on how to better prevent financial crises turned to the specific incentives within the financial system and the existing supervision, which focuses primarily on individual institutions, as these may have encouraged the buildup of debt-financed imbalances. Thus, a greater emphasis should be placed on macroprudential analysis and regulation (Deutsche Bundesbank, 2011b, and Basel Committee on Banking Supervision, 2010). The aim of macroprudential policy is to contain systemic risk, thus strengthening the resilience of the financial system as a whole. This approach is designed to ensure that externalities within the financial system, notably the procyclicality and interconnectedness of financial institutions, can be addressed appropriately.[1] Consequently, existing supervisory tools must be expanded or adjusted to prevent systemic risk from arising in the future and to considerably reduce the likelihood of credit and asset price bubbles.

...Price Stability: Still the Primary Objective of Monetary Policy

Against this background, monetary policy and its tools must remain focused on price stability

[1] A range of tools aimed at curtailing both procyclicality and network risk is currently under discussion. See Bank for International Settlements (2010) and Galati and Moessner (2011).

and should not be overburdened with other objectives. In fact, the credibility of monetary policy depends not only on the clarity of its objectives but also on transparency regarding its limitations. Adopting financial stability as an additional, independent monetary policy objective runs the risk of arousing unrealistic expectations about the effectiveness of monetary policy tools. Nevertheless, central banks as institutions may still be given the additional task of pursuing financial stability, as long as they also have the appropriate set of additional independent tools available to them. Indeed, central banks' expertise constitutes a forceful argument for them to continue to play a prominent role in analyzing and assessing financial stability. The advantage of having independent tools for price stability and financial stability is evident when there is a need for monetary and macroprudential policies to be adjusted in different ways. Nevertheless, as developments in money and financial markets are of key importance for both monetary policy and macroprudential policy, there are likely interdependencies that should be taken into account. For example, bank lending is important not only for the monetary transmission process but is also a link for macroprudential policy. This creates the opportunity for policy decisions in both spheres to complement each another, but it also harbors the danger of them counteracting each other or even canceling each other out.

There is no single answer to the question of how necessary or advantageous a coordination of policy areas would be (Committee on the Global Financial System, 2010). Recent research has provided some initial clues and corroborates the view that the inflation rate can be stabilized quite well if macroprudential policy has its own tools and works alongside monetary policy (Beau, Clerc, and Mojon, 2011, and Christensen, Meh, and Moran, 2010).

However, harmful effects with respect to inflation rate volatility can arise if monetary policymakers ignore the impact of macroprudential tools on the financial markets (see Angelini, Neri, and Panetta, 2010). If central banks make decisions regarding both macroprudential and monetary policy tools, additional fluctuations in the inflation rate compared with the monetary policy status quo can be virtually ruled out, and such fluctuations could even be reduced overall (see also Bean et al., 2010).

These preliminary results should be interpreted with caution. First, the underlying dynamic stochastic general equilibrium models only approximately reproduce the complex interactions between the real and the financial sectors. Second, such research is only in its infancy; at present, only a few models allow a simultaneous analysis of monetary and macroprudential policy (see Beau, Clerc, and Mojon, 2011). Nevertheless, the results confirm that there *should be* a clear allocation of objectives and tools to achieve the aims of both policies. Assuming that there will be a satisfactory exchange of information between both monetary and macroprudential policymakers in the future, the existing studies give no cause to fear that the objective of price stability must be compromised.

Price Stability Should Still Be Understood to Mean Low Inflation Rates

Even though the pre-crisis consensus regarding price stability as the primary objective of monetary policy remains valid, it can be asked whether the experience of the crisis should have implications for the specific form that the objective of price stability takes. Specifically, there have been concerns that the credible commitment to ensure a low rate of inflation might restrict the leeway for monetary policy stabilization, since in the event of massive interest rate cuts the lower bound for nominal interest rates would be hit quite quickly. Two competing approaches have been suggested to deal with this alleged shortcoming: a higher inflation target (Williams, 2009, and Blanchard, Dell'Ariccia, and Mauro, 2010) and a switch to targeting the price level or, more precisely, the price-level path.[2] Neither of these two alternatives convinces me.

As regards a higher inflation target, it is not only the substantial and ongoing welfare losses accompanying a rise in the inflation target that

[2] Eggertsson and Woodford (2003) already proposed price-level targeting in connection with the deflation experienced in Japan.

argue against this proposal but also, above all, the loss of credibility for monetary policy associated with such a discretionary measure (Weber, 2010b). The resulting destabilization of inflation expectations would make it significantly more difficult for the central bank to achieve its (possibly higher) inflation target and to safeguard macroeconomic stability.

Compared with a strategy of inflation targeting, price-level targeting does offer a number of advantages, at least in theory. It opens up the option of influencing private sector inflation expectations and thus of combating deflationary risks in the event of a crisis. However, it is doubtful whether a change in the target specification in the event of acute deflationary risk would be suitable for achieving the desired positive effect on private sector inflation expectations (Walsh, 2010). A more serious problem is that a strategy of price-level targeting is associated with a few additional drawbacks compared with optimal monetary policy, casting doubt on whether such a change of strategy would be beneficial (see Deutsche Bundesbank, 2011a, and Gerberding, Gerke, and Hammermann, 2010).

All in all, this means that neither raising the inflation target nor switching to price-level targeting would be appropriate from an economic stability point of view. Instead, this problem must be tackled at its root; the existing wrong incentives and regulatory loopholes must be eliminated to render crises less likely and less severe. It is, in any case, questionable whether the leeway available to monetary policy at the lower bound of the nominal short-term money market rates actually was that limited. The effectiveness of central banks' unconventional measures during the crisis gives no cause to view the lower bound of the interest rate as a binding restriction on the effectiveness of monetary policy.

PARTICULAR LESSONS FOR MONETARY POLICY IN THE EURO AREA

All the issues I have mentioned until now concern more or less all central banks and every monetary policymaker. I now turn my focus to the particular challenges for monetary policy in the euro area. These arise from the sovereign debt crisis, which is the major challenge for economic and monetary union. The circumstances surrounding the debt crisis are aggravating the conduct of the Eurosystem's common monetary policy, which is geared toward maintaining price stability in the euro area as a whole.

Heterogeneity as a Challenge for Monetary Policy

One of the aggravating factors is heterogeneity in terms of growth, inflation, and competitiveness. With regard to the euro-area countries' economic performance, we are currently observing a widening divergence. Broadly speaking, there is a considerable growth gap between the core and the periphery, or to put it more precisely, *some* peripheral countries of the euro area.

In my view, the economic heterogeneity of the euro area is a non-issue. Why should heterogeneity be a problem for the single monetary policy? After all, the dispersion of growth rates, as measured by the weighted standard deviations of quarterly growth rates, is not significantly greater than in the first years of the EMU. With regard to inflation variance, we now see even lower values than then. Furthermore, the U.S. economy is characterized by considerable heterogeneity, too, and that does not impede the Federal Reserve's monetary policy, either. And in much the same way as the Federal Open Market Committee is focused on the United States as a whole, the Governing Council of the European Central Bank must take a euro-area–wide perspective: While national developments have to be taken into consideration, monetary policy cannot be tailored to the specific needs of individual member states.

The real problem with heterogeneity—and that is a concern to me—is that a number of countries have obviously failed to meet the obligations and requirements of a currency union. The persistent problems of countries in refinancing their debt are only the *symptoms* of the problems, not the problem itself. The financial crisis

has revealed unsustainable developments in some member countries— developments that were already in existence before the crisis: too much public spending, unproductive use of capital inflows, loss of competitiveness. These were just some of the shortcomings that had been carelessly neglected, not least by the financial markets. Painful adjustment processes, including structural reform and budget consolidation, are essential to restore the ability of the countries concerned to live up to the demands of the single monetary policy.

Fiscal Stabilization Measures Were Necessary but They Undermined the Basic Founding Principles of the EMU

Ensuring financial stability in the euro area required and justified fiscal aid for Greece and the establishment of a temporary stabilization mechanism. Nevertheless, these particular measures have undermined the foundations of the EMU.

The establishment of the EMU was based on principles deemed necessary to make the euro a stable currency. According to the principle of subsidiarity, economic policies other than monetary policy remain the responsibility of national governments. With regard to fiscal policy, rules and institutional arrangements were established to ensure sound fiscal policies in the member states. Furthermore, a "no-bailout" clause stipulated the national responsibility of each country for repaying its own public debt.

Rules for sound public finances are of particular importance in a monetary union since the incentives for excessive borrowing are even greater in a monetary union than they are anyway. Excessive borrowing can also place a strain on the conduct of a stability-oriented monetary policy. Unsound public finances are the Achilles' heel of a monetary union of independent states.

Purchases of government bonds for monetary policy purposes, for example, harbor the risk of blurring the boundaries between monetary and fiscal policy, particularly given high government deficits and debt levels. Such actions might harm the credibility of monetary policy. A little earlier, I said that monetary policy must remain focused on price stability and should not be overburdened with other objectives. This principle applies not only with respect to financial stability, but also fiscal policy.

During the financial crisis, the Eurosystem—like the central banks of other major economic regions—took unconventional monetary policy measures on an unprecedented scale. The ample provision of liquidity was effective in offsetting the consequences of the abrupt decline in market liquidity, in maintaining monetary policy transmission, and, ultimately, in helping to prevent the real economy from sliding into a prolonged depression. On the other hand, unlimited provision of central bank liquidity to banks without a sustainable business model cannot be a long-run solution. Again, monetary policy should not act as a substitute for tasks of other policy areas. In particular, monetary policy should not and cannot persistently replace the repair of banks' balance sheets. The phasing-out of non-standard measures has to be continued; the objective is to return to the pre-crisis operational framework which has proven its effectiveness and flexibility during the crisis.

Economic Governance in the Euro Area Needs Reform

Since the fiscal stabilization measures in favor of euro-area peripheral countries have undermined the basic principles of the EMU, it is obvious that there has to be a fundamental and far-reaching reform of economic governance in the euro area. First, the European leaders agreed that the fiscal rules must be tightened since their application in practice had proven too weak. Second, they agreed that macroeconomic imbalances should be addressed earlier and more effectively. The crisis demonstrated that sound public finances are a necessary, but not sufficient, condition for financial and economic stability. Ireland, for instance, was among the least-indebted countries of the euro area before the crisis erupted. Finally, the leaders agreed to establish a permanent stabilization mechanism since it is an illusion to believe that a reform of economic governance

might prevent the reoccurrence of fiscal crises in the future.

In March, the leaders agreed on what they view as a comprehensive package. Its measures certainly represent a step in the right direction, but they are not a "quantum leap towards strengthening the institutional framework of EMU" (European Central Bank, 2010, p. 4), which is required to reinforce economic governance in the euro area. Ultimately, the future success of the EMU will hinge crucially on the member states' willingness to comply with the tighter set of rules.

CONCLUSION

Tomorrow, it will be 17 years ago to the day since Helmut Schlesinger, one of my predecessors as president of the Bundesbank, gave the Eighth Homer Jones Memorial Lecture. In his speech, "On the Way to a New Monetary Union: The European Union," he explained the historic dispute between "monetarists" and "economists" (Schlesinger, 1994). In the particular context of European monetary integration, these terms had a totally different meaning than our general understanding. "'Monetarists,'" he explained, "believed that monetary integration has to start first and that economic and political integration would follow." "'Economists,'" however, "believed that economic convergence between the national economies must occur before…a monetary union."

The "monetarists" prevailed, but they erred in their belief that the introduction of the single currency would automatically act as a locomotive for the political union of Europe. There is no political union so far and there is little expectation that this might change significantly in the foreseeable future. Therefore, national executive and legislative branches will remain responsible for economic and fiscal policies over the medium to long term. Intergovernmental fiscal transfers beyond the rather moderate and earmarked payments from the European Union budget (approximately 1 percent of gross national product) are hardly acceptable; overburdening the financial solidarity of the people might jeopardize the idea of European integration.

The "economists," on the other hand, had a point in demanding more economic convergence. Their worries were, by the way, taken into account by the implementation of convergence criteria that must be fulfilled before a country can join the euro area. The underlying problem of the current crisis is, however, not a lack of convergence ex ante or heterogeneity per se; rather, it is the lack of willingness on the part of a number of member states to meet the requirements of the membership in a monetary union. If they fail to correct these deficiencies swiftly and thoroughly, stability-oriented monetary policy in the EMU will become increasingly difficult, all the more so as monetary policy has been profoundly challenged by the financial crisis.

The major lessons that central bankers in the euro area and elsewhere should take to heart are the following: First, monetary policy has to consider the implications of financial instability for price stability; monetary and credit aggregates can provide helpful information in this regard. Second, since the policy rate remains too blunt a tool to tackle financial imbalances, the objective of financial stability requires its own, macroprudential set of tools, whereas maintaining price stability should remain the primary objective of monetary policy. Third, price stability should continue to be seen as a stable and low inflation rate. Finally, without stability-oriented prudent fiscal policy, it will be increasingly difficult for monetary policy to ensure price stability at low interest rates.

REFERENCES

Angelini, Paolo; Neri, Stefano and Panetta, Fabio. "Grafting Macroprudential Policies in a Macroeconomic Framework: Choice of Optimal Instruments and Interaction with Monetary Policy." Unpublished manuscript, Banca d'Italia, March 5, 2010; www.bis.org/repofficepubl/hkimr201007.15.pdf.

Bank for International Settlements. *80th Annual Report*. Basel, Switzerland: Bank for International Settlements, 2010; www.bis.org/publ/arpdf/ar2010e.pdf.

Basel Committee on Banking Supervision. "BASEL III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Basel, Switzerland: Bank for International Settlements, December 2010; www.bundesbank.de/download/bankenaufsicht/pdf/basel3_capital_rules.pdf.

Bean, Charles; Paustian, Matthias; Penalver, Adrian and Taylor, Tim. "After the Fall," in *Macroeconomic Challenges: The Decade Ahead*. Proceedings of the 2010 Jackson Hole Economic Policy Symposium, Jackson Hole, Wyoming, August 26-28, 2010. Kansas City, MO: Federal Reserve Bank of Kansas City, 2010; www.kansascityfed.org/publicat/sympos/2010/bean-paper.pdf.

Beau, Denis; Clerc, Laurent and Mojon, Benoît. "Macro-Prudential Policy and the Conduct of Monetary Policy." Unpublished manuscript, Banque de France, January 2011; www.benoitmojon.com/pdf/Beau-Clerc-Mojon.pdf.

Blanchard, Olivier J.; Dell'Ariccia, Giovanni and Mauro, Paolo. "Rethinking Macroeconomic Policy." *Journal of Money, Credit, and Banking*, September 2010, *42*(Suppl. 1), pp. 199-215.

Christensen, Ian; Meh, Césaire and Moran, Kevin. "Bank Leverage Regulation and Macroeconomic Dynamics." Unpublished manuscript, Bank of Canada, 2011.

Committee on the Global Financial System. "Macroprudential Instruments and Frameworks: A Stocktaking of Issues and Experiences." CGFS Paper No. 38, Committee on the Global Financial System, Bank for International Settlements, May 2010; www.bis.org/publ/cgfs38.pdf.

Deutsche Bundesbank. "Price-Level Targeting as a Monetary Policy Strategy." *Monthly Report*, January 2011a, pp 31-45; www.bundesbank.de/download/volkswirtschaft/mba/2010/201001mba_en_price.pdf.

Deutsche Bundesbank. "Approaches to the Measurement and Macroprudential Treatment of Systemic Risk." *Monthly Report*, March 2011b, pp. 37-52.

Deutsche Bundesbank. "The Implications of the Financial Crisis for Monetary Policy." *Monthly Report*, March 2011c, pp. 53-68.

Eggertsson, Gauti B. and Woodford, Michael. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity*, 2003, *1*, pp. 139-211; www.newyorkfed.org/research/economists/eggertsson/BrookingsPaper.pdf.

European Central Bank. "Reinforcing Economic Governance in the Euro Area." European Central Bank, June 20, 2010; www.ecb.int/pub/pdf/other/reinforcingeconomicgovernanceintheeuroareaen.pdf.

Galati, Gabriele and Moessner, Richhild. "Macroprudential Policy: A Literature Review." BIS Working Paper No. 337. Bank for International Settlements, February 2011, www.bis.org/publ/work337.pdf.

Gerberding, Christina; Gerke, Rafael and Hammermann, Felix. "Price-Level Targeting When There Is Price-Level Drift." Discussion Paper, Series 1, No. 23/2010; Deutsche Bundesbank Research Centre, December 30, 2010; www.bundesbank.de/download/volkswirtschaft/dkp/2010/201023dkp.pdf.

Papademos, Lucas D. and Stark, Jürgen, eds. *Enhancing Monetary Analysis*. Frankfurt, Germany: European Central Bank, 2010.

Schlesinger, Helmut. "On the Way to a New Monetary Union: The European Monetary Union." Federal Reserve Bank of St. Louis *Review*, May/June 1994, *76*(3), pp. 3-10; http://research.stlouisfed.org/publications/review/94/05/Union_May_Jun1994.pdf.

Walsh, Carl. "The Future of Inflation Targeting." Unpublished manuscript, University of California, Santa Cruz, September 2010; http://people.ucsc.edu/~walshc/MyPapers/walsh_ace_20100927.pdf.

Weber, Axel A. "From Academic to Policy Maker." *BIS Review* No. 56, June 2007; www.bis.org/review/r070601a.pdf.

Weber, Axel A. Comment on Jordi Galí, "The Monetary Pillar and the Great Financial Crisis," in *The Great Financial Crisis: Lessons for Financial Stability and Monetary Policy. A Colloquium in Honour of Lucas Papademos*. Frankfurt, Germany: European Central Bank, May 20-21, 2010a.

Weber, Axel A. "Der IWF spielt mit dem Feuer (The IMF Is Playing with Fire)." *Financial Times Deutschland*, February 25, 2010b (available in German only); www.bundesbank.de/download/presse/publikationen/20100225_namensbeitrag_weber.pdf.

Williams, John C. "Heeding Daedalus: Optimal Inflation and the Zero Lower Bound." *Brookings Papers on Economic Activity*, Fall 2009, pp. 1-37.

Decomposing the Education Wage Gap: Everything but the Kitchen Sink

Julie L. Hotchkiss and Menbere Shiferaw

The authors use a multitude of data sources to provide a comprehensive, multidimensional decomposition of wages across both time and educational status. Their results confirm the importance of investments in and use of technology, which has been the focus of most of the previous literature. The authors also show that demand and supply factors played very different roles in the growing wage gaps of the 1980s and 1990s.

Federal Reserve Bank of St. Louis *Review*, July/August 2011, *93*(4), pp. 243-71.

BACKGROUND

There is a clear consensus in the economics literature that the gap in wages between more highly skilled and less-skilled workers has been increasing. Research findings on this topic agree that the gap began to widen considerably in the 1970s (for example, see Piketty and Saez, 2003). Much of the focus on the growing wage gap is motivated by its implications for income inequality. Whether income inequality serves as an engine of economic growth by providing powerful incentives or acts as a hindrance to economic potential, a clear picture of the driving forces behind its growth is essential to inform the debate.

Much of the literature places the blame for the growing skills wage gap on increasing returns to postsecondary education. Ingram and Neumann (2006), however, argue that years of education is a weak measure of skill in the analysis of wage distribution and that much more skill heterogeneity exists among workers. They find that the return to years of education remains constant after controlling for skills. However, given the high degree of correlation between education and skill and the fact that education is typically the mechanism through which one achieves a higher skill level, this paper focuses on education-based wage differentials rather than skill differentials and refers to individuals with more education as highly skilled workers and those with less education as lower-skilled workers.

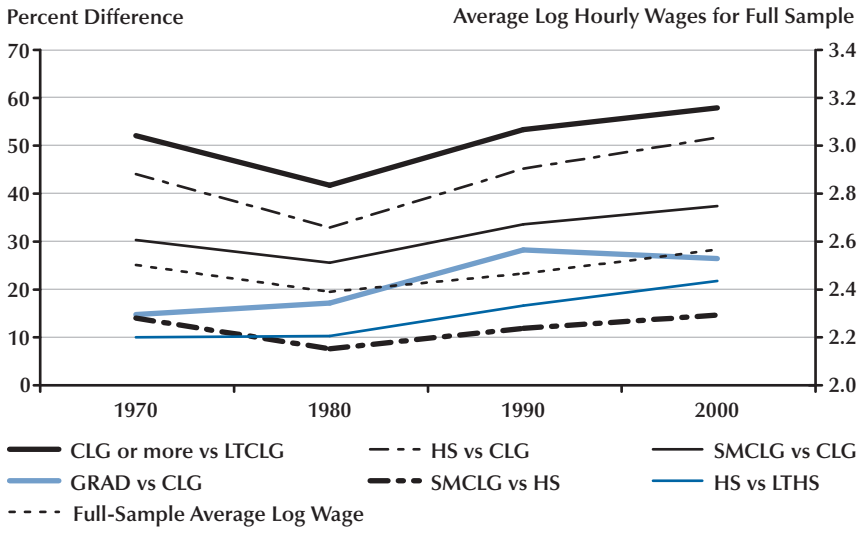
Figure 1 uses data from this article's analysis to show how the wage gaps between education groups have changed from 1970 to 2000. Guvenen and Kuruscu (2007) find that the overall wage inequality between the college group and the high school group rose only modestly during the 1970s because the between-group inequality was actually falling as within-group inequality was rising. This is consistent with the means plotted in Figure 1; the gap between high school and college and the gap between college and more and less than college (between-group comparisons) fell fairly dramatically, but the gap between high school and less than high school and the gap

Julie L. Hotchkiss is a research economist and policy adviser at the Federal Reserve Bank of Atlanta and adjunct professor at Georgia State University. Menbere Shiferaw is a quantitative research analysis specialist at the Federal Reserve Bank of Atlanta. The authors thank Mookie Hojiwala and Anne Flatness at the Bureau of Economic Analysis for assistance with trade data; David Autor, David Dorn, and Christopher Cunningham for assistance with commuting zone questions; and Mark Leonard, participants at the Western Economic Association meetings, and participants at the Federal Reserve System Committee on Applied Microeconomics meetings for comments. Valuable research assistance was provided by Andrew Balthrop, Nicole Baerg, and Fernando Rios-Avila.

© 2011, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

Figure 1

Education Wage Gap Over Time



NOTE: LTHS, less than high school; HS, high school; LTCLG, less than college; SMCLG, some college; CLG, college; GRAD, more than college.

between graduate and college (both could be considered more within-group comparisons) have risen. After 1980, however, the wage gaps between all but one pair of education groups grew, with some moderation of that growth since 1990.

If the labor market can be thought of as two sectors—one that employs skilled workers and one that employs less-skilled workers—the literature suggests multiple supply and demand reasons for the earnings gap growth. The most widely hypothesized reason for the increase in the earnings gap is an increase in demand for skilled workers resulting from technological change, or skill-biased technological change. As industries/firms increase the adoption of computer-based technologies into their production processes in response, for example, to the decline in the price of technology or the abundance of relatively cheap skilled labor, their demand for skilled workers increases. The "skilled worker" in this case includes those who know how to use the technology and those whose productivity is enhanced by computers.

Autor, Katz, and Kearney (2006) find that computerization has not only increased the demand for highly skilled workers (those with abstract thinking-type jobs complemented by computers), but has also decreased the demand for intermediate-skilled workers (those with routine task-type jobs easily replaced by computers). This increase in demand for skilled workers—either *ceteris paribus* or accompanied by a decline in demand for intermediate-skilled, less-educated workers—will increase the education wage gap.

As the demand for skilled labor increases, the returns to a college education should also increase, which, in turn, should lead to an increase in the supply of educated workers, which should put downward pressure on the skills wage gap. However, the wage gap has continued to increase. Consistent with this observation, Crifo (2008) argues that the increased demand for skill among educated workers results in fewer workers with ordinary skills seeking higher education. The net result is a reduction in the supply of educated workers available to meet the growing demand,

thus contributing an additional factor that increases the wage gap.[1]

Card and Lemieux (2001) analyze the wage gap between college and high school graduates for younger and older men and find that the education wage gap for older workers has remained relatively stable while the gap among younger workers has risen sharply since the mid-1970s. Their explanation, also consistent with analysis in Topel (1997), is that the relative supply of young college-educated workers has slowed, while that of older college-educated workers has remained steady. Thus, because the current demand for college labor is increasing faster than the supply, wage inequality continues to increase. Lemieux (2006) provides additional documentation that increasing returns to postsecondary education account for most of the growth in wage inequality.

Card and DiNardo (2002), among others, are critical of skill-biased technological change as the source of the growing wage gap (especially since 1980). The primary basis for this criticism is that although technology continued to advance dramatically through the 1990s, the growth in skill-based earnings inequality was much slower than in the 1980s. In addition, researchers have identified a number of alternative potential contributors to the growing wage gap. Some examples include (i) declining unionization, as in Card and DiNardo (2002), which would result in lower wages among workers in sectors more likely to be unionized—the less-skilled; (ii) the increased labor force participation of women, as in Topel (1997), which would increase the supply of workers to traditionally lower-paying occupations; (iii) shifts in immigration source countries, as in Topel (1997), which has more recently increased the supply of less-skilled workers from Latin America; and (iv) shifts in product demand, as in Autor, Katz, and Kearney (2006). Piketty and Saez (2003) cite a trend in reporting stock options as wages and changing social norms regarding what is an acceptable "high wage" as contributors to the measured growth in the wage gap. Topel (1997) explores a number of potential supply-side contributors to the wage gap and finds that the weight of evidence for the growing gap falls on increasing returns to education for explaining the growth in earnings inequality.

Our paper joins this vast literature in an attempt to contribute a better understanding of the relative contributions of different supply and demand factors in explaining the growing earnings inequality between education levels. The contribution of the analysis in this paper includes using a multitude of data sources in an attempt to capture more of the variation across demand and supply factors that affect workers' wages across educational groups. As Kranz (2006) identified, many previous studies focus on either demand or supply factors. While Kranz's (2006) goal was to exhaust both supply and demand factors in the aggregate, comparing changes in the wage gap across countries, our goal is to do so in an analysis at the individual-worker level.

In addition, contributions of the composition of groups of workers and how their characteristics translate into wages are decomposed not only across groups, but also across time in a fairly straightforward way to directly address the question of the relative importance of different contributors to the changing wage gap. The analysis is at an individual level but incorporates local labor market variations through regressors, such as immigration, mobility, and unemployment rates, at the commuting zone (CZ) level. The advantage of using regressors at the CZ level, as opposed to regressors measured at the metropolitan statistical area or county level, is that this area measure better characterizes the actual labor market in which a worker's wages are determined. For example, in addition to the possibility that immigration status may affect a worker's wage, it is well known that immigrants tend to be geographically concentrated; thus, capturing this labor market-specific concentration—as well as changes in concentration—might be important in explaining wage differences across education groups.

The analysis herein confirms the previously documented importance of technology in explaining the wage gap growth during the 1980s and 1990s. However, our specification allows us to move beyond this simple conclusion and identify the mechanisms through which technology boosted the wages of both highly skilled and lower-skilled workers.

[1] Goldin and Katz (2008, especially Chap. 3) also share the view that the supply of educated workers has not recently kept up with the demand.

METHODOLOGY AND DATA

The strategy used to examine changes in the education wage gap over time is a straightforward, reduced-form approach that relates numerous supply and demand factors to the measured change in the wage gap between workers with varying levels of educational attainment. The analysis is at the individual level, which allows for a truly marginal analysis of the impact of the change in each of the factors on the observed change in the wage gap between two periods.

Methodology

The determinants of the measured wage of two education groups (A and B) are estimated in three time periods (1980, 1990, and 2000). The change in the wage gap (*WG*) between the two education groups and between two time periods (*j* and *k*) can be expressed as

$$(1)\quad \begin{aligned} &WG^k_{A,B} - WG^j_{A,B} \\ &= \left[\ln W^k_A - \ln W^k_B\right] - \left[\ln W^j_A - \ln W^j_B\right], \end{aligned}$$

where log wages of worker with education *i* in time period *t* are described as

$$(2)\quad \ln W^t_i = \mathrm{X}^t_i\beta^t + \mathrm{Y}^t_i\alpha^t + \mathrm{Z}^t_i\delta^t + \varepsilon^t_i,$$

where X^t_i is a vector of demand factors in time *t* that would be expected to affect the wage of this worker and would typically be measured at the industry, occupation, or CZ level; Y^t_i is a vector of supply factors (mostly measured at the individual or CZ level); and Z^t_i is a vector of CZ, institutional, and other characteristics expected to affect the labor market environment in which wages are being determined.

Full descriptions of the regressors and their expected contribution to wage determination are provided in Table A1 (Appendix A). Worker demand regressors include characteristics that describe or are brought to the labor market by employers. Specifically, these include industry-level investment in computers and computer software, individual-level expected use of computers at work, industry-level value added, and industry and occupation CZ employment shares.

Supply regressors include characteristics that describe or are brought to the labor market by workers. Specifically, these include lagged values of immigrant penetration; demographics such as race, gender, and marital status; human capital measures, which include age and expected home computer use; an indicator for the presence in a CZ of at least one postsecondary institution offering a bachelor's degree; the share of the CZ workforce that is female; and lagged values of CZ population and share of the population with the worker's same level of education.

Institutional characteristics are factors not specifically brought by either employers or workers but which still describe the environment of the labor market. These include the extent of unionization within a worker's industry, the CZ unemployment rate, mobility rate of the population in a worker's CZ, and industry and occupational dummy variables.

The wage gap estimated for each pair of skill groups and years is decomposed as follows:

$$(3)\quad \begin{aligned} &WG^k_{A,B} - WG^j_{A,B} \\ &= \left[\ln W^k_A - \ln W^k_B\right] - \left[\ln W^j_A - \ln W^j_B\right] \\ &= \left[\mathrm{T}^k_A\Omega^k_A - \mathrm{T}^k_B\Omega^k_B\right] - \left[\mathrm{T}^j_A\Omega^j_A - \mathrm{T}^j_B\Omega^j_B\right] \\ &= \Omega^k_A\left(\mathrm{T}^k_A - \mathrm{T}^j_A\right) + \mathrm{T}^j_A\left(\Omega^k_A - \Omega^j_A\right) \\ &+ \left[-\Omega^k_B\left(\mathrm{T}^k_B - \mathrm{T}^j_B\right)\right] + \left[-\mathrm{T}^j_B\left(\Omega^k_B - \Omega^j_B\right)\right], \end{aligned}$$

where T = [XYZ] and $\Omega = [\beta\,\alpha\,\delta]'$.[2] This decomposition is structured to determine how much of the wage gap growth between years *j* and *k* can be explained by changes in the endowments of skill groups (e.g., use of a computer at home, mobility) and how much can be explained by changes in how the respective labor markets value those endowments (differences in estimated coefficients across time). If a term is estimated to be positive, the difference (in college or high school graduates' characteristics between the two years or in estimated valuation of those characteristics) contributes positively to the growing

[2] Also see Wellington (1993), who uses this same decomposition to explore changes in the male/female wage gap.

skills wage gap. If a term is estimated to be negative, it reduces the measured skills wage gap.

The decomposition has four terms. The first term, $\Omega_A^k(T_A^k - T_A^k)$, indicates the contribution to the wage gap growth of changes in endowments of workers in skill group A between years *j* and *k*. The second term, $T_A^j(\Omega_A^k - \Omega_A^j)$, indicates the contribution to the wage gap growth of the change in valuation of endowments of workers in skill group A between years *j* and *k*. The third term, $[-\Omega_B^k(T_B^k - T_B^j)]$, indicates the contribution to the wage gap growth of changes in endowments of workers in skill group B between years *j* and *k*. And the fourth term, $-T_B^j(\Omega_B^k - \Omega_B^j)$, indicates the contribution to the wage gap growth of the change in valuation of endowments of workers in skill group B.

Data

The data for the wage gap analysis are from several sources. Details and variable descriptions of data sources can be found in Appendix A. Major data sources include the Integrated Public Use Microdata Series, National Income and Product Accounts, Department of Commerce, National Bureau of Economic Analysis, and the Current Population Survey. The main data source is the Integrated Public Use Microdata Series, from which individual-level data on wages, human capital, demographics, and institutional factors were extracted. We include all workers 18 to 64 years of age. In addition to providing the individual-level wage, education, and other demographic characteristics, these data also provide the CZ-level characteristics included in the regression (the construction of CZs is also described in Appendix A). CZ characteristics are constructed using the same sample of 18- to 64-year-olds. CZ-level characteristics are expected to capture the importance of changing local labor market characteristics in determining changes in the skills wage gap. As pointed out by Autor and Dorn (2008), this level of aggregation is preferred to using (i) metropolitan statistical areas, which exclude individuals not located in a metropolitan area, and (ii) counties, which reflect artificial geographic boundaries.

Consistent with most of the literature on skills wage gaps or income inequality, such as Lemieux (2006), we make several decisions regarding top-coded and outlier observations (in hours or earnings). Our outlier restrictions are binding on the top end in that it is highly unlikely that top-coded earnings would have survived our outlier restrictions. We drop all observations with reported hours top-coded at 99 hours per week; this amounts to 0.21 percent of the 1980 sample, 0.38 percent of the 1990 sample, and 0.06 percent of the 2000 sample. In addition, using real 2000 dollars, we drop observations if individuals earned less than $1 per hour or more than $1,000 per hour. These restrictions result in a loss of 0.55 percent of the 1980 sample, 0.27 percent of the 1990 sample, and 0.14 percent of the 2000 sample.

Regressors are separated into groups based on the mechanism through which they are expected to affect wages. For example, demand for more highly skilled workers is expected to be related to the increase in employer investment in computer hardware and software. If employment in a worker's industry represents a relatively smaller share of overall employment in the worker's local labor market, it is expected that demand for workers, and thus wages, will be lower in that industry. In addition, increases in immigration that bring a competing skill type to a local market are expected to exert downward pressure on the wages of workers of that skill type. Table 1 presents sample means for the regressors used in the analysis, separated by whether the regressor is expected to capture the influences of demand, supply, demographic, or institutional factors on wages; sample distributions across industries and occupations are also provided.

Clearly, the classification of regressors as supply or demand influences is somewhat arbitrary. Generally, we classify factors that come to the labor market through the worker as supply factors and factors that come to the labor market through the employer as demand factors. The number of observations ranges from roughly 1.5 million high school graduates and 375,000 college graduates in 1980 to 1.8 million high school graduates and 922,000 college graduates in 2000. The characteristics of workers, employers, and

Table 1

Sample Means by Year

Regressors	1980	1990	2000
Demand factors			
Computer investment ($ billions)	0.529	1.082	2.000
	(0.639)	(1.252)	(2.650)
Software investment ($ billions)	0.291	1.065	3.456
	(0.320)	(0.994)	(3.399)
Probability of computer use at work	0.228	0.378	0.513
	(0.138)	(0.180)	(0.206)
Industry value added ($ trillions)	0.159	(0.207)	(0.280)
	(0.140)	(0.180)	(0.231)
Industry employment share	0.811	0.829	0.813
	(0.056)	(0.047)	(0.049)
Occupation employment share	0.813	0.831	0.816
	(0.061)	(0.053)	(0.054)
Supply factors			
Probability of computer use at home	0.523	0.566	0.809
	(0.166)	(0.147)	(0.096)
Female share of CZ labor force	0.427	0.454	0.466
	(0.019)	(0.015)	(0.013)
Share of CZ with less than high school diploma	0.261	0.163	0.133
	(0.067)	(0.054)	(0.047)
Share of CZ with high school diploma	0.391	0.349	0.323
	(0.052)	(0.060)	(0.061)
Share of CZ with college degree	0.089	0.135	0.157
	(0.020)	(0.038)	(0.043)
Share of CZ with postgraduate degree	0.073	0.067	0.081
	(0.024)	(0.024)	(0.030)
Share of CZ born in North America (excluded)	0.921	0.897	0.857
	(0.077)	(0.107)	(0.127)
Share of CZ born in Latin America or Caribbean	0.027	0.043	0.071
	(0.043)	(0.066)	(0.080)
Share of CZ born in Europe or Asia	0.040	0.047	0.061
	(0.031)	(0.040)	(0.052)
Share of CZ born in other non–North American countries	0.013	0.013	0.011
	(0.016)	(0.016)	(0.012)
University or college in CZ = 1	0.982	0.968	0.961
	(0.133)	(0.177)	(0.193)
Demographics			
Age (years)	36.22	37.58	39.27
	(12.85)	(11.94)	(12.17)
Female = 1	0.446	0.459	0.491
	(0.497)	(0.498)	(0.500)

NOTE: Standard errors are listed in parentheses. The white race category may include respondents of Hispanic ethnicity.

Table 1, cont'd

Sample Means by Year

Regressors	1980	1990	2000
Demographics, cont'd			
White (excluded) = 1	0.873	0.879	0.783
	(0.333)	(0.326)	(0.412)
Black = 1	0.101	0.087	0.097
	(0.301)	(0.281)	(0.296)
Asian = 1	0.018	0.026	0.038
	(0.135)	(0.159)	(0.190)
Other race = 1	0.008	0.008	0.082
	(0.091)	(0.090)	(0.274)
Married with spouse present = 1	0.622	0.613	0.576
	(0.485)	(0.487)	(0.494)
Institutional factors			
CZ unemployment rate	0.061	0.059	0.050
	(0.020)	(0.016)	(0.015)
Percent of workers covered by union in industry	0.201	0.160	0.119
	(0.148)	(0.128)	(0.119)
Mobility rate of CZ population	0.120	0.122	0.124
	(0.060)	(0.054)	(0.051)
Industries			
Natural resources and mining = 1	0.030	0.030	0.025
	(0.171)	(0.170)	(0.156)
Construction = 1	0.066	0.077	0.076
	(0.248)	(0.266)	(0.265)
Manufacturing = 1	0.259	0.215	0.158
	(0.438)	(0.411)	(0.365)
Transportation and utilities = 1	0.058	0.059	0.055
	(0.233)	(0.235)	(0.228)
Wholesale trade = 1	0.048	0.052	0.037
	(0.214)	(0.223)	(0.188)
Retail trade = 1	0.169	0.141	0.123
	(0.375)	(0.348)	(0.328)
Financial activities = 1	0.047	0.056	0.067
	(0.212)	(0.230)	(0.251)
Information = 1	0.019	0.026	0.030
	(0.138)	(0.160)	(0.170)
Professional and business services = 1	0.062	0.067	0.094
	(0.241)	(0.251)	(0.291)
Education and health services = 1	0.176	0.202	0.207
	(0.381)	(0.402)	(0.405)
Leisure and hospitality = 1	0.023	0.027	0.083
	(0.150)	(0.162)	(0.276)
Other services (excluded) = 1	0.043	0.047	0.045
	(0.202)	(0.212)	(0.208)

NOTE: Standard errors are listed in parentheses. The white race category may include respondents of Hispanic ethnicity.

Table 1, cont'd

Sample Means by Year

Regressors	1980	1990	2000
Occupations			
Managerial and professional specialty (excluded) = 1	0.216	0.253	0.264
	(0.412)	(0.435)	(0.441)
Technical sales and administrative support = 1	0.300	0.322	0.302
	(0.458)	(0.467)	(0.459)
Service = 1	0.114	0.079	0.137
	(0.318)	(0.270)	(0.344)
Farming, forestry, and fishing = 1	0.017	0.023	0.028
	(0.128)	(0.151)	(0.165)
Precision production, craft, and repair = 1	0.135	0.129	0.117
	(0.342)	(0.335)	(0.321)
Operators, fabricators, and laborers = 1	0.217	0.194	0.153
	(0.412)	(0.395)	(0.360)
Wage and education variables			
Hourly wage	14.113	15.47	17.86
	(17.937)	(20.266)	(26.443)
Less than high school = 1	0.228	0.134	0.121
	(0.419)	(0.340)	(0.326)
High school = 1	0.393	0.348	0.328
	(0.489)	(0.476)	(0.470)
College = 1	0.098	0.148	0.158
	(0.297)	(0.355)	(0.365)
Postgraduate = 1	0.084	0.079	0.084
	(0.277)	(0.269)	(0.277)

NOTE: Standard errors are listed in parentheses. The white race category may include respondents of Hispanic ethnicity.

CZs have changed over time as might be expected. For example, the amount of money invested by firms in computer hardware and software has increased almost 4 times and 12 times, respectively, between 1980 and 1990 and between 1990 and 2000, while the probability of workers using computers at work has more than doubled over both time periods. In addition, computer use at home has increased by 56 percent[3]; education levels overall have increased; the share of the CZ born in Latin America has increased more than the share born in other parts of the world; the population has aged; marriage and unionization rates have declined; and the shares of workers employed in financial activities, information, leisure and hospitality, and professional and business services occupations have all increased, as expected.

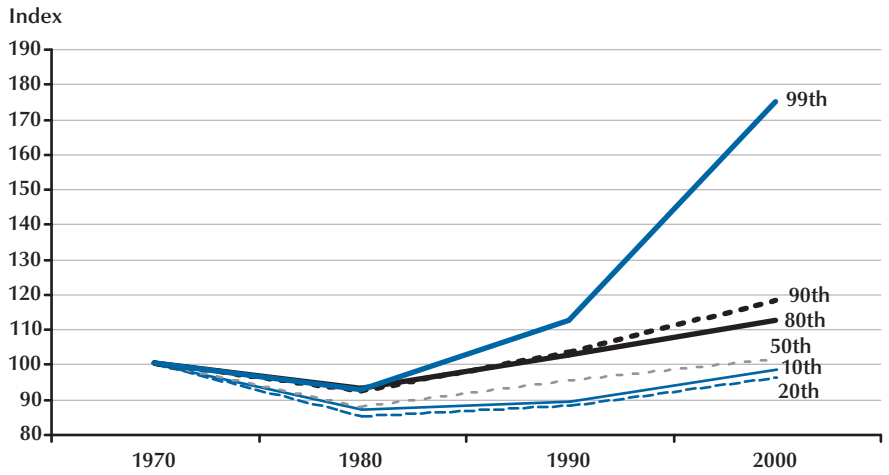
Recent investigations of the growth in real wages find the greatest growth in the upper portion of the earnings distribution.[4] Only Lemieux (2006) makes a direct link between the upper portion of the earnings distribution and the highest levels of education. Figures 2 and 3 plot normalized hourly wages by worker percentiles and education levels, respectively, to compare the data

[3] The probability of computer use at home seems high at 52 percent. Note that the Current Population Survey supplement from 1984 is used as a proxy for computer use in 1980.

[4] For example, see Guvenen and Kuruscu (2007), Lemieux (2006), Ginther and Rassier (2006), Autor, Katz, and Kearney (2006), Piketty and Saez (2003), and Topel (1997).

Figure 2

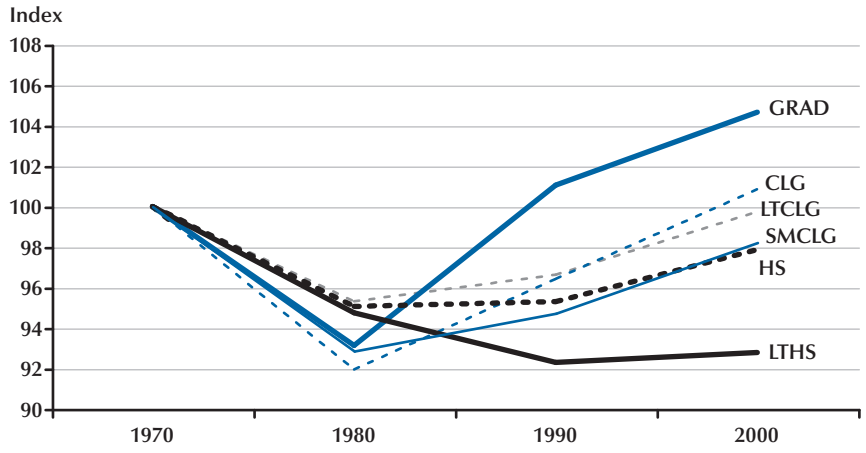
Indexed Hourly Wages Across Worker Percentiles



NOTE: Hourly wages are indexed to the value of hourly wages in 1970.

Figure 3

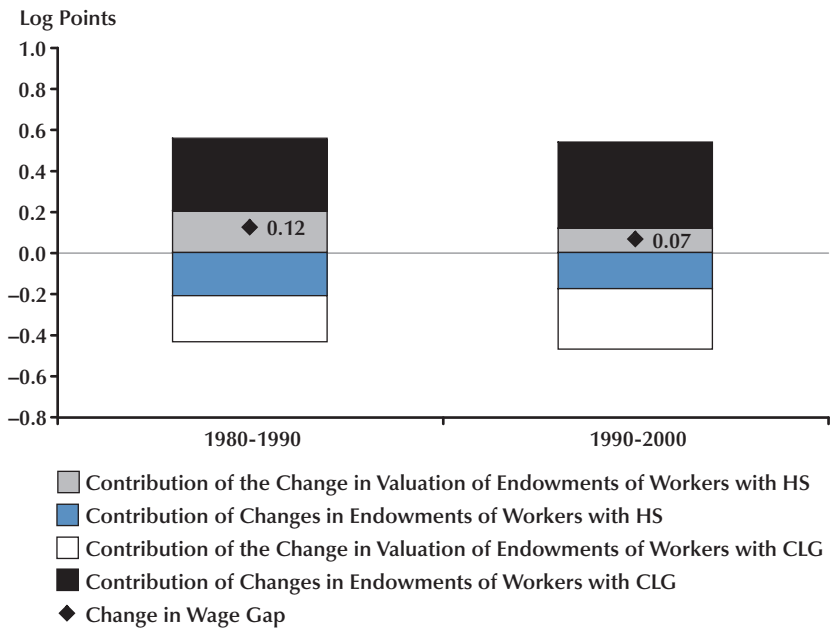
Indexed Hourly Wages Across Education Levels



NOTE: Hourly wages are indexed to the value of hourly wages in 1970. LTHS, less than high school; HS, high school; LTCLG, less than college; SMCLG, some college; CLG, college; GRAD, more than college.

Figure 4

Relative Contributions of Changes in Endowments and Valuation of Endowments to the Changing Wage Gap Between High School and College Graduates



NOTE: CLG, college; HS, high school.

used in this analysis with those in previous analyses. Figure 2 confirms that the most dramatic growth in wages between 1980 and 2000 (especially between 1990 and 2000) occurred in the upper portion of the wage distribution—among workers in the 99th percentile. Figure 3 illustrates how this growth across the wage distribution translates into growth across education levels. While the growth among workers with a postgraduate degree outpaced growth for workers of lower education levels, the wage gap between the highest and next-highest education level (postgraduate versus college) shrank slightly, while the gap between college graduates and high school graduates continued to grow through 2000.

RESULTS

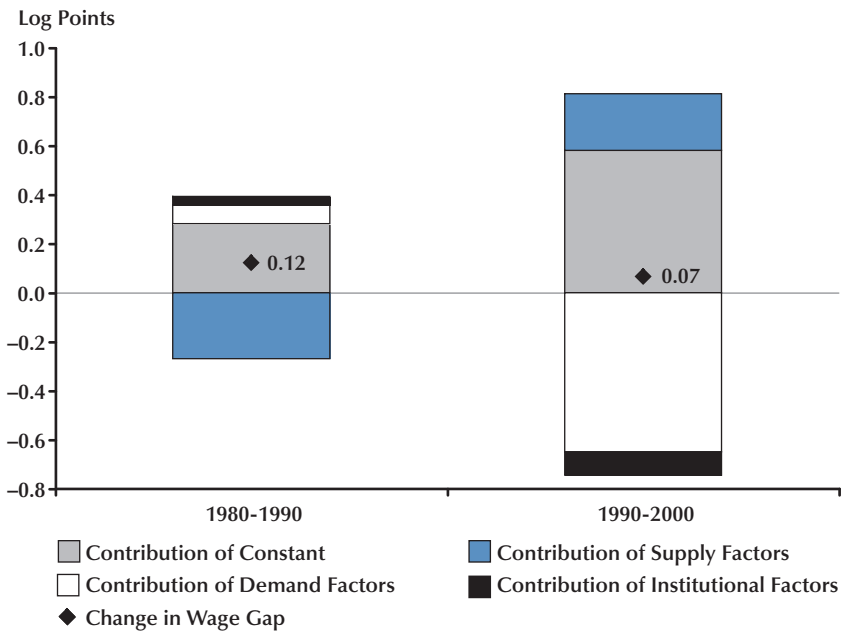
Tables 2 and 3 show the decompositions of changes in the wage gap between college and high school graduates from 1980 to 1990 (Table 2) and from 1990 to 2000 (Table 3). Figures 4 and 5 reproduce these results graphically to more easily visualize the relative contributions of changes in endowments of each educational group, contributions of changes in how those endowments translate into wages, and how different groups of regressors (e.g., supply vs. demand) compare with each other. Appendix B contains the estimated parameter coefficients for each year and each education level.

Relative Contributions of Changes in Endowments and Coefficients

Considering the endowments of workers with different education levels and how those endowments translate into wages, the relative contributions are fairly consistent across the two decades (see Figure 4). Changes in college graduates' endowments and the labor-market valuation of

Figure 5

Relative Contributions of Changes in Supply and Demand (and Other) Factors to the Changing Wage Gap Between High School and College Graduates



high school graduates' endowments (the coefficient effect) worked to increase the wage gap during both decades. However, changes in high school graduates' endowments and the changes in the labor-market valuation of college graduates' endowments exerted downward pressure on the wage gap in both decades. The implication is that, overall, both high school and college graduates were increasing their wage-enhancing characteristics (both individual and job-related) during both decades. The increasing endowments among college graduates, however, exceeded those experienced by high school graduates. As discussed in the next section, technology investments and increased computer use were the driving forces behind this greater endowment effect for college graduates.

The declining value of those characteristics (the coefficient effect) also exerted opposing pressures on the wage gap. The decline in valuation was greater among college graduates, particularly in the 1990s, which helps to explain the slowdown in the growth of the wage gap during that decade. As discussed in more detail in the next section, the driving force behind this large negative coefficient effect in the 1990s among college graduates was the significant decline in labor market return to occupational employment share.

Relative Contributions of Demand, Supply, and Institutional Factors

Figure 5 graphically illustrates the relative contributions of demand and supply factors to the changing wage gap between high school and college graduates across the 1980s and 1990s. There are some striking differences. But first we note that the significant contribution of unexplained factors in the determination of the wage gap across both decades is apparent through the size of the contribution of the constant term. An important potential component of the constant term is the change over time in the relative ability

Table 2

Decomposition of the Change in the Wage Gap Between College and High School Graduates (1980-1990)

Components of decomposition	College graduates: Contribution of changes in endowments	College graduates: Contribution of changes in coefficients	High school graduates: Contribution of changes in endowments	High school graduates: Contribution of changes in coefficients	Total (row sum)
Total (column sum)	**0.3528*** (0.0041)**	**–0.2241*** (0.0044)**	**–0.2108*** (0.0020)**	**0.2049*** (0.0022)**	**0.1229*** (0.0067)**
Demand factors	**0.2614*** (0.0041)**	**–0.0274 (0.0630)**	**–0.1479*** (0.0015)**	**–0.0103 (0.0255)**	**0.0759 (0.0681)**
Technology demand (industry hardware and software investment, probability of worker using computer at work)	0.2171*** (0.0039)	0.1323*** (0.0161)	–0.1192*** (0.0014)	0.0184*** (0.0045)	0.2486*** (0.0173)
Industry demand (industry value added, CZ employment share in worker's industry)	0.0423*** (0.0008)	0.1234*** (0.0436)	–0.0279*** (0.0004)	–0.0766*** (0.0196)	0.0612 (0.0478)
Occupation demand (CZ employment share in worker's occupation)	0.0020*** (0.0006)	–0.2831*** (0.0625)	–0.0007** (0.0003)	0.0479* (0.0253)	–0.2340*** (0.0674)
Supply factors	**0.1005*** (0.0025)**	**0.2114*** (0.0493)**	**–0.0767*** (0.0013)**	**–0.5043*** (0.0242)**	**–0.2692*** (0.0550)**
Low-skilled immigrant supply (percent of CZ population born in Central America, Caribbean, or South America)	–0.0054*** (0.0004)	–0.0005 (0.0010)	0.0067*** (0.0002)	–0.0047*** (0.0004)	–0.0040*** (0.0012)
Highly skilled immigrant supply (percent of CZ population born in non–Latin American countries)	0.0128*** (0.0003)	0.0534*** (0.0025)	0.0006*** (0.0002)	–0.0644*** (0.0013)	0.0024 (0.0029)
Technology supply (probability of worker using computer at home)	0.0120*** (0.0008)	–0.0405*** (0.0143)	–0.0001*** (0.0000)	0.0568*** (0.0045)	0.0282** (0.0150)
Other supply (percent of CZ population of worker's skill group, lagged a decade; CZ population, lagged a decade; college or university in CZ; percent of CZ labor force that is female)	0.0572*** (0.0021)	0.4465*** (0.0397)	–0.0409*** (0.0012)	–0.4448*** (0.0217)	0.0179 (0.0453)
Demographics (age, gender, race, ethnicity, marital status)	0.0239*** (0.0007)	–0.2476*** (0.0202)	–0.0429*** (0.0003)	–0.0472*** (0.0092)	–0.3137*** (0.0222)
Institutional factors CZ unemployment rate, industry extent of unionization, mobility rate of population in CZ, worker's industry, and occupation dummy variables	**–0.0091*** (0.0006)**	**–0.1283*** (0.0118)**	**0.0138*** (0.0003)**	**0.1586* (0.0072)**	**0.0350** (0.0138)**
Constant	**0.0000**	**–0.2798*** (0.0755)**	**0.0000**	**0.5610*** (0.0355)**	**0.2811*** (0.0835)**

NOTE: Standard deviations are listed in parentheses; these have been estimated using the delta method accounting for the sampling variation in the regressors; see Phillips and Park (1988), Oehlert (1992), and Jann (2008). ***, **, and * indicate significance at the 99, 95, and 90 percent confidence levels, respectively.

Table 3

Decomposition of the Change in the Wage Gap Between College and High School Graduates (1990-2000)

Components of decomposition	College graduates: Contribution of changes in endowments	College graduates: Contribution of changes in coefficients	High school graduates: Contribution of changes in endowments	High school graduates: Contribution of changes in coefficients	Total (row sum)
Total (column sum)	**0.4195*** (0.0082)**	**–0.2960*** (0.0084)**	**–0.1755*** (0.0040)**	**0.1181*** (0.0043)**	**0.0661*** (0.0131)**
Demand factors	**0.2191*** (0.0049)**	**–0.6962 (0.0606)**	**–0.0469*** (0.0017)**	**–0.1261*** (0.0266)**	**–0.6502*** (0.0663)**
Technology demand (industry hardware and software investment, probability of worker using computer at work)	0.2184*** (0.0049)	0.0397* (0.0208)	–0.0613*** (0.0016)	0.1526*** (0.0060)	0.3493*** (0.0223)
Industry demand (industry value added, CZ employment share in worker's industry)	–0.0073*** (0.0007)	–0.1502*** (0.0403)	0.0065*** (0.0004)	0.1140*** (0.0204)	–0.0370 (0.0451)
Occupation demand (CZ employment share in worker's occupation)	0.0080*** (0.0008)	–0.5857*** (0.0593)	0.0079*** (0.0004)	–0.3927*** (0.0270)	–0.9626*** (0.0652)
Supply factors	**0.2227*** (0.0079)**	**–0.0533 (0.0592)**	**–0.1717*** (0.0041)**	**0.2334*** (0.0294)**	**0.2310*** (0.0667)**
Low-skilled immigrant supply (percent of CZ population born in Central America, Caribbean, or South America)	–0.0033*** (0.0003)	0.0049*** (0.0009)	0.0072*** (0.0002)	–0.0027*** (0.0005)	0.0608*** (0.0011)
Highly skilled immigrant supply (percent of CZ population born in non–Latin American countries)	0.0182*** (0.0004)	–0.0106*** (0.0021)	–0.0157*** (0.0002)	0.0332*** (0.0011)	0.0251*** (0.0024)
Technology supply (probability of worker using computer at home)	0.1118*** (0.0073)	0.3045*** (0.0327)	–0.1240*** (0.0039)	–0.1485*** (0.0074)	0.1438*** (0.0346)
Other supply (percent of CZ population of worker's skill group, lagged a decade; CZ population, lagged a decade; college or university in CZ; percent of CZ labor force that is female)	0.0776*** (0.0018)	–0.3672*** (0.0411)	–0.0281*** (0.0008)	0.3534*** (0.0258)	0.0357 (0.0486)
Demographics (age, gender, race, ethnicity, marital status)	0.0184*** (0.0006)	0.0151 (0.0187)	–0.0111*** (0.0003)	–0.0019 (0.0105)	0.0204 (0.0215)
Institutional factors CZ unemployment rate, industry extent of unionization, mobility rate of population in CZ, worker's industry, and occupation dummy variables	**–0.0223*** (0.0008)**	**–0.0201* (0.0107)**	**0.0432*** (0.0005)**	**–0.0960*** (0.0073)**	**–0.0952*** (0.0130)**
Constant	**0.0000**	**0.4736*** (0.0798)**	**0.0000**	**0.1068*** (0.0397)**	**0.5804*** (0.0891)**

NOTE: Standard deviations are listed in parentheses; these have been estimated using the delta method (Phillips and Park, 1988), accounting for the sampling variation in the regressors (see Jann, 2008). ***, **, and * indicate significance at the 99, 95, and 90 percent confidence levels, respectively.

of college and high school graduates. Hendricks and Schoellman (2011) present evidence that a fair amount of the growth in the college wage premium can be attributed to the growth in the relative ability (or "quality") of college graduates compared with high school graduates. Such a change in quality is unmeasured and will, thus, be captured only by the constant term.

Of arguably greater interest here than the role of unmeasurables, however, is that changes in supply, demand, and institutional factors have had completely opposite effects on the wage gap across the two decades. During the 1980s, demand and institutional factors increased the wage gap, while supply factors, as a whole, exerted downward pressure on the wage gap. The opposite was true for the 1990s—supply factors increased the wage gap, while demand and institutional factors decreased it. The most dramatic reversal was among demand factors. Tables 2 and 3 provide details of the relative contributions.

Demand Factors. Consistent with the skill-biased technological change literature, the largest single contributor to the wage-gap-enhancing change in college graduates' endowments was the investment by their employers in technology and their use of computers at work, both in the 1980s and the 1990s.[5] At the same time, employers of high school graduates were investing in technology and those workers were also increasingly likely to use computers at work, but these changes were not nearly large enough to offset the growth along this dimension among college graduates, particularly in the 1990s. During the 1990s, however, the change in the use of computers at home (a supply factor) by high school graduates was the single largest contributing endowment factor exerting downward pressure on the wage gap (–0.1240). And this downward pressure slightly exceeded the upward pressure of the growing use of home computers by college graduates (0.1118). Perhaps this reflects that increased computer use allowed high school graduates to catch up in terms of computer-specific human capital, especially since home computer use by high school graduates was essentially non-existent in the 1980s.

Nonetheless, as in Krueger (1993), we find that computer use at work is rewarded more than computer use at home. For college graduates, a 10-percentage-point increase in the probability of using a computer at work translated into a 3 percent, 7 percent, and 9 percent increase in wages in 1980, 1990, and 2000, respectively (see estimation results in Appendix B). Analogous rewards were 5 percent, 6 percent, and 3 percent for high school graduates. This growing valuation of computer use at work by employers of college graduates (along with the returns workers experience from their employer's technology investments) and the declining (but still positive) valuation of computer use by employers of high school graduates explain why the coefficient effect for technology demand factors is positive for both high school and college graduates in both decades. It also illustrates what others have found: It was not only the increased use of technology among college graduates that translated into faster wage growth, but also the greater translation of technology investment and use into higher wages for college graduates that expanded the wage gap.

The boost to the wage gap from increased technology use and investment between 1980 and 1990 (0.2486) was almost completely offset by downward pressure imposed by changing occupational demand (–0.2340). Between 1990 and 2000, this downward pressure of changing occupational demand is three times larger than the continued upward pressure on the wage gap imposed by changing technology investment and use. This accounts for the bulk of the flip between the 1980s and 1990s in the direction of the contribution of demand factors. As in Autor, Katz, and Kearney (2006), we measure occupational demand as the share of employment accounted for by each occupation; the greater the share of employment in a particular occupation, the greater the demand for workers with those occupational skills. Generally, the empirical results presented here are consistent with the theoretical conclu-

[5] As in Autor, Levy, and Murnane (2003), we measure employers' investments in technology as the total spent on all computer and peripheral equipment and software. Even if new devices were introduced between the 1980s and 1990s, this aggregated measure should be reflective of the total investment.

sions drawn by Autor, Katz, and Kearney (2006) that market forces likely played an important role in the determination of the wage gap, especially during the 1990s. The downward pressure on the wage gap as a result of changing occupational demand between 1990 and 2000 came from the reduced rewards to employment in occupations dominated by college graduates (even more so than during the 1980s) and the increased rewards to employment in occupations dominated by high school graduates. While the share of jobs populated by high school and college graduates did not substantially change between 1990 and 2000, the labor market rewards of employment in those occupations did. Specifically, a 1-percentage-point increase in the CZ share of employment in a worker's occupation increased wages among high school graduates by 0.05 percent in 1990 but by 0.53 percent in 2000 (see Appendix B)—thus the relatively large negative coefficient effect in the "Occupation demand" category in Table 3 (–0.3927). At the same time, the analogous coefficient among college graduates decreased from 0.17 to –0.51, putting further downward pressure on the wage gap (–0.5857).

Autor, Levy, and Murnane (2003) conclude that technological change caused relative demand shifts favoring educated labor (also see Katz and Murphy, 1992). The results from the analysis here suggest that the rewards to that shift in demand toward educated labor were primarily flowing to college graduates through the increased use of and investment by employers in technology. This is consistent with Autor, Levy, and Murnane's (2003) conclusions that technological change caused, rather than reflected, the demand shift toward educated labor (as seen here in both the 1980s and 1990s results).

In addition, the growing rewards to high school graduates through increasing occupational share in the 1990s (as opposed to primarily through technological change) are consistent with Autor, Katz, and Kearney's (2006) evidence of a polarization of the labor market in the 1990s; the marginal productivity of manual task input (supplied by less-educated workers) is complementary with a rise in routine task input (supplied primarily by lower-cost computer capital). There is very little evidence here of this effect in the 1980s, which is, again, consistent with Autor, Katz, and Kearney's (2006) monotonic shift in occupational demand during that decade.

The relatively innocuous impact of the changing industrial employment share is consistent with the findings of Wheeler (2005) and Katz and Murphy (1992) that rising inequality within industries is more important than rising inequality between industries in explaining the growing education gap in both decades.

Supply Factors. During the 1980s, supply factors, as a whole, put downward pressure on the wage gap. The most significant supply factor driving the growing education wage gap during the 1980s was the valuation of demographics (–0.3137), most notably the valuation of demographics of college graduates (–0.2476), the largest contributor to which was age. Between 1980 and 1990, the oldest of the baby boomers were entering their 40s, with the youngest baby boomers graduating from college and entering their 20s. In addition, increasing numbers of workers with a college degree were entering the workforce (although at a decreasing rate; see Card and Lemieux, 2001). The net result, it appears, was that earlier college-educated boomers were facing significant competition as the youngest of their cohort began graduating from college, putting downward pressure on college wages—thus, the wage gap.

The largest supply factor contributing to the wage gap change during the 1990s was computer use at home (0.1438). Even though high school graduates increased their computer use slightly more than college graduates during this decade, the increased use gave a much larger boost to college graduates' wages (a 0.3045 contribution to the wage gap change) than to high school graduates' wages (a –0.1485 contribution to the wage gap change), making for a net positive contribution to the wage gap. This may be because high school graduates were increasingly less likely to apply their newly acquired computer skills on the job. This accounts for the bulk of the flip between the 1980s and 1990s in the direction of the contribution of supply factors.

Another significant supply factor change is the share of workers with the same education level (lagged) in the individual's CZ.[6] Changes in this factor were relatively unimportant in the 1980s but contributed a relatively significant share to wage gap growth in the 1990s. Changes in both the endowment and coefficient effects related to this factor contributed to its sizable contribution. First, college-educated workers became more geographically concentrated and high school graduates became less geographically concentrated (endowment changes). Second, being located in a CZ with a large share of workers with the same skill level was increasingly a bonus for college graduates but became a penalty for high school graduates—a continuation of the decline in return to this characteristic that was also seen between 1980 and 1990. This result is consistent with the finding of others, such as Giannetti (2001) and Hotchkiss, Pitts, and Robertson (2008), that once a workforce has a large enough concentration of highly skilled workers, the workers themselves benefit from the rents generated by skill complementarities. This finding also suggests that the supply effects found at an aggregate level by Card and Lemieux (2001) (fewer available college-educated workers boosts their wages) do not necessarily trickle down to the individual level; an individual college graduate captures rents from locating in a labor market with others of the same education level, *ceteris paribus*.

While Topel (1997) found that the percent of the labor force that is female did not have much impact on growing wage inequality, decomposing that supply factor into endowment and coefficient effects highlights a notable shift from the 1980s to the 1990s. Between 1980 and 1990, the coefficients on the share of the workforce that is female changed from negative (more females in the labor force put downward pressure on wages) to positive. This had the effect of raising both college and high school graduate average wages (making the college graduate coefficient effect for this regressor positive and the high school graduate coefficient effect negative).

In contrast, between 1990 and 2000, the coefficients on the percentage of the CZ labor force that is female declined for both college and high school graduates, making the impact of the change just the opposite of what occurred during the previous decade. Much has been made of highly educated women "opting out" of the labor force during the 1990s (for example, see Hotchkiss, Pitts, and Walker, 2010). If this took the form of women working fewer hours or in jobs requiring less skill, this opt-out phenomenon could be contributing to the dramatic downward pressure on the wage gap from the percent of the CZ labor force that is female.

Topel (1997) also found that immigration was not particularly important for explaining growing wage inequality during the 1980s. We also found this to be the case for both the 1980s and the 1990s, likely because of the small fraction of the workforce made up by immigrants.

Institutional Factors. Changes in factors that we categorize as institutional increased the wage gap between 1980 and 1990 but decreased the wage gap during the 1990s. Institutional factors are those characteristics that describe the labor market and differ from the characteristics brought to the labor market by employers and workers. Card and DiNardo (2002) point to declining unionization as a major contributor to the growing wage gap between education groups. However, in addition to being a relatively minor contributor in this analysis, controlling for other wage-determining factors at the individual level results in the contribution of unionization (both the change in unionization rates and the change in return to unionization) exerting *downward* pressure on the wage gap during both decades, although the impact of that downward pressure was much smaller in the 1990s.

Changes in mobility worked in favor of high school graduate wages in the 1990s but had little impact on the changing wage gap in the 1980s. In 1990, there appears to have been a wage penalty for working in a CZ with high levels of mobility for both college and high school graduates, although the penalty was greater among college graduates. In 2000, that penalty became larger for college graduates but became a bonus for high

[6] Details that follow relating to the categories of "Other supply" and "Institutional factors" are not reported individually in Tables 2 and 3 but can be easily constructed using the means in Table 1 and the parameter estimates in Appendix B.

school graduates—hence the fairly significant downward pressure on the wage gap. It was also in 2000 that the return to being employed in an occupation with a high employment share increased significantly for high school graduates. The increasing return to mobility may reflect a degree of flexibility among high school graduates that allowed them to take advantage of increased demand for the occupations in which they are employed.

One might also expect to find lower average wages in CZs with an abundance of slack labor. The positive coefficient on the unemployment rate, however, is consistent with the presence of sticky wages (for an example, see Gottschalk, 2005). For any given equilibrium level of wages (characterized by all of the other regressors included in the estimation), the higher the unemployment rate, the higher the observed wage in that labor market is likely to be (the higher the observed wage is above the equilibrium wage). This is not an estimated causal relationship between unemployment and the wage level, but rather, merely a cross-sectional correlation holding all other labor market characteristics constant. The result does not invalidate the frequently replicated negative relationship between wage growth and the unemployment rate (for example, as seen in Aaronson and Sullivan, 2001).

SENSITIVITY ANALYSIS

One of the main points of the analysis in this paper is that focusing on just one potential contributor to the change in the education wage gap over time runs the risk of biasing the conclusions. This section illustrates just how sensitive the decomposition is to exclusions of various regressors. Three alternative specifications are estimated: (i) excluding the industry and occupation dummy variables, (ii) excluding the technology demand variables, and (iii) excluding all CZ-level regressors. The resulting changes in the decompositions across specifications are illustrated (along with the baseline decompositions) in Figures 6 and 7.

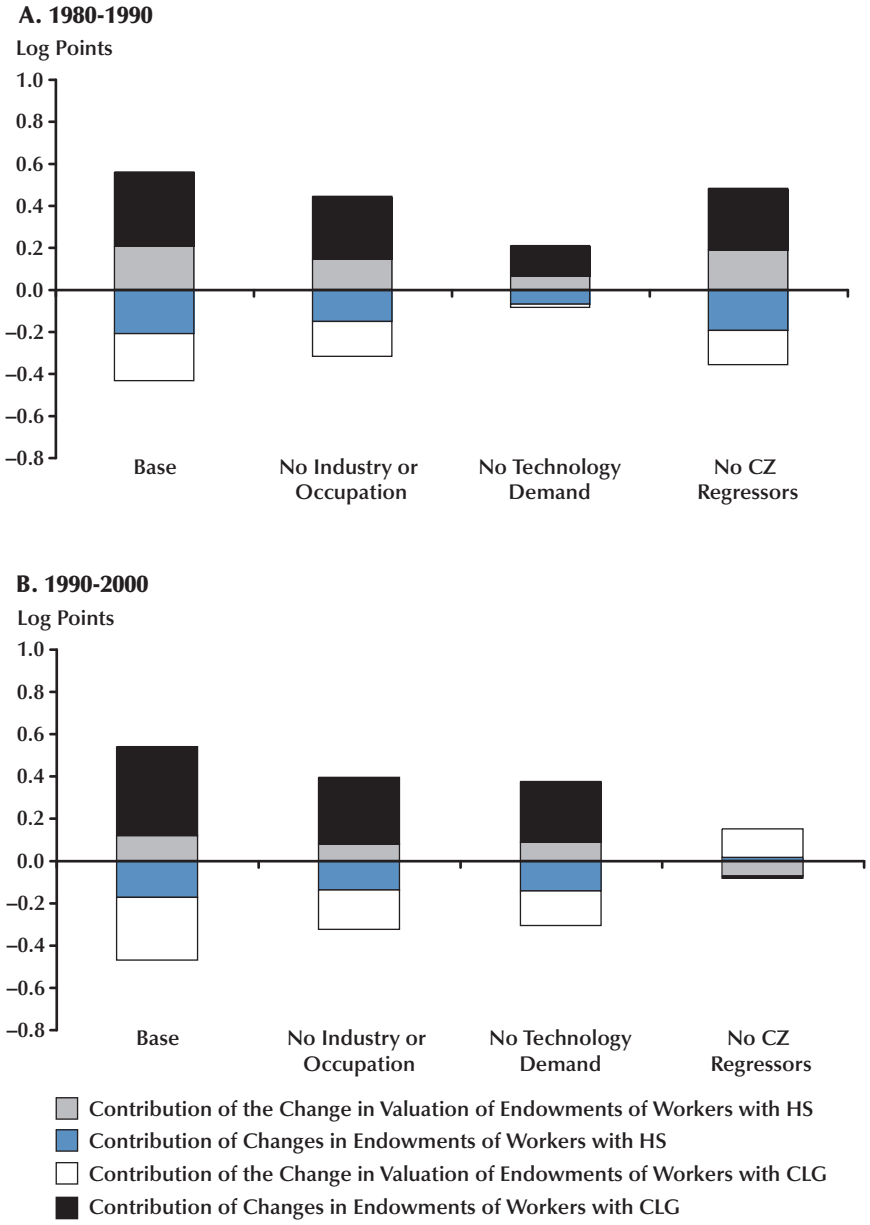
With only one exception, none of the different specifications altered the relative contributions of changes in endowments and coefficients (Figure 6). Although the individual terms in the decomposition differed in size from the base specification, the relative contributions reflected in the baseline decomposition remained unchanged. The exception was the relative contributions of endowments and coefficients to the observed change in the wage gap between 1990 and 2000 when CZ-level variables are excluded from the analysis; the overwhelming source of the difference in this case is the increase in the unexplained portion of the valuation of endowments among college graduates—as reflected in the estimate of the intercept term. Removing technology demand from the estimation considerably decreased the changes in the contribution of endowments of both college graduates and high school graduates in both decades. The most dramatic effect was the reduction of the change in the valuation of endowments of college graduates during the 1980s.

Removing the industry and occupation dummy variables primarily affected the contribution of the industry and occupation CZ employment shares. This change in contribution manifested itself through an increase in the relative contribution of the CZ occupation employment share to a growing wage gap. This, in turn, reduced the growing advantage of high school graduates over college graduates in demand for their occupational fields. This pattern of change was the same across both decades (see in Figure 6 how the gray section of the second bar in both panels is smaller than the gray section of the first bar). The implication is that excluding occupation and industry fixed effects would have resulted in underestimating the complementary role that demand for high school graduates' skills (as measured by demand for occupational shares of high school graduates) played as the demand for technological skills increased.

The motivation for removing the technology demand factors was to determine which other factors would take the place of this dominant influence on the change in the wage gap. The primary effect of removing technology demand factors was an increase in the relative contribution of supply factors to the growing wage gap. This occurred primarily through an increased

Figure 6

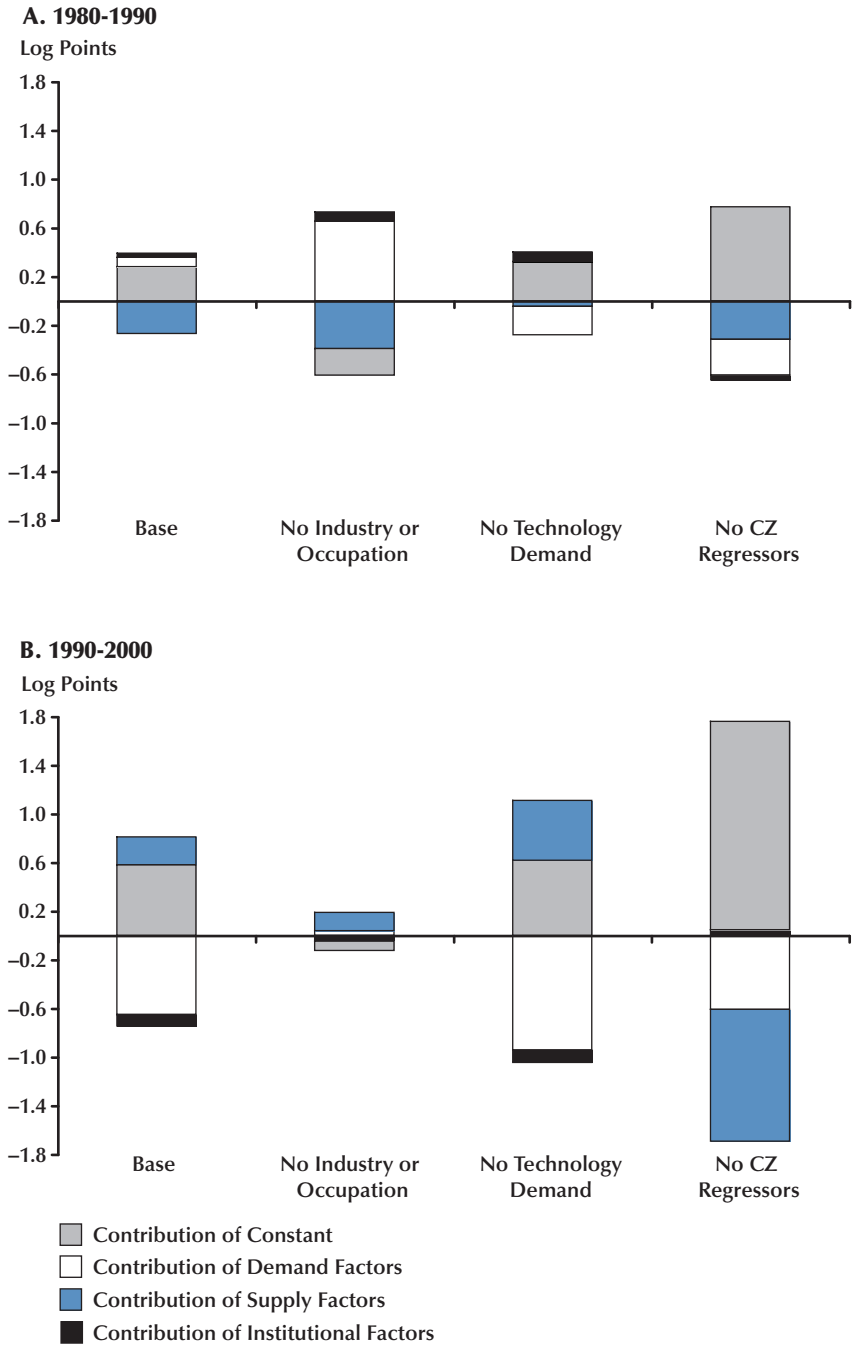
Sensitivity Analysis: Contributions of Changes in Endowments and Coefficients Across Different Specifications



NOTE: CLG, college; CZ, commuting zone; HS, high school.

Figure 7

Sensitivity Analysis: Contributions of Changes in Demand, Supply, and Institutional Factors Across Different Specifications



NOTE: CZ, commuting zone.

contribution of technology supply (see how the blue section of the third bar of both panels in Figure 7 compares with the blue section of the first bar; it is a smaller negative contribution in panel A and a larger positive contribution in panel B).

Excluding CZ-level regressors had a differential effect in the 1980s and 1990s. In the absence of CZ regressors in the 1980s, the contribution of demand factors to the wage gap increase (the white section of the fourth bar in panel A of Figure 7) was reduced significantly, compared with the baseline, primarily through the reduced importance of industry and occupation employment shares. In the 1990s, the contribution of supply factors to the wage gap increase (the blue section of the fourth bar in panel B of Figure 7) was significantly reduced, compared with the baseline, mainly through the reduced importance of home computer use. The increase in the contribution of the intercept was largest in this specification across both decades.

For the most part, with the exception of excluding CZ-level regressors, the relative contributions of changes in endowments and coefficients remain the same across different specifications. However, the relative contributions of supply and demand factors do change in fairly significant ways. Of course, those changes are partially dependent on the categorization of regressors into supply and demand influences, but once there is agreement on that point, it is clearly important to include as many measures as possible of potential influence. It is particularly important to include measures of geographic differences across education groups and time when trying to identify primary contributors to the changing wage gap.[7]

CONCLUSIONS AND IMPLICATIONS

The analysis in this paper provides a thorough reduced-form investigation of the relative contributions of supply and demand factors to the growing wage gap between high school and college graduates during the 1980s and the 1990s. Most importantly, the analysis identifies the mechanism through which technological change boosted wages of both groups of workers in each decade. Specifically, in both decades wage gains from increased demand for college graduates flowed through their increased use of technology (and technological investments by their employers), rather than from merely an increase in demand for educated workers. However, the main rewards from technology to high school graduates flowed through increased demand for their particular skills (which are theorized to be complementary to technological advancements), rather than through the use of technology itself. These results provide empirical evidence in support of the theoretical arguments of Autor, Katz, and Kearney (2006) that the labor market of the 1990s experienced a polarization; the marginal productivity of manual-task input (supplied by less-educated workers) is complemented by a rise in routine-task input (supplied primarily by lower-cost computer capital).

In general, the results are mostly consistent with those in the previous literature; however, the individual-level analysis in this paper provides an advantage over some aggregate analyses. For example, whereas Card and Lemieux (2001) found that reduced aggregate supply boosted wages of college graduates, the results here indicate that the marginal effect of a growing concentration of college graduates (increased supply in a geographic area) had an increasingly positive impact on college wages over the two decades, consistent with evidence of rents generated by skill complementarities, as found by Giannetti (2001).

The analysis also demonstrated that supply and demand wage-determining factors had opposite effects in the growth of the wage gap during the 1980s and 1990s; however, changes in endowments of workers with college degrees were largely responsible for the increasing wage gap in both decades. Consistent with the skill-biased technological change literature, technological change—the increased investments in technology and

[7] Others have documented the importance of geography on wage differences and wage growth. For example, see Bartik (1991), DuMond, Hirsch, and Macpherson (1999), Hirsch (2005), Easton (2006), Hirsch, König, and Möller (2009), and Black, Kolesnikova, and Taylor (2009).

computer use by workers (both college and high school graduates)—was the single largest contributing endowment change that affected the wage gap across both decades, even after controlling for as many other demand, supply, and institutional factors as possible.

In addition to contributing to our overall understanding of the dynamics of the wage gap between workers of different education levels during the 1980s and 1990s and the roles that supply and demand factors in each decade played in determining the wage gap, the analysis in this paper provides an even more general lesson. Focusing on only one factor in a complicated market process runs the risk of losing perspective of that factor's relative importance in the determination process or missing the impact of that factor's interaction with other market forces. The sensitivity analysis demonstrated the importance of including as many measures of potential influence as possible when trying to identify sources in the changing wage gap, particularly measures of geographic differences across education groups.

REFERENCES

Aaronson, Daniel and Sullivan, Daniel G. "Cross-State Evidence on the Relationship Between Unemployment and Wage Growth." *Chicago Fed Letter*, May 2001, No. 165; www.chicagofed.org/digital_assets/publications/chicago_fed_letter/2001/cflmay2001_165.pdf.

Autor, David and Dorn, David. "Inequality and Specialization: The Growth of Low-Skill Service Jobs in the United States." Unpublished manuscript, Department of Economics, Massachusetts Institute of Technology, July 20, 2008; www.carloalberto.org/files/autor-dorn-specialization-services-july21-2008.pdf.

Autor, David H.; Katz, Lawrence F. and Kearney, Melissa S. "The Polarization of the U.S. Labor Market." NBER Working Paper No. 11986, National Bureau of Economic Research, January 2006; www.nber.org/papers/w11986.pdf.

Autor, David H.; Levy, Frank and Murnane, Richard J. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics*, November 2003, *118*(4), pp. 1279-333.

Bartik, Timothy J. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991.

Black, Dan A.; Kolesnikova, Natalia A. and Taylor, Lowell. "Earnings Functions When Wages and Prices Vary by Location." *Journal of Labor Economics*, 2009, *27*(1), pp. 21-47.

Card, Daniel E. and DiNardo, John E. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*, October 2002, *20*(4), pp. 733-83.

Card, Daniel E. and Lemieux, Thomas. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *Quarterly Journal of Economics*, May 2001, *116*(2), pp. 705-46.

Crifo, Patricia. "Skill Supply and Biased Technical Change." *Labour Economics*, October 2008, *15*(5), pp. 812-30.

DuMond, J. Michael; Hirsch, Barry T. and Macpherson, David A. "Wage Differentials Across Labor Markets and Workers: Does Cost of Living Matter?" *Economic Inquiry*, October 1999, *37*(4), pp. 577-98.

Easton, Todd. "Metropolitan Wage Levels of Less-Educated Workers: 1986-1999." *Industrial Relations*, April 2006, *45*(2), pp. 119-46.

Giannetti, Mariassunta. "Skill Complementarities and Migration Decisions." *Labour*, 2001, *15*(1), pp. 1-31.

Ginther, Donna K. and Rassier, Dylan G. "Work Activities as Unmeasured Skills: Do Work Activities Explain Earnings Inequality?" Presented at the Society of Labor Economists Eleventh Annual Meeting, May 2006.

Goldin, Claudia D. and Katz, Lawrence F. *The Race Between Education and Technology*. Cambridge, MA: Harvard University Press, 2008.

Gottschalk, Peter. "Downward Nominal Wage Flexibility: Real or Measurement Error?" *Review of Economics and Statistics*, November 2005, *87*(3), pp. 556-68.

Guvenen, Fatih and Kuruscu, Burhanettin. "A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000." NBER Working Paper No. 13095, National Bureau of Economic Research, May 2007; www.nber.org/papers/w13095.pdf?new_window=1.

Hendricks, Lutz and Schoellman Todd. "Student Abilities During the Expansion of U.S. Education, 1950-2000." Unpublished manuscript, Department of Economics, University of North Carolina–Chapel Hill, 2011; www.lhendricks.org/Research/paper_abil.pdf.

Hirsch, Barry T. "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." *Industrial and Labor Relations Review*, July 2005, *58*(4), pp. 525-51.

Hirsch, Boris; König, Marion and Möller, Joachim. "Is There a Gap in the Gap? Regional Differences in the Gender Pay Gap." IZA Discussion Paper No. 4231, Institute for the Study of Labor, June 2009; http://ftp.iza.org/dp4231.pdf.

Hotchkiss, Julie L.; Pitts, M. Melinda and Robertson, John C. "The Push-Pull Effects of the Information Technology Boom and Bust: Insight from Matched Employer-Employee Data." *Economic Development Quarterly*, August 2008, *22*(3), pp. 200-12.

Hotchkiss, Julie L.; Pitts, M. Melinda and Walker, Mary Beth. "Assessing the Impact of Education and Marriage on Labor Market Exit Decisions of Women." Working Paper No 2010-2, Federal Reserve Bank of Atlanta, February 2010; www.frbatlanta.org/documents/pubs/wp/wp1002.pdf.

Ingram, Beth F. and Neumann, George R. "The Returns to Skill." *Labour Economics*, February 2006, *13*(1), pp. 35-59.

Jann, Ben. "The Blinder-Oaxaca Decomposition for Linear Regression Models." *Stata Journal*, 2008, *8*(4), pp. 453-79.

Katz , Lawrence F. and Murphy, Kevin M. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 1992, *107*(1), pp. 35-78.

Kranz, Daniel F. "Why Has Wage Inequality Increased More in the USA than in Europe? An Empirical Investigation of the Demand and Supply of Skill." *Applied Economics*, April 2006, *38*(7), pp. 771-88.

Krueger, Alan B. "How Computers Have Changed the Wage Structure: Evidence from Microdata: 1984-1989." *Quarterly Journal of Economics*, February 1993, *108*(1), pp. 33-60.

Lemieux, Thomas. "Post-Secondary Education and Increasing Wage Inequality." NBER Working Paper No. 12077, National Bureau of Economic Research, 2006; www.nber.org/papers/w12077.pdf.

Moretti, Enrico. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics*, July-August 2004, *121*(1-2), pp. 175-212.

Nervis, Allan. *The State Universities and Democracy*. Urbana, IL: University of Illinois Press, 1962.

Oehlert, Gary W. "A Note on the Delta Method." *American Statistician*, February 1992, *46*(1), pp. 27-9.

Phillips, Peter C.B. and Park, Joon Y. "On the Formulation of Wald Tests of Nonlinear Restrictions." *Econometrica*, September 1988, *56*(5), pp, 1065-83.

Piketty, Thomas and Saez, Emmanuel. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics*, February 2003, *118*(1), pp. 1-39.

Tolbert, Charles M. and Sizer, Molly. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update." ERS Staff Paper No. 9614, Economic Research Service, Rural Economy Division, U.S. Department of Agriculture; September 1996; www.ers.usda.gov/briefing/rurality/lmacz/LMACZ1990.pdf.

Topel, Robert H. "Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality." *Journal of Economic Perspectives*, Spring 1997, *11*(2), 55-74.

Wellington, Alison J. "Changes in the Male/Female Wage Gap, 1976-1985." *Journal of Human Resources*, Spring 1993, *28*(2), pp. 383-411.

Wheeler, Christopher H. "Evidence on Wage Inequality, Worker Education, and Technology." Federal Reserve Bank of St. Louis *Review*, May/June 2005, *87*(3), pp. 375-93; http://research.stlouisfed.org/publications/review/05/05/Wheeler.pdf.

APPENDIX A

Variable Descriptions and Data Sources

Variable Descriptions and Data Sources Overview

The data used for the analysis in this paper are from a number of sources. The primary data source is the Integrated Public Use Microdata Series (IPUMS) and was obtained from the Minnesota Population Center at the University of Minnesota. Commuting zone (CZ)-level regressors are constructed using the individual-level data in the IPUMS. In particular, average demographics and labor market characteristics are constructed based on CZs with data from the IPUMS.

Data for industry-level investment in technology are obtained from the National Income and Product Accounts (NIPA). Data for industry value added, designed to capture overall product demand—and thus worker demand—also come from NIPA.

Data for computer use at work and home and unionization by industry are obtained from the Current Population Survey (CPS). Data from the National Center for Education Statistics (NCES) are used to obtain zip codes for all institutions of higher learning, which are then mapped onto CZs. Detailed descriptions and sources of all variables used in the analysis are provided in Table A1.

Table A1

Variable Descriptions and Construction and Data Sources

Dependent variable: Individual log hourly wage. All dollar values are deflated to 2000 values using the personal consumption expenditures (PCE) chain-type price deflator. All regressors, even if CZ (*k*)- or industry (*j*)-specific, are measured at the individual level (*i*). See the next section for information related to construction of CZs.

Regressors	Description	Data source
Demand factors		
$computer_j$ $software_j$	Industry-specific (three-digit NAICS) dollar investment in high-tech equipment and software; millions of dollars. Expected to capture industry demand for technologically astute workers.	NIPA
$comwork_i$	Measures an individual's use of a computer at work. A reduced-form OLS model is estimated using the CPS to determine a person's probability of using a computer at work. The parameter estimates are then applied to the IPUMS to obtain a predicted probability of an individual using a computer at work. The earliest CPS survey of computer and Internet use was conducted in 1984; this supplement is used as a proxy for computer use in 1980.	CPS Computer and Internet Use Supplement
VA_j	Industry-specific value added, measured as the dollar value of output minus the value of intermediate inputs. Expected to capture total derived demand for workers.	NIPA
$EmplShare_{kj}$ $EmplShare_{ki}$	Share of total workforce in CZ *k* that is employed in the worker's industry *j* (occupation *i*). Expected to capture local labor market demand for employment across industries.	IPUMS
Supply factors		
$comhome_i$	Measures an individual's use of a computer at home. A reduced-form OLS model is estimated using the CPS to determine a person's probability of using a computer at home. The parameter estimates are then applied to the IPUMS to obtain a predicted probability of an individual using a computer at home. The earliest CPS survey of computer and Internet use was conducted in 1984; this supplement is used as a proxy for computer use in 1980.	CPS Computer and Internet Use Supplement

NOTE: CPS, Current Population Survey; IPUMS, Integrated Public Use Microdata Series; NAICS, North American Industry Classification System; NCES, National Center for Education Statistics; NIPA, National Income and Product Accounts; OLS, ordinary least squares.

Table A1, cont'd

Variable Descriptions and Construction and Data Sources

Regressors	Description	Data source
Supply factors, cont'd		
$flfper_k$	This is a measure of the percent of the CZ labor force that is female. Others have concluded that female workers are lower-paid substitutes for low-skilled men, and their presence could drive down wages of low-skilled workers.	IPUMS
$perskillX_k$	Percent of the CZ population that is of skill group X (e.g., high school only, college graduate).	IPUMS
$mCenCaribbSouthAm_k$ $EuropeAsia_k$ $mothernonNAm_k$	Percent of CZ population born in Central America, the Caribbean, or South America; percent of CZ population born in Europe or Asia; and percent of CZ population born in other non–North American areas (e.g., African countries, Arctic regions). Immigrant shares are expected to capture the effect of immigration on local wage determination.	IPUMS
$schldummy_k$	Dummy variable set equal to 1 if CZ has at least one college or university that offers a bachelor's degree. Zip codes of schools (obtained from the NCES) were mapped onto the CZ. Other work has used a dummy variable indicating the presence of a land-grant university only (see Nervis, 1962, and Moretti, 2004).	U.S. Department of Education NCES
Demographics		
age_i	Age of individual (and its squared value).	IPUMS
$female_i$	Dummy variable set equal to 1 if individual is female.	IPUMS
$white_i$ (excluded)	Dummy variable set equal to 1 if white. All race variables are constructed from IPUMS variable *race*. May include respondents of Hispanic ethnicity.	IPUMS
$black_i$ $asian_i$ $otherrace_i$	Dummy variables set equal to 1 if black, Asian (Chinese, Japanese, or other Asian or Pacific Islander), or any other race.	IPUMS
$marriedsp_i$	Dummy variable set equal to 1 if married with a spouse present.	IPUMS
Institutional factors		
$urate_k$	CZ-level unemployment rate; constructed using individual labor force data from the IPUMS. Expected to capture current local labor market conditions.	IPUMS
$union_j$	Industry-specific unionization percent. Expected to measure the degree of noncompetitive wage-setting mechanisms present in worker's industry.	CPS, Annual Social and Economic Supplement
$mfluidity_k$	Percent of the CZ that lived in a different state five years ago. Expected to capture the mobility of workers in the local labor market; a greater degree of mobility makes a labor market more competitive.	IPUMS
ind_j occ_m	Dummy variables for broad industry (j) and occupation (m) classifications. Expected to capture occupation- and industry-specific determinants of wages not otherwise controlled for.	IPUMS

Method for Assigning Commuting Zones to Individuals

Data on CZs are extracted from the IPUMS. The original data were constructed by Tolbert and Sizer (1996), who used 1990 Census data on journey-to-work county commuting flows to construct 741 CZs (clusters of counties with strong commuting ties). We use the same (1990) definition of CZs for all analysis years (1980, 1990, and 2000) for a consistent definition of a labor market area throughout our analysis.

A CZ is assigned to an individual in the sample by matching CZs to either *public use microdata* (PUMA) for 1990 and 2000 or a similarly defined *county group* (CNTYGRP) for 1980. Because each PUMA or CNTYGRP can contain multiple CZs, we use the following method to assign each observation in a PUMA or CNTYGRP to a CZ (similar to the method used by Autor and Dorn, 2008):

(i) The CZ dataset is merged into the IPUMS dataset that contains PUMA/CNTYGRP by county federal information processing standards (FIPS) codes.

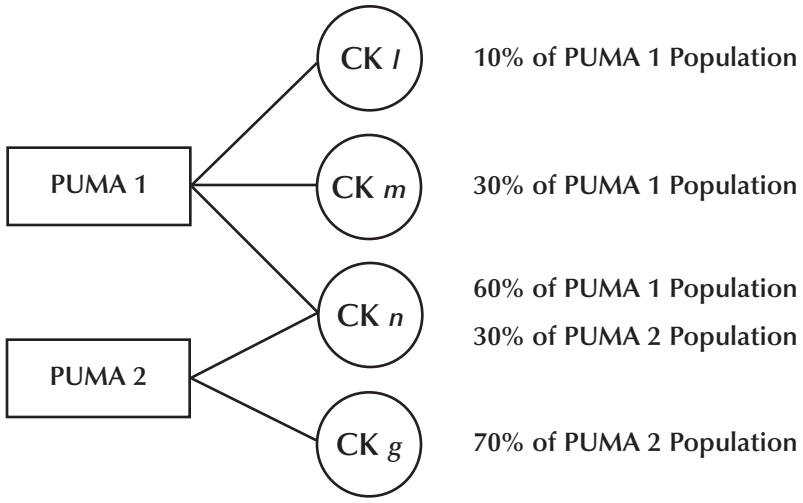
(ii) Depending on the year, between 68 and 82 percent of CZs are matched exclusively to one PUMA or CNTYGRP. In 2000, 1,677 of the 2,052 PUMAs (82%) match to a single CZ. In 1990, 1,348 of the 1,726 PUMAs (78%) match to a single CZ. In 1980, 788 of the 1,154 county groups (68%) match to a single CZ.

(iii) When the match between CZ and PUMA/CNTYGRP is not exclusive, a random assignment strategy is used to distribute the PUMA/CNTYGRP population across the appropriate CZs.

   (a) Population weights are created for each CZ within a PUMA or CNTYGRP. The weights are equal to the share of the PUMA or CNTYGRP population in each CZ.

   (b) Each IPUMS observation within a PUMA or CNTYGRP is assigned a value from a uniform random variable distribution.

   (c) Each person is then assigned a CZ based on the CZ's population share weight and the person's uniform distribution value. For example (see diagram below), if PUMA 1's population is distributed across CZ *l* (10 percent), CZ *m* (30 percent), and CZ *n* (60 percent), then individuals from PUMA 1 with a uniform draw between 0 and 0.10 will be assigned to CZ *l*; individuals with a draw between 0.10 and 0.40 will be assigned to CZ *m*; and the remaining population is assigned to CZ *n*.



APPENDIX B

OLS Parameter Estimates of Log Wage Equations by Education and Year

	1980		1990		2000	
Regressors	High school	College	High school	College	High school	College
Demand factors						
Computer investment ($ billions)	0.0064***	0.0155***	–0.0410***	–0.0521***	–0.0054***	–0.0071***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.000)	(0.000)
Software investment ($ billions)	0.1336***	0.1847***	0.1262***	0.1469***	0.0169***	0.0182***
	(0.002)	(0.005)	(0.001)	(0.002)	(0.000)	(0.000)
Probability of computer use at work	0.5058***	0.2708***	0.5629***	0.6902***	0.2806***	0.9438***
	(0.017)	(0.030)	(0.014)	(0.023)	(0.015)	(0.027)
Industry value added ($ trillions)	0.2848***	0.3761***	0.3976***	0.4947***	0.0309***	0.0775***
	(0.010)	(0.019)	(0.006)	(0.009)	(0.005)	(0.006)
Industry employment share	0.5047***	0.7450***	0.5758***	0.8754***	0.5337***	0.7913***
	(0.017)	(0.040)	(0.017)	(0.035)	(0.018)	(0.033)
Occupation employment share	0.1071***	0.5031***	0.0477**	0.1705***	0.5281***	–0.5083***
	(0.022)	(0.054)	(0.023)	(0.049)	(0.024)	(0.048)
Supply factors						
Probability of computer use at home	0.2569***	0.2889***	0.1344***	0.2287***	0.4537***	0.6486***
	(0.006)	(0.014)	(0.008)	(0.016)	(0.014)	(0.042)
Female share of CZ labor force	–0.4929***	–0.7329***	0.7022***	0.4925***	0.3309***	–0.3157***
	(0.031)	(0.066)	(0.036)	(0.066)	(0.039)	(0.063)
CZ population (lagged, millions)	2.1561***	1.7968***	0.3476***	0.4070***	0.0097***	0.0105***
	(0.050)	(0.099)	(0.009)	(0.013)	(0.000)	(0.001)
Share of CZ with high school diploma (lagged)	0.1485***	—	0.1168***	—	–0.2650***	—
	(0.009)		(0.009)		(0.011)	
Share of CZ with college degree (lagged)	—	1.5064***	—	0.7076***	—	1.1516***
		(0.080)		(0.051)		(0.031)
Share of CZ born in Latin America or the Caribbean (lagged)†	–1.0993***	–0.3310***	–0.6587***	–0.3697***	–0.5268***	–0.1895***
	(0.038)	(0.075)	(0.017)	(0.028)	(0.013)	(0.020)
Share of CZ born in Europe or Asia (lagged)	1.5807***	0.5972***	3.1575***	1.7684***	1.8391***	1.3050***
	(0.034)	(0.067)	(0.031)	(0.044)	(0.025)	(0.033)
Share of CZ born in other non–North American countries (lagged)	–1.5441***	0.0038	0.1026*	1.1482***	1.0670***	1.8213***
	(0.060)	(0.117)	(0.057)	(0.084)	(0.053)	(0.069)
University or college in CZ	0.0521***	0.0236***	0.0241***	0.0227***	0.0143***	0.0173***
	(0.003)	(0.009)	(0.002)	(0.005)	(0.002)	(0.005)
Demographics						
Age	0.0378***	0.0595***	0.0386***	0.0462***	0.0363***	0.0498***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Age squared	–0.0003***	–0.0005***	–0.0003***	–0.0004***	–0.0003***	–0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	–0.3478***	–0.2489***	–0.3615***	–0.3038***	–0.2961***	–0.3064***
	(0.002)	(0.005)	(0.002)	(0.004)	(0.002)	(0.003)

NOTE: Standard deviations are listed in parentheses. Because of the potential endogeneity, CZ population, shares of CZ with different education degrees, and immigration shares are all lagged a decade (e.g., the 1990 value is used in the 2000 regression). ***, **, and * indicate significance at the 99, 95, and 90 percent confidence levels, respectively. †The excluded immigration share in CZ is North American. ‡The excluded race is white. §The excluded industry is other services. ¶The excluded occupation is managerial and profession specialty. OLS, ordinary least squares.

APPENDIX B, cont'd

OLS Parameter Estimates of Log Wage Equations by Education and Year

	1980		1990		2000	
Regressors	High school	College	High school	College	High school	College
Demographics, cont'd						
Black‡	–0.0083***	0.0360***	–0.0098***	0.0530***	0.0132***	0.0625***
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
Asian	–0.0330***	–0.0710***	–0.0891***	–0.0546***	–0.0778***	–0.0283***
	(0.004)	(0.006)	(0.004)	(0.004)	(0.003)	(0.004)
Other race	–0.0122**	–0.0559***	–0.0014	–0.0256**	–0.0501***	–0.0696***
	(0.005)	(0.015)	(0.005)	(0.012)	(0.002)	(0.004)
Married with spouse present	0.0621***	0.0773***	0.0882***	0.1066***	0.1022***	0.0951***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Institutional factors						
CZ unemployment rate	1.6755***	1.9837***	0.4808***	1.6019***	1.2773***	1.4432***
	(0.039)	(0.079)	(0.046)	(0.077)	(0.051)	(0.083)
Unionization	0.3190***	0.2478***	0.3788***	0.0914***	0.3080***	–0.2073***
	(0.005)	(0.011)	(0.005)	(0.009)	(0.005)	(0.008)
Mobility rate of CZ population	0.2707***	0.1205***	–0.0130	–0.0836***	0.1233***	–0.2159***
	(0.010)	(0.022)	(0.010)	(0.018)	(0.011)	(0.018)
Industries§						
Natural resources and mining	0.2234***	0.3976***	0.2513***	0.4682***	0.1216***	0.3037***
	(0.004)	(0.010)	(0.004)	(0.007)	(0.004)	(0.009)
Construction	0.1631***	0.3168***	0.0912***	0.2777***	0.0569***	0.2745***
	(0.003)	(0.008)	(0.003)	(0.006)	(0.003)	(0.006)
Manufacturing	0.1711***	0.3838***	0.1904***	0.4394***	0.1399***	0.4362***
	(0.003)	(0.007)	(0.003)	(0.005)	(0.002)	(0.005)
Transportation and utilities	0.2098***	0.3789***	0.2132***	0.4274***	0.1261***	0.4730***
	(0.004)	(0.009)	(0.003)	(0.006)	(0.003)	(0.006)
Wholesale trade	0.0125**	0.1855***	–0.0018	0.2669***	0.0411***	0.3399***
	(0.005)	(0.011)	(0.005)	(0.007)	(0.004)	(0.006)
Retail trade	–0.0577***	0.0473***	–0.1532***	0.0105*	–0.0473***	0.1941***
	(0.004)	(0.008)	(0.003)	(0.006)	(0.003)	(0.006)
Financial activities	–0.0062	0.0898***	–0.1528***	–0.0249***	0.1255***	0.3701***
	(0.004)	(0.009)	(0.004)	(0.007)	(0.004)	(0.005)
Information	0.2319***	0.2340***	0.0809***	0.2947***	0.0197***	0.2444***
	(0.005)	(0.010)	(0.004)	(0.006)	(0.004)	(0.006)
Professional and business services	0.1079***	0.2565***	–0.0607***	0.0331***	–0.0032	0.2176***
	(0.003)	(0.007)	(0.003)	(0.006)	(0.003)	(0.005)
Education and health services	0.0561***	0.2219***	0.0057**	0.1961***	–0.0355***	0.2039***
	(0.003)	(0.007)	(0.003)	(0.005)	(0.003)	(0.005)
Leisure and hospitality	0.0836***	0.2187***	0.1184***	0.2656***	0.0108***	0.1771***
	(0.004)	(0.010)	(0.004)	(0.007)	(0.003)	(0.006)

NOTE: Standard deviations are listed in parentheses. Because of the potential endogeneity, CZ population, shares of CZ with different education degrees, and immigration shares are all lagged a decade (e.g., the 1990 value is used in the 2000 regression). ***, **, and * indicate significance at the 99, 95, and 90 percent confidence levels, respectively. †The excluded immigration share in CZ is North American. ‡The excluded race is white. §The excluded industry is other services. ¶The excluded occupation is managerial and profession specialty. OLS, ordinary least squares.

APPENDIX B, cont'd

OLS Parameter Estimates of Log Wage Equations by Education and Year

	1980		1990		2000	
Regressors	High school	College	High school	College	High school	College
Occupations¶						
Technical sales and administrative support	–0.1218***	–0.1379***	–0.1353***	–0.1834***	–0.0907***	–0.2108***
	(0.002)	(0.005)	(0.002)	(0.004)	(0.002)	(0.004)
Service	–0.2730***	–0.2806***	–0.3066***	–0.3621***	–0.2094***	–0.4460***
	(0.004)	(0.010)	(0.004)	(0.008)	(0.004)	(0.007)
Farming, forestry, and fishing	–0.4395***	–0.4745***	–0.4565***	–0.6334***	–0.3228***	–0.6094***
	(0.006)	(0.014)	(0.004)	(0.010)	(0.004)	(0.010)
Precision production, craft, and repair	–0.0820***	–0.1667***	–0.0976***	–0.2608***	–0.0784***	–0.3255***
	(0.002)	(0.005)	(0.002)	(0.004)	(0.002)	(0.004)
Operators, fabricators, and laborers	–0.1796***	–0.3232***	–0.2115***	–0.4465***	–0.1664***	–0.5683***
	(0.003)	(0.008)	(0.003)	(0.006)	(0.003)	(0.006)
Constant	0.6925***	–0.3260***	0.1315***	–0.6058***	0.0247	–0.1322**
	(0.024)	(0.056)	(0.026)	(0.051)	(0.030)	(0.062)
Observations	1,506,546	375,090	1,610,134	684,110	1,822,896	922,376
Adjusted R^2	0.2530	0.2548	0.2589	0.2418	0.2043	0.2045

"Frictions in Financial and Labor Markets": A Summary of the 35th Annual Economic Policy Conference

Rodolfo Manuelli and Adrian Peralta-Alva

This article contains synopses of the papers presented at the 35th Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis held October 21-22, 2010. The conference theme was "Frictions in Financial and Labor Markets." Leading participants in this field presented their research and commentary.

Federal Reserve Bank of St. Louis *Review*, July/August 2011, *93*(4), pp. 273-92.

The Thirty-Fifth Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis was held October 21-22, 2010. The papers presented at the conference covered a variety of approaches and topics within the general theme of frictions in financial and labor markets. One group of papers directly addresses the question of the impact of frictions in financial markets—defined as a departure from the complete market, perfectly competitive Arrow-Debreu equilibrium—on economic performance. In "Quantifying the Impact of Financial Development on Economic Development," Greenwood, Sanchez, and Wang study the impact of increases in the relative (to the rest of the economy) efficiency of financial intermediaries in output and total factor productivity (TFP). For a calibrated version of their model they conclude that financial frictions can account for large changes in output and measured TFP. A somewhat different conclusion is reached by Midrigan and Xu in "Finance and Misallocation: Evidence from Plant-Level Data." In that paper, the authors study a different financial friction—a borrowing limit—and find that when the model must match the observed distribution of the growth rate of the output of individual firms, the contribution of market imperfections to TFP is rather small.

In "Middlemen in Limit-Order Markets," Jovanovic and Menkveld analyze the role of middlemen in asset markets who are assumed to have superior information and, hence, potentially improve the allocation of resources as they can "direct" each asset to its best use. They find that, depending on the distribution of information of potential asset traders, the presence of middlemen can either increase or decrease efficiency. They also confront the model with data that are consistent with the introduction of middlemen but their results are ambiguous. The last paper that most directly discusses the role of financial frictions is "Financial Markets and Unemployment," by Monacelli, Quadrini, and Trigari. They study a situation in which firms and workers bargain for wages but the total surplus—the object to be divided—decreases in relation to the amount of debt carried by the firm. They show that, in

Rodolfo Manuelli is a professor and director of graduate studies in the department of economics at Washington University in St. Louis and a visiting scholar at the Federal Reserve Bank of St. Louis. Adrian Peralta-Alva is a senior economist at the Federal Reserve Bank of St. Louis. Author-provided papers delivered at the conference can be found at http://research.stlouisfed.org/conferences/policyconf/35program.html.

© 2011, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

response to a positive productivity shock, firms will choose to borrow more since this lowers their current surplus and thus the wage demands of their workers.

A second set of papers looks at the role of search frictions in labor and goods markets. In "Joint-Search Theory: New Opportunities and New Frictions," Guler, Guvenen, and Violante consider the employment-search problem of a couple. They show that, in the absence of a market that permits perfect risk-sharing, location decisions and employment decisions are related and, due to the costs of separation from one's partner, some workers would reject job offers that appear to be above their reservation wage. A similar idea—searching for a price in this case—drives the price dispersion results in "Equilibrium Price Dispersion and Rigidity: A New Monetarist Approach," by Head, Liu, Menzio, and Wright. They show that when individuals differ in their ability to search for the lowest price, (i) the optimal pricing policy of a firm involves periods of price stickiness (when average prices are changing) and (ii) price dispersion may occur in equilibrium even when there is no inflation.

Finally, two papers deal with the effect of frictions on income distribution. In "Intergenerational Redistribution in the Great Recession," Glover, Heathcote, Krueger, and Ríos-Rull study how a recession—not unlike the recent one in the United States—influences the welfare of different generations. They show that (i) asset prices will likely fall more than warranted by fundamentals and (ii) this has a negative effect on relatively older households. At the other end of the spectrum, younger households see their labor income drop but are able to purchase some assets (from the older generations) at bargain prices. Their welfare does not decrease as much as that of the older cohorts and, in some cases, it may increase. In "Social Security, Benefit Claiming, and Labor Force Participation: A Quantitative General Equilibrium Approach," İmrohoroğlu and Kitao consider the quantitative implications of three alternative Social Security reforms: reductions in benefits, increases in normal retirement age, and increases in the earliest retirement age. They find that these proposals will have long-run positive effects. Even though this is not the focus of their paper, it seems that such changes could hurt current retirees and individuals close to retirement.

Overall, the research at the conference succeeded in focusing attention of academic economists and policymakers alike on the role of frictions in the economy.

INDIVIDUAL CONTRIBUTIONS

In the following sections we briefly describe the essential elements of the individual conference papers. The aim is not to provide a complete description of the environment and results, but rather to convey the major methodological and factual contributions of the research. In some cases, our analysis goes beyond the conference version of the paper and tries to draw inferences relevant for policymakers.

QUANTIFYING THE IMPACT OF FINANCIAL DEVELOPMENT ON ECONOMIC DEVELOPMENT

How do changes in the productivity of the financial intermediary sector affect the level of output? This is the question studied by Greenwood, Sanchez, and Wang. In their model, financial intermediaries exist because they better detect misreporting by firms.

The main theoretical finding is that increases in the efficiency of financial intermediaries (relative to the rest of the economy) increase output through two channels. First, the set of firms that receive funding shrinks and includes more high-productivity firms. Second, the size of the loan that each firm receives—which in all cases falls short of the perfect information level—increases for the most-productive firms and decreases for the least-productive.

Greenwood, Sanchez, and Wang calibrate their model to match the relevant data for the United States and use it to predict the impact on a given country's output level if it adopted Luxembourg's financial system. The basic model does an excellent job matching the cross-country

evidence and suggests that large gains are possible with efficiency-increasing developments in the financial intermediary sector.

The Model

The production function of a firm is given by

$$y = x\theta k^{\alpha}\ell^{1-\alpha},$$

where x is an aggregate productivity shock (common to all firms), θ (which can take two values $\theta_1 < \theta_2$) is a firm-specific shock, and k and ℓ are, respectively, capital and labor. The key assumption is that although the type of firm is public knowledge (i.e., the set $\tau = (\theta_1,\theta_2)$ is known), the particular realization of θ is not.

Since firms need to borrow funds to purchase capital, they contract with a financial intermediary. The distinguishing feature of this intermediary is its access to a monitoring technology. Greenwood, Sanchez, and Wang describe the properties of this monitoring technology in terms of the probability of detecting a cheater—a firm that claims to have received a low-productivity draw, θ_1, when in fact it enjoys high productivity, θ_2—as a function

$$P_{ij}\left(\ell_{mj}, k, z\right),$$

where $P_{ij}(\ell_{mj},k,z)$ is the probability of detecting fraud when a firm announces that its productivity is θ_j when in fact it is θ_i. This probability increases the number of workers assigned to monitoring, ℓ_{mj}, as well as the productivity of the financial sector, z. It decreases with the size of the loan, capturing the idea that larger (and more complex) loans are more difficult to monitor.

Greenwood, Sanchez, and Wang analyze the optimal contract between financial intermediaries and firms. They find that

(i) the set of projects that is financed—that is, the set of $\tau = (\theta_1,\theta_2)$ that gets loans—shrinks as the relative efficiency of the financial sector increases (i.e., as z/x increases). Moreover, this "shrinkage" is associated with increases in the average efficiency of the funded firms;

(ii) as the efficiency of the financial intermediaries increase, some low-return firms fail to obtain funding, while high-return firms receive larger loans. This increases output and measured TFP; and

(iii) increases in financial intermediary efficiency result in higher wages.

Quantitative Results

Greenwood, Sanchez, and Wang consider alternative measures of intermediation costs (interest rate spreads, the capital-to-output ratio, or overhead costs) and calibrate the model using U.S. data on firm size distribution and output per worker. Then they use the model to ask some counterfactual questions about the United States—a mature economy in which increases in productivity in finance match overall increases in productivity—and Taiwan—a developing country that has experienced a significant increase in the *relative* productivity of its financial sector. They find that

(i) in the United States, about 30 percent of the growth in output per capita in the 1974-2004 period (from $22,352 to $41,208) can be attributed to productivity improvements in the financial sector, z. Stated differently, had the level of productivity of financial intermediaries remained at its 1974 level, output per capita would have grown from $22,352 to $33,656. The difference is accounted for by the banking sector; and

(ii) in Taiwan, over the same period about 50 percent of the increase in output was due to improvements in z.

Greenwood, Sanchez, and Wang also use the model to understand the contribution of changes in the efficiency of finance on the cross-sectional distribution of output levels. To this end, they assume that the model holds; in addition, with data on interest rate spreads and output per capita, they estimate, for each country j, the levels of aggregate productivity, x^j, and financial sector productivity, z^j, that are consistent with the evidence. Since there is no obvious real-world analog of the parameter z, they regress the value of z^j for country j on a measure of financial develop-

ment (the ratio of private credit to gross domestic product [GDP]) and find that the correlation is high. This suggests that their identification procedure captures actual changes in efficiency.

With a parameterized model for a sample of over 40 countries, they find that the United States has the highest level of productivity outside the financial sector (i.e., the highest level of x^j), while Luxembourg has the highest level of financial sector efficiency (highest level of z^j). Then they ask the following question: "By how much would the output per capita in country *j* increase if instead of its own (estimated) value of the efficiency of the financial sector, z^j, it had the level of financial efficiency enjoyed by Luxembourg?" Their findings include the following:

(i) World output would rise by 65 percent by moving all countries to the best financial practices.

(ii) Dispersion in cross-country output would fall by about 19 percentage points, from 77 percent to 58 percent.

(iii) Measured increases in world TFP would exceed 17 percent.

(iv) The countries that would benefit the most from this switch would see increases in output per capita of over 130 percent. Turkey, Uganda, Brazil, and Nigeria are in this category.

(v) The countries that would benefit the least would still experience significant increases in output that exceed 10 percent. This group includes New Zealand, Finland, Austria, and Ireland.

(vi) With no technological change in the financial sector, U.S. GDP would have grown at an annual rate of 1.4 percent instead of the observed 2.0 percent rate.

A careful analysis of the robustness of these results to alternative measures of intermediation costs is provided in the paper.

The paper also identifies the mechanism through which changes in the efficiency of the financial sector affect output. Greenwood, Sanchez, and Wang define the degree of distortion as the average value of the excess expected return over the cost of capital. High values of this indicator are associated with low investment and a source of inefficiency that has the flavor of "money left on the table": Some projects with high expected returns are not financed (and hence never implemented). Their findings include the following:

(i) The average distortion level ranges from a high of 49.8 percent (Uganda) to a low of 4.6 percent (Luxembourg).

(ii) The coefficient of variation across plants in a country can be high; it ranges from 32.7 percent (Uganda) to 1.94 percent (Luxembourg). This dispersion corresponds to differences across firms in the shadow price of capital, which implies that, relative to the first best, the country is in the interior of its production possibilities frontier.

(iii) The average (world) distortion level is 23.4 percent, with an average coefficient of variation of 14.6 percent. If all countries adopted Luxembourg's financial efficiency, the mean distortion drops to 2.6 percent and the average standard deviation to 1.1 percent, a very significant decrease.

The authors discuss alternative identification strategies, including the role of internal finance and robustness checks for the specification of the production function (constant elasticity of substitution instead of Cobb-Douglas), as well as unmeasured investment in intangible capital. They find that, in all cases, the differences in productivity of the financial intermediary sector account for a significant fraction of the differences in output per capita. Moreover, policies that result in increases in that productivity relative to overall productivity can have large effects on output.

Conclusion

Greenwood, Sanchez, and Wang find that finance matters for productivity and that changes in the efficiency of monitoring—a key function of financial intermediaries—can have a large impact on output. For policymakers the model

illustrates the dangers of adopting policies that, in some sense, result in lower levels of efficiency in the financial sector, as well as the benefits associated with promoting the adoption of best practices.

FINANCE AND MISALLOCATION: EVIDENCE FROM PLANT-LEVEL DATA

Midrigan and Xu study the theoretical and quantitative effects of capital market imperfections on measured aggregate TFP. The imperfections take the form of a borrowing limit that depends on a firm's asset position. The model is rich enough to match a large number of moments of the distribution of the output growth rate of individual firms; the major finding is that financial repression does not have a large impact on measured TFP.

No Entry and No Exit

Midrigan and Xu study two different versions of their basic economy. In the first, all individuals must operate a technology (i.e., they are entrepreneurs) and there is a fixed supply of labor. In the second model, individuals can choose to be either entrepreneurs or workers, depending on their managerial skills.

The problem faced by an entrepreneur who chooses investment, savings, and consumption to maximize expected utility is

$$\max E\left[\sum_{t=0}^{\infty}\beta^t U\left(C_{it}\right)\right]$$

subject to

$$Y_{it} = A_{it}\left(L_{it}^{\alpha}K_{it}^{1-\alpha}\right)^{\eta},$$
$$WL_{it} + K_{it} \le \lambda B_{it},$$
$$C_{it} + B_{it+1} = Y_{it} + (1-\delta)K_{it} + (1+r)\left(B_{it} - WL_{it} - K_{it}\right).$$

The first constraint is simply the production function, and A_{it} is the specific technology shock that follows some exogenous stochastic process. The second constraint is the financing constraint. It says that the expenditure on inputs—labor input WL_{it} and capital input K_{it}—cannot exceed a multiple λ of previous-period saving, B_{it}. In this formulation, $\lambda = 1$ corresponds to no borrowing: All expenditures must be financed using previous saving, while $\lambda = \infty$ captures perfect capital markets.

The optimal decision rules in this economy are

$$F_L\left(L_{it}, K_{it}\right) = \left(1+\tilde{r}_{it}\right)W,$$
$$F_K\left(L_{it}, K_{it}\right) = \tilde{r}_{it} + \delta,$$

where

$$\tilde{r}_{it} = r + \mu_{it},$$

where μ_{it} is the shadow value (Lagrange multiplier) of the financing constraint. Thus, in this model, the financing constraint effectively implies that different firms face different effective prices for inputs and, hence, that the economy must be in the interior of the production possibilities set. Improvements in the environment that decrease the financing constraint (i.e., that lower the Lagrange multiplier, μ_{it}) reduce the cross-sectional dispersion of input prices and, in principle, will bring the economy closer to the frontier.

Midrigan and Xu argue that the key driver of inefficiency in this setting is the standard deviation of the specific productivity shock, A_{it}. The larger the variability, the greater the need of external sources to finance expansion and, hence, the more important any frictions in financial markets become.

Quantitative Results

Midrigan and Xu use data from a large dataset of firms in Korea and Colombia to estimate the parameters of the model. The key stochastic process is the individual productivity level, A_{it}. They require that the estimated process be such that the model's predictions for the distribution of growth rates of individual firms match the data. Their major finding is that if they force the model to account for the skewness of growth rates, individual productivity processes do not vary much. This, in turn, implies that the cost of financial imperfections is not very large. To see this, consider the extreme case of a constant level of A_{it}. In this scenario, perhaps after some initial period,

firms can accumulate enough assets to purchase the optimal level of inputs and, hence, financial repression cannot have any lasting effects.

Given the estimated process for individual productivity, Midrigan and Xu pick the parameter λ to match the debt-to-GDP ratio for Korea, Colombia, and a country such as the United States. They find that financial restrictions cannot account for the large estimated differences in TFP. To be precise, they estimate that U.S. TFP is 1.3 percent lower than it would have been without frictions, while in Korea—which has a relatively developed financial system—the loss is 3.6 percent, and in Colombia—which has more restrictions on financial intermediation—the loss is 5.2 percent. Since standard estimates imply that the TFP gap between poor countries and the United States is about 60 percent, the model implies that imperfect capital markets explain only 4 percent of that 60 percent.

Entry and Exit

In a second version of the basic model, Midrigan and Xu allow for entry and exit. They assume that there is some exogenous "death rate" (and a similar birth rate) and that individuals can choose, in every period, whether to be an entrepreneur or a worker. Thus, the one-period profit of an entrepreneur is

$$\pi(B_{it}, A_{it}) = \max_{K,L} A_{it} F(L,K) - (1+r)WL - (r+\delta)K,$$

subject to

$$WL + K \leq \lambda B_{it}.$$

The individual-agent problem is then

$$\max E\left[\sum_{t=0}^{\infty} \beta^t U(C_{it})\right]$$

subject to

$$C_{it} + B_{it+1} = (1+r)B_{it} + \max\left[\pi(B_{it}, A_{it}), W\right],$$

where the last term indicates that the individual will choose to be a worker if $W > \pi(B_{it}, A_{it})$ and an entrepreneur otherwise.

As Midrigan and Xu note, when new potential entrepreneurs are born it is important to determine the correlation, if any, between their entrepreneurial skills, A_{it}, and their endowment of saving, B_{it}. In the quantitative exercise, the authors consider two possible relationships between these two variables.

In this environment, financial frictions can affect productivity along two margins. First, as in the model with no entry and no exit, capital market imperfections imply that each firm faces a potentially different shadow price of capital, which creates inefficiencies. Second, the same imperfections can influence the "identity" of the individuals who become entrepreneurs. The potential cost includes both low-productivity individuals who eventually manage a firm and high-productivity individuals who, for lack of access to credit markets, choose to become workers.

Quantitative Results

Midrigan and Xu initially assume that individuals are born with no wealth. Thus, a potentially new entrepreneur has $B_{it} = 0$ at birth. As before, the authors choose the stochastic process for productivity to match the distribution of growth rates and find that the losses associated with financial constraints are much larger than in the previous case. For the United States the loss increases from 1.3 percent to 4.4 percent, for Korea from 3.6 percent to 10.6 percent, and for Colombia from 5.2 percent to 13.1 percent. Even though these estimates fall short of explaining the bulk of estimated differences, they suggest that capital market imperfections can have a significant impact on output. Interestingly, almost all the losses correspond to misallocation of inputs across firms and very few to misallocation of entrants into entrepreneurship.

One source of inefficiency in this case is the binding constraint on new high-productivity firms. In the absence of financing constraints, these firms would be "born large." However, the lack of access to credit prevents them from reaching a high scale initially. This loss of potential output reduces measured TFP. However, the counterpart of this scenario is that these productive but "asset-poor" firms grow very rapidly:

They start small but, since they are very productive, they grow fast. When $B_{it} = 0$, the model grossly overpredicts the growth rate of young firms and, hence, this is not a good identification assumption.

As an alternative, Midrigan and Xu assume that initial wealth is positively related to the potential demand for inputs. They justify this assumption as a proxy for the existence of venture capitalists who can obtain a signal about A_{it} and provide start-up funds to the entrepreneur. They choose the correlation to match the growth rate of small firms. With this new parameterization they find that the loss of TFP is very similar to that of the model with no entry and no exit.

Conclusion

In a dynamic model of entrepreneurship with the requirement that the endogenous distribution of the growth rates of individual firms' output and the aggregate debt-to-GDP ratio must match the data, the major finding is that capital market imperfections can explain only a small fraction of the measured differences in TFP between rich and financially developed economies and poor—and typically financially constrained—economies.

MIDDLEMEN IN LIMIT-ORDER MARKETS

Jovanovic and Menkveld study the allocative efficiency implications of financial intermediaries. They develop a series of theoretical models that are useful in understanding the role of a special type of financial intermediary: middlemen. They model middlemen as reducing frictions since they have access to better information than buyers and sellers in asset markets. They show that (i) the introduction of middlemen can either improve or reduce welfare (efficiency) and (ii) the outcome depends on the information structure. When adverse selection is not important, middlemen reduce efficiency, whereas in environments in which asymmetric information about common values creates selection problems, middlemen can improve efficiency by revealing information.

The authors analyze the impact of a trading system that they view as facilitating the introduction of middlemen. Their findings are mixed. Consistent with their theory, middlemen reduce bid-ask spreads but their introduction decreases trading volume. The welfare consequences of this experiment are ambiguous and more work in this area is necessary.

The Model

Jovanovic and Menkveld study a classic asset-allocation problem: One individual—the buyer—with private valuation for the asset equal to $x + z$ can trade with another individual—the seller—with private valuation given by $y + z$. In this setting, x and y are independently drawn from some distribution, denoted by F, and their values are known only to the individual parties. On the other hand, the common component z may or may not be known to both buyers and sellers. In this simple environment, first-best efficiency requires that the asset be held by whoever has the higher private valuation. Thus, if $x > y$ the seller should keep the asset, while if $x \leq y$ the asset should be transferred to the buyer.

Jovanovic and Menkveld study a sequence of environments that differ in terms of their information structure. The simplest case is one in which neither party knows the value of z and there are no middlemen. Even in this case, the outcome depends on the allocation of monopoly rights.

Consider the case in which the seller "comes to the market first"; then the seller can post a bid interpreted as the price at which he or she is willing to sell the asset. Normalizing the expected value of the common component equal to zero, the seller's payoff is

$$\Pi^s = \max_p p\left(1 - F(p)\right) + F(p)x.$$

The interpretation is simple: If the price p is announced, the buyer will accept the bid only if $y \geq p$, which happens with probability $1 - F(p)$. Thus, in the case of a sale, the seller receives $p(1 - F(p))$. If the buyer's valuation is below p—an event with probability $F(p)$—the seller keeps the asset and enjoys utility $F(p)x$.

Next let us look at the opposite case: The buyer comes to the market first and posts a price

Table 1

Summary Results

Regime	No. of Trades	Welfare	Spread
First best	T	W	0
No middlemen	$0.5 \times T$	$0.94 \times W$	0.5
Middlemen	$0.6 \times T$	$0.85 \times W$	0.25

at which he or she is willing to purchase the asset. In this case, the seller's payoff is

$$\Pi^s = \max(y-p)F + (p)(1-F(p)) \times 0.$$

What is the role of an intermediary—a middleman in financial jargon—in this market? One view, noted by Jovanovic and Menkveld, is that high-frequency trading programs have an information-processing advantage over individuals. Thus, one way to model this advantage is to assume that the high-frequency middleman trader knows the common value z. Since a middleman is just an intermediary, the authors assume that he or she has no private valuation for the good—that is, if the middleman ends up with the asset, the payoff is just z.

The authors study a competitive version in which there is free entry into the middleman market. This guarantees that middlemen make zero profits in equilibrium. The middleman makes a bid to the seller and then, when the buyer arrives, posts an asking price. Even though the buyer does not know the common component z, it can be inferred from the middleman's bid. Jovanovic and Menkveld show that the asking price (by the middleman) is

$$p^a = z + u,$$

where

$$u = \arg\max s(1-F(s)),$$

while the middleman's bid to the seller is

$$p^b = z + \pi,$$

where

$$\pi = \max s(1-F(s)).$$

Note that in the presence of middlemen both buyers and sellers "learn" the actual value of the common component z and that this property does not depend on the particular properties of the distribution of z.

It is natural to ask whether the presence of middlemen improves the efficiency of financial markets. To this end, the authors compute an example (although the results are much more general); the results (which can be further generalized) are summarized in Table 1.

In this version of the model, the presence of middlemen unambiguously decreases welfare. It is well known from the work of Myerson and Satterthwaite (1983) that no mechanism can attain the first-best level of welfare. In this case, when buyers and sellers are symmetrically informed about the common component of the value of the asset and there are no middlemen, the level of welfare is about 94 percent of the first best. The volume of trade is only half of what would be observed in the first best and this reveals the source of the inefficiency: There are instances in which the buyer has a higher valuation for the asset and the price announced by the seller is too high and, hence, no trade takes place. Finally, there is a significant spread (defined as the average of the prices posted by sellers and buyers) in this case. Enter the middlemen (the last row of Table 1). In this case, more trades are executed (about 60 percent of the first-best volume) and the spreads decrease (from 0.5 to 0.25); nevertheless, welfare is lower. Thus, the simple evidence of more competition—more volume and lower spreads—does not imply higher efficiency. The reason is simple: In some cases, middlemen fail to execute both sides of the trade and they end up holding the asset. Since their valuation is lower than either the buyer's or the seller's, this is a source of welfare losses. In this environment, the conclusion is unambiguous: Middlemen reduce welfare.

What happens when buyers and sellers are asymmetrically informed about z? In this case, it is possible to show that, in the absence of middlemen, the volume of trade converges to zero as the variance of the common component grows without bound. The intuition for this is the no-

trade theorem of Milgrom and Stokey (1982). If the payoffs to buyers and sellers were to be renormalized by dividing by the standard deviation of the common shock—an innocuous transformation given risk neutrality—those payoffs would be as follows:

$$\text{Buyer: } \frac{x}{\sigma} + z,$$

$$\text{Seller: } \frac{y}{\sigma} + z.$$

Then as $\sigma \to \infty$, the valuations converge to z—that is, the valuations are common to both traders—and if the private information of the buyers and sellers is just a signal of the true z, the no-trade theorem implies that there will be no trade in equilibrium.

Enter the middlemen again. In this case, the middlemen act just as in the previous case. Even if the traders initially were asymmetrically informed, the bid and ask prices posted by the middleman reveal the true value of z. This eliminates the asymmetry of information about the common component, which alleviates the adverse selection problem. Thus, in this case the presence of middlemen increases welfare.

Jovanovic and Menkveld go on to study dynamic versions and discuss mechanism design issues, but the message of their theoretical results is fairly robust: Whether middlemen improve economic efficiency depends on the traders' information structure.

Empirical Analysis

The authors analyze the advent of middlemen by exploiting the introduction of a high-frequency, trading-friendly venue as an instrument. The equity exchange Chi-X started trading Dutch index stocks on April 16, 2007. Unlike the incumbent exchange, Euronext, it did not charge traders for limit orders (i.e., posted prices), modifications, or executions; limit orders that led to execution received a rebate. The authors use data from the first 77 trading days of 2007 and 2008 to establish the treatment effect. To control for time effects they use Belgian index stocks as the untreated sample since those stocks were trading on Euronext but not on Chi-X.

They were able to identify a trader who nets out trades between the two systems. This trader has several of the characteristics of a middleman: Trades are most active when there are changes in an aggregate index of stock values, and his trading position is zero about half the time even though volume is typically high.

Jovanovic and Menkveld find that the entry of middlemen (in this case, Chi-X) is accompanied by an increase in liquidity supply and a drop in volume. The bid-ask spread did not increase for Dutch stocks but went up by 35 percent for Belgian stocks. Thus, from the perspective of the model the introduction of middlemen decreases spreads by 35 percent. The number of trades was unaffected by the entry of middlemen but volume declined.

Conclusion

The paper shows that, theoretically, the introduction of middlemen can have an ambiguous impact on welfare. The authors find some evidence of middlemen-like trading associated with the introduction of a new trading system. They find that spreads are lower after the middlemen appear, which is consistent with their theory. However, they also find that trading volume decreases and, hence, the impact on welfare is ambiguous.

FINANCIAL MARKETS AND UNEMPLOYMENT

Monacelli, Quadrini, and Trigari consider whether changes to the availability of credit may amplify unemployment fluctuations, relative to a standard business cycle model with search and matching.

The key transmission mechanism of the paper works as follows: Higher debt reduces the period-by-period surplus produced by a firm; workers and firms bargain over this total surplus, and thus higher debt allows firms to lower wages and labor costs. This idea is motivated by results from the empirical corporate finance literature. For instance, Matsa (2010) tests the premise that collective bargaining imposes a greater threat to a firm when

the firm maintains higher levels of excess liquidity. Firms thus have an incentive to use debt to shield their liquidity from workers during bargaining. He finds that union bargaining power leads firms to increase financial leverage in a quantitatively significant way. Similarly, Bronars and Deere (1991) find a positive correlation between leverage and the degree of unionization (which may affect the bargaining power of workers).

The mechanism linking financial frictions to unemployment fluctuations considered by these authors is fundamentally different from traditional channels wherein credit facilitates investment or provides additional resources to finance production input costs. The quantitative analysis of this paper suggests that the response of employment and unemployment to credit shocks is, in fact, statistically significant and of relevant magnitude. Hence, financial frictions may indeed amplify the impact of exogenous shocks on employment and output.

The Model

The economy is constituted by risk-neutral households, which can be either employed or unemployed. While unemployed, households can costlessly search for a job. Households save in two types of assets: shares of firms and bonds. Firms are created through the matching of a posted vacancy and a worker. The number of matches is determined by a standard matching function that depends on the number of unemployed and vacancies. Matches break with probability λ every period. All firms are subject to a common productivity shock, which varies exogenously through time. Lending is done by competitive intermediaries who pool a large number of loans. At every period, the following events take place (in sequential order): (i) wage bargaining, (ii) financial decisions, and (iii) default.

Each firm employs one worker, and output is thus equal to aggregate productivity, z_t. Dividends are output plus the net change in debt minus the wage bill, namely,

$$d_t = z_t - w_t - b_t + \frac{b_{t+1}}{R},$$

where R is the interest rate charged on debt. Firms maximize the expected discounted value of dividends. Firms know wages are determined by Nash bargaining. More important, firms understand that debt levels, b_t, affect the outcome of the bargaining process, and take this into account when choosing their optimal debt levels. If $w_t = g_t(b_t)$, a firm's problem can be summarized by

$$J_t(b_t) = \max_{b_{t+1}} \left\{ z_t - g_t(b_t) - b_t + \frac{b_{t+1}}{R} + \beta(1-\lambda)E_t J_{t+1}(b_{t+1}) \right\}$$

$$s.t.$$

$$\varphi_t \beta E_t J_{t+1}(b_{t+1}) \geq \frac{b_{t+1}}{1+r}.$$

The constraint of the firm's problem is the limited participation constraint. It states that the amount of debt the firm can undertake cannot be larger than a fraction of the value of the firm. The firm's credit limit is affected by an exogenous factor ϕ_t; changes in its value are called credit shocks.

Observe that the additive nature of the firm's objective implies that new debt does not depend on current wages or current debt. Therefore, all firms will choose to carry forward the same level of debt. This is analytically very convenient since there is no need to keep track of the firm's distribution of debt holdings.

Since agents are risk neutral, the interest rate is constant and given by $r = 1/\beta - 1$. The interest on corporate bonds takes into account that repayment occurs with probability λ. As markets are competitive, $R(1 - \lambda) = 1 + r$.

Firms considering entry must pay a fixed entry cost. Production starts one period after entry, and thus no labor costs are incurred on entry. In equilibrium, firms will enter the market until the entry cost is equal to the value of posting a vacancy.

An important feature of the model that simplifies the bargaining problem is that the value of participation for workers is not affected by capital income. This can be done because the impact of changes in the dividend of an individual firm is negligible for an individual worker. Hence, the current payoff of being employed is the wage. The current payoff of being unemployed is an exogenously given unemployment benefit, a. The overall return for the household is

$$\hat{H}(b_t, w_t) = w_t + \beta E_t\left[(1-\lambda)H_{t+1}(b_{t+1}) + \lambda U_{t+1}\right],$$

where the value of being employed is

$$H_t(b_t) = g_t(b_t) + \beta E_t\left[(1-\lambda)H_{t+1}(b_{t+1}) + \lambda U_{t+1}\right].$$

In turn, the value of being unemployed, given that the probability of finding a job is p_t, is recursively defined by

$$U_t = a + \beta E_t\left[p_t H_{t+1}(B_{t+1}) + (1-p_t)U_{t+1}\right].$$

As is common in this type of bargaining problems, the wage is set so that the firm and the household split the net surplus of the match,

$$S_t(b_t) = z_t - a - b_t + \frac{b_{t+1}}{R} + \beta E\left[(1-\lambda-p_t\eta)S_{t+1}(b_{t+1})\right].$$

Workers will thus capture a fraction η of this surplus.

Results

The main theoretical finding is that firms choose to carry as much debt as the constraint allows, as long as workers have strictly positive bargaining power. The rationale is as follows: If the firm increases its debt by one unit, it can pay the present value of this amount in dividends today; however, the effective cost of repayment is lower than one. Repayment is made only with probability $1-\lambda$, and, more importantly, a higher level of debt reduces the part of the surplus that eventually goes to the worker.

The response of the model to a positive productivity shock z_t is as follows: Higher productivity generates an employment expansion. Financial frictions amplify the effects of this shock because higher productivity also increases the value of the firm—and thus the amount it can borrow—and the debt it will undertake. Ultimately, higher debt reduces the cost of labor, which motivates firms to expand employment further.

The quantitative analysis starts by considering a somewhat standard parameterization of the model. Impulse response analysis illustrates that the amplification of employment fluctuations that can be generated by this model is quantitatively relevant. Finally, a structural vector autoregressive approach is used. The shocks are identified using short-term restrictions derived from the theoretical model. The impact of credit shocks on employment is indeed statistically significant. Moreover, productivity shocks lead to credit expansions.

Conclusion

Monacelli and Trigari consider a new mechanism that may amplify the effect of shocks on unemployment and the business cycle. By incurring debt, firms lower the surplus available for bargaining with workers and thus lower labor costs. Firms thus maximize their debt capacity (limited by an endogenous constraint whereby creditors can recover only an exogenous fraction of the value of the firm). Productivity shocks, in addition to their standard effects on income and employment, change the value of the firm, and thus its debt capacity and labor costs. This amplification mechanism may be of interest to policymakers as a potential factor to account for the persistently high levels of unemployment in the United States since its recent recession.

JOINT-SEARCH THEORY: NEW OPPORTUNITIES AND NEW FRICTIONS

Guler, Guvenen, and Violante study the joint job-search and location problem of a household formed by a couple that perfectly pools income. Previous models assume these decisions are made by individuals acting in isolation. The framework of analysis builds on the well-known single-agent search models of McCall (1970) and Mortensen (1970), in which the unemployed receive a wage offer from an exogenous wage distribution every period. Rejecting an offer means remaining unemployed. Once an offer is accepted by an individual, he or she will be employed at the corresponding wage for the infinite future.

The key theoretical finding is that two-agent households face new opportunities and new

frictions relative to one-agent households. New opportunities arise because one member of the couple can get income immediately without sacrificing the choice of "searching" for a better job (since the other member of the household may continue to receive offers). Further, by switching roles as better offers come along, a two-agent household may manage to climb the wage ladder even in the absence of an on-the-job search (per individual). The new frictions stem from the fact that job offers may come from different locations and couples face an additional cost when they do not live together. For certain model parameterizations, labor outcomes under joint job-search and location decisions are quantitatively very different from those of standard single-agent models.

The Model

Individuals live forever and all participate in the labor force. Time is continuous and there is no aggregate uncertainty. Single individuals maximize the expected lifetime utility from consumption. Everyone is entitled to certain consumption, b, when unemployed and receives wage offers at a constant rate, a, from a predetermined wage distribution, $F(w)$. The discount factor is r. In the standard single-agent case, the moment an offer is accepted income is forever fixed at such wage rate. The optimal strategy is characterized by a unique reservation wage such that any offer above such reservation level will be accepted, and rejected otherwise.

Members of a two-agent household can be in different states and are faced with different options in each. First, both members may be unemployed and choose to remain so. Alternatively, they can both accept their offers and stay with the associated wages forever. The most interesting case occurs when one member of the household is currently employed and the other is unemployed. The corresponding value functions U (both unemployed), T (both employed), and Ω (one employed and one unemployed) are defined as follows:

$$rU = u(2b) + 2\alpha \int \max\{\Omega(w) - U, 0\} dF(w)$$

$$rT(w_1, w_2) = u(w_1 + w_2)$$

$$r\Omega(w_1) = u(w_1 + b)$$

$$+\alpha \int \max \begin{Bmatrix} T(w_1, w_2) - \Omega(w_1), \\ \Omega(w_2) - \Omega(w_1), 0 \end{Bmatrix} dF(w_2).$$

Since time is continuous, the probability of a simultaneous arrival of offers when both agents are unemployed is zero and is thus ignored in the definition of U. Couples with only one employed partner can do the following: The unemployed partner accepts a job offer and, simultaneously, the previously employed partner quits and starts looking for new offers (the gain of the couple is then $\Omega(w_2) - \Omega(w_1)$). This "breadwinner cycle" strategy is obviously not available to one-agent households.

Results

If agents are risk neutral, the symmetry of the model causes the couple to behave exactly as two independent single agents. For a joint labor search to make any difference agents must be risk averse. Risk aversion causes an individual agent within the couple to accept offers that a single-agent household would reject because accepting an offer does not prevent the couple from continuing to search. Similarly, if one partner is working, the unemployed member may reject offers that single individuals would accept. A breadwinner cycle does emerge in equilibrium. Partners alternate between who works and who searches, depending on the offers received by each. When faced with the same job offer sequence, single-agent households simply accept a job and then never quit. In the long run, the wages of both individuals in the two-agent household are higher under joint search due to the breadwinner cycle.

The authors consider modifications to the key model assumptions to understand the robustness of their results. One interesting modification allows employed agents to receive job offers (on-the-job search). If the rate of arrival of offers is the same for unemployed and employed individuals, then the model with dual earners yields the same solution as the single-agent case, even with risk-

averse agents. The breadwinner cycle may also disappear if agents have access to risk-free borrowing and lending and face "loose" debt constraints. The basic intuition for this scenario is that borrowing effectively substitutes for the consumption smoothing provided within the two-agent household, making it irrelevant. Each partner can implement search strategies independent of the other partner's actions. Hence, they act again as in the single-agent case.

In the next step of the analysis, the authors extend the baseline model to study the problem of choosing among multiple locations. For single agents this extension is straightforward and only requires adjusting reservation wages for the cost of moving. In contrast, if the couple dislikes living in different locations, then optimal choices may be affected in important ways. First, if the cost of living apart is paid period by period, it can easily add up to more than the one-time relocation cost paid by a single agent. Joint search thus adds new frictions. Couples now face the following possible states: (i) Both can be employed in the same location (an absorbing state), or (ii) both can be employed in different locations (assumed to be another absorbing state). (iii) One partner can be employed and the other unemployed, or (iv) both can be unemployed and searching. Locations are assumed to be symmetric and thus partners will live together if one is unemployed. Couples must now account for inside and outside offers in their value functions.

In this framework, and even under risk neutrality, two-agent households may behave differently from single-agent households. First, each member of the two-agent household will be less selective than a single-agent household—that is, one individual of a couple will accept a wage offer that the same individual would reject if single because couples face a worse wage distribution than single agents since some wage configurations are only attainable living apart. Second, tied stayers (a partner who rejects an offer he or she would accept if single) and tied movers (a partner who follows the other partner even when individual calculations would dictate otherwise) can easily emerge in equilibrium. Both possibilities involve a high cost by each individual agent of a couple compared with being single. However, these choices are optimal from the perspective of their household.

The authors then consider a set of simulations to illustrate the quantitative implications of the theory. They show there are reasonable parameterizations under which the joint-search problem does yield results that are substantially different (quantitatively) from an analysis based on single agents. Specifically, if the cost of living apart is relatively high, then the unemployment rate in a two-agent job-search model can be as high as 13 percent versus 5 percent in the single-agent model with an otherwise comparable parameterization.

Conclusion

It is conceivable that members of a two-agent household make their job and location decisions jointly. Standard models abstract from this. Guler, Guvenen, and Violante develop a model in which two-agent households pool income perfectly to study the implications of this joint decision process. They provide a thorough analysis of conditions under which a couple's behavior will differ from that of both agents making choices independently. Theory shows the insurance provided by a household partnership introduces the possibility of a breadwinner cycle. If living apart represents a flow cost for each member of the couple, then this friction may cause individuals to reject offers or to abandon jobs that they would accept or keep, correspondingly, if single.

The analysis in this paper suggests important issues, such as the design of unemployment compensation or policies aimed at the participation or attachment to the labor force, may be better studied in a framework that explicitly models joint decisions within a household.

EQUILIBRIUM PRICE DISPERSION AND RIGIDITY: A NEW MONETARIST APPROACH

Head, Liu, Menzio, and Wright study why some sellers do not adjust their prices as soon as the aggregate price level changes. This "price stickiness," which appears to be a well-established

feature of the data, plays a fundamental role in modern monetary models. In particular, it allows monetary policy to have real effects. The majority of such models either assume prices are sticky or impose an exogenous cost of changing prices.

Head et al. consider instead a framework in which firms can always adjust their prices at no cost. The model is a dynamic general equilibrium monetary version of Burdett and Judd's (1983) model. There is a single good and a continuum of identical consumers and firms. Firms post prices and households take as given a price distribution. Households can sample only a finite number of price offers.

The main message of this paper is that price dispersion may emerge in equilibrium, even without inflation. Monetary policy is neutral. Yet, the theory can match key statistical properties of prices in the U.S. data. The authors conclude that the fact that prices are sticky, per se, does not necessarily imply money is not neutral or that particular policy recommendations must be followed.

The Model

Time is discrete. As in Lagos and Wright (2005), in every period two markets open sequentially. The first is decentralized. Buyers and firms come together through a frictional matching process. Importantly, buyers are anonymous. Hence, trade can only occur with fiat money, which is supplied by the government. Households face the following possibilities: With probability a_0 they cannot buy any good in this period; with probability a_1 they can buy from exactly one firm; and with probability a_2 they are able to buy from two firms. Firms post prices. The household sees the firms and posted prices, p, from which it is possible to buy and must choose how much of the good to purchase, q_t. This household choice is constrained by the amount of money carried from the previous period, m_t. The corresponding value function is

$$V_t(m_t, p) = \max_{q_t} u(q_t) + W_t(m_t - pq_t)$$
$$s.t.$$
$$0 \le pq_t \le m_t.$$

Here, W denotes the value function of entering the second market (described below). The overall lifetime utility of households that enters this market with m_t units of money can then be written as follows:

$$U_t(m_t) = \alpha_0 W_t(m_t) + \alpha_1 \int V_t(p, m_t) dF_t(p)$$
$$+\alpha_2 \int V_t(p, m_t) d\left[1 - \left(1 - F_t(p)\right)^2\right].$$

Notice that the lowest price charged by the two firms is a random variable with distribution $1 - [1 - F_t(p)]^2$.

Firms post a nominal price, taking as given a buyer's money holdings and the price distribution, $F_t(p)$. The price to post and a production level are then chosen to maximize expected profits.

The second market is centralized and money or credit is allowed. Households work, purchase goods, and, in doing so, determine how much money they will have by the closing of this market. The associated value function is

$$W_t(m_t) = \max_{h_t, x_t, m_{t+1}} v(x_t) - h_t + \beta U_{t+1}(m_{t+1})$$
$$s.t.$$
$$w_t x_t + m_{t+1} \le w_t h_t + m_t + D_t + T_t$$
$$x_t, m_{t+1} \ge 0.$$

Here, x_t denotes consumption, h_t hours worked, and w_t hourly wages. Government transfers are denoted as T_t and firms' dividends D_t. A key assumption of the model is that preferences are quasilinear. This implies wealth effects wash out and thus all households choose to carry the same amount of money. The distribution of money is therefore degenerate.

The paper ultimately focuses on stationary monetary equilibrium whereby real money holdings, consumption, and hours worked remain constant, while prices grow at the same rate as the money supply.

Results

Profit maximization implies (i) the price distribution is continuous and (ii) its support must be an interval $[\underline{p}_t, \overline{p}_t]$ (either mass points or gaps in the support would allow an individual firm to obtain a discrete increase in profits by adjusting

its prices by a discrete amount while keeping its customer base relatively constant). In equilibrium, firms must be indifferent to posting any price on the support of the distribution. Hence, more structure is needed to derive quantitative predictions.

Consider the set of firms whose prices lie on the intersection of the new support for prices (that results from monetary injections) and the old support for prices. The authors assume that at any given period a constant fraction ρ of these firms keeps their prices fixed, while the others shift prices randomly. All firms whose prices do not belong to the new support of the price distribution shift to a random place in the new support. Formally,

$$\text{If } p_t \geq \mu \underline{p}_t = \underline{p}_{t+1}$$

$$\text{then } p^*_{t+1}(p_t) = \left\{ \begin{array}{c} p_t \text{ with prob. } \rho \\ p' \text{ with prob. } 1-\rho \end{array} \right\}.$$

$$\text{If } p_t < \mu \underline{p}_t = \underline{p}_{t+1} \text{ then } p^*_{t+1}(p_t) = p'.$$

In the above equation, an asterisk denotes equilibrium values; ρ' is sampled from a distribution compatible with stationary equilibrium whereby $F_{t+1}(p) = F^*_t(p/\mu)$. A nice feature of this pricing policy is that when ρ equals 1, the model attains the smallest number of price changes and the highest average price duration. When its value is set to 0, firms change prices every period.

Not surprisingly, money is neutral in this model because the overall distribution of prices is perfectly flexible. However, money is not superneutral since inflation affects the nominal interest rate and thus real money holdings.

The next step in the analysis considers a calibrated version of the model. A set of specific parametric forms for costs, utility, and so on are selected, as is standard in the literature. When taking the model to the data, it is assumed that the decentralized market corresponds to the U.S. retail sector. Parameter values are chosen to match the average real interest and inflation rates of the U.S. data for 1988-2004. The empirical distribution of price changes of the retail sector and an average markup of 30 percent are also calibration targets. The model does an excellent job in matching these targets.

The authors then test the predictions of the model regarding the effects of inflation on the frequency and magnitude of price adjustments and contrast them to their empirical counterparts. Qualitatively speaking, the model is consistent with the data. Quantitatively, the model slightly overestimates the impact of inflation. The model reasonably accounts for the fraction of positive and negative price adjustments as inflation changes.

Conclusion

Head et al. show that individual sellers may optimally choose not to adjust prices as soon as the aggregate price level changes even though it is feasible and costless to do so because, in equilibrium, expected profits are equal for all possible prices a firm may post. The key friction is a market with matching where buyers can sample only a finite number of prices. Monetary policy is neutral, yet the model accounts for price facts very well. The key conclusion is that price stickiness does not necessarily imply that money is not neutral.

INTERGENERATIONAL REDISTRIBUTION IN THE GREAT RECESSION

If markets are incomplete, how does a negative productivity shock—a recession—affect the welfare of young and old individuals? This is the question investigated by Glover, Heathcote, Krueger, and Rios-Rull. The key observation is that households' portfolios vary systematically with age. In particular, older households hold a larger fraction of their wealth in the form of assets instead of human wealth (labor income). Consider then a negative shock that affects both asset returns and wages. This shock may induce older households to sell some of their assets and, ultimately, cause asset prices to fall more than wages. In this case, the welfare of the old decreases by more than that of the young. To see this, note that the young lose part of their labor income but are able to purchase assets at very low prices. Thus, there is a mechanism that compensates them for

the loss of income. There is no similar mechanism that reduces the impact of the recession for the old. Thus, if these forces are at work, we expect that one of the effects of a recession is redistribution of income from relatively old households to relatively young ones.

The Model

Glover et al. use an overlapping-generations model. They assume that preferences of a typical individual are given by

$$U = E\left\{\sum_{i=1}^{N}\left(\Pi_{j=1}^{i}\right)u\left(c_i\right).\right.$$

Individuals live for N periods (deterministic lifetime) and it is assumed that there are N cohorts alive at any given time, each of identical size.

Labor is traded in spot markets and households can transfer wealth across periods by trading shares of the representative firm. The model does not allow for capital accumulation.

Grover et al. study three versions of this economy distinguished by the portfolios that households are allowed to hold. The basic details are as follows:

- *Economy A*: In this economy, households can trade only shares in the risky firm. There are no safe assets (e.g., bonds).
- *Economy B*: In this economy, households can trade shares in the risky firm and a risk-free bond. However, they cannot choose their portfolio. It is exogenous and chosen to match the evidence for the United States.
- *Economy C*: In this version, households can trade both assets and they can choose the composition of their portfolios. In this economy, asset prices are endogenous.

Quantitative Results

Glover et al. calibrate the model using U.S. data to get a sense of the quantitative effects of a recession. The basic idea is to apply each version (economies A-C) and pick all relevant parameters to match the observed life-cycle profile of earnings, net worth, and portfolio composition. The calibrated version also matches the ratios of (i) aggregate safe assets to aggregate net worth and (ii) aggregate net worth to aggregate (over a 10-year period) labor income. They assume that a recession is given by a 10 percent drop in aggregate productivity that lasts 10 years. Thus, the model assumes that the expected duration of a period of high productivity is 66.7 years and the average duration of a recession is 11.8 years.

The nature of the experiment is to explore the consequences of a transition from the aggregate high-productivity shock to a low-productivity shock. Since the three economies differ in terms of access to financial instruments, it is useful to report the results separately.

Economy A. For this version of the model, Glover et al. find that the stock price falls 20 percent during a recession (a period of 10 years in real time). In the post-recession period, stock prices slightly overshoot their long-run values. In terms of welfare, the consequences of a recession are monotonic in age: Older generations suffer more. The actual magnitudes for the baseline calibration amount to a 15 percent decline in consumption. For the youngest households this loss is just over 1 percent.

Economy B. In this economy, the decline in asset prices is similar to that of Economy A and, in general, asset prices move in the same direction. The price of bonds declines even more than the price of stocks and the pattern of welfare losses replicates that of the one-asset economy.

Economy C. In this economy, asset prices decline by more than in the other two economies. This reflects the fact that young households are heavily leveraged (relative to the other economies) and, hence, their wealth takes a larger hit. The resulting decline in consumption by the younger generations results in a larger decline in asset prices. In this economy, the old hold more safe assets than in the data and hence their welfare does not suffer as much when the recession hits. In terms of welfare—measured in consumption-equivalent units—the oldest generation experiences the largest decline (11.12 percent), although a much smaller decline than in economies A and B (around 14 to 15 percent). Moreover, the youngest generation actually is slightly better off since they can purchase assets at a much

lower price. The welfare for all intermediate generations declines by more than in the other two economies.

Conclusion

Glover et al. analyze the distributional consequences of a large recession across different age cohorts. The model is highly stylized but, across different versions, a clear pattern emerges: The oldest generation appears to suffer the most from the recession since they must live the rest of their lives in that state. The very young are the least affected, as the majority of their lives will be spent in high-productivity times and they also benefit from the additional drop in asset prices when they are net buyers.

Even though asset markets are complete in a very narrow sense, there are no markets for intergenerational risk-sharing. From the perspective of a policymaker seeking to provide such insurance, a policy of subsidizing the old by running a deficit (issuing bonds) that will be paid off by future generations appears as an interesting policy alternative to run through this model.

SOCIAL SECURITY, BENEFIT CLAIMING, AND LABOR FORCE PARTICIPATION: A QUANTITATIVE GENERAL EQUILIBRIUM APPROACH

İmrohoroğlu and Kitao consider the quantitative implications of three alternative Social Security reforms: (i) a reduction in benefits and payroll taxes, (ii) an increase in the earliest retirement age (to 64 from 62), and (iii) an increase in the normal retirement age (to 68 from 66). A key motivation for this study is addressing the forecasted shortfall in the Social Security system. (According to the "2009 OASDI Trustees Report" of the Social Security Administration, either a large reduction in benefits or an increase in taxes may be required to prevent such a shortfall.) The framework of this analysis is a large-scale overlapping-generations model with incomplete insurance markets, endogenous savings, and a social security system similar to that of the United States. The model also includes health shocks, which are known to have important implications for precautionary savings and labor decisions. In contrast to previous studies within this branch of the literature, the model allows for endogenous decisions in both benefit claiming and labor force participation, which are at the core of the reforms considered by the authors.

A carefully calibrated version of the model suggests reforms (i) and (iii) (reducing benefits and payroll taxes and increasing the normal retirement age) have the largest impact. These reforms make agents save and work more to obtain better self-insurance, and the social security budget improves significantly. Reform (ii) (increasing the earliest retirement age) has a modest impact since the savings obtained from early retirees are compensated by higher expenditures at later ages.

The Model

The economy is populated by overlapping generations of individuals who live up to J periods. Their life span in uncertain. The population grows at a constant rate. Each individual is subject to a health status shock (driven by a Markov process that depends on age). The probability of living one more year is a function of the health status. Individuals start with no assets. Each individual has one unit of time per period that can be used in market or leisure activities. Agents may save in a risk-free asset but borrowing is not allowed. There is an exogenously given age-specific productivity profile that affects effective hours available for each activity. Agents face an idiosyncratic labor productivity shock every period. Finally, each individual is subject to idiosyncratic health expenditure shocks, and the distribution of these shocks depends on the health status of the individual. Health expenditure shocks are partially insured because of the presence of private health insurance. Each agent may or may not have access to private (employer-provided) health insurance. Access is determined in the first period of life and remains fixed into the infinite future. All of the elderly have access to government-provided health insurance (Medicare). Health insurance programs pay a fraction of gross expenditures

and charge a premium. In equilibrium, premiums are such that health insurance programs are actuarially fair.

The government operates a pay-as-you-go pension system. The details of the system are carefully constructed to mimic its U.S. counterpart. Benefits are a concave function of lifetime earnings, and benefits depend on whether the agent chooses to retire early or late. Some taxes are set to match the U.S. data counterparts. The labor income tax rate is set to balance the government budget.

Results

The model is calibrated to match data from the Medical Expenditure Panel Survey (MEPS), which includes self-reported health status and medical expenditure data. Age-dependent labor productivity is from Hansen (1993), while the idiosyncratic component is based on the estimates by Heathcote, Storeslettern, and Violante (2008), whose aim is to match the observed U.S. labor income inequality. The calibrated version of the model matches the targeted U.S. data well. This suggests that the model may be a reasonable instrument to perform quantitative counterfactual experiments (and to derive the possible implications of the proposed reforms to Social Security).

The first quantitative experiment assumes the social security tax and benefits are simultaneously reduced by 50 percent. The precautionary motif causes a very strong increase in savings such that aggregate capital increases by 10 percent. For instance, savings by retirement age increase by 18.8 percent. The total labor income tax declines, and thus the labor supply increases by about 3 percent. Higher wages and hours worked ultimately translate to a Social Security budget surplus of about 0.3 percent of GDP.

The second policy considered by the authors, which raises the earliest retirement age by two years, has small implications on aggregate macro variables. Furthermore, despite the government savings from not paying benefits to individuals 62 and 63 years of age, their benefits at a later age are also higher and offset savings.

The final experiment, which increases the normal retirement age by two years, has important quantitative implications. Old-age participation increases, more individuals postpone claiming their benefits and, as a result, the Social Security deficit becomes a surplus of 1.32 percent of GDP. The aggregate capital stock and assets at retirement age increase modestly (2.4 percent and 5.5 percent), relative to the first policy reform.

The last part of the paper evaluates the implications of some modifications to the theoretical framework. One extension of considerable interest introduces forecasted changes in demographic structure into the model (the benchmark experiment takes the current demographic structure as given into the infinite future). Of course, this results in an even more serious deficit for the system absent any reform. The first and last policy reforms have quantitatively similar implications, and both turn deficits into surpluses for the Social Security system.

Conclusion

İmrohoroğlu and Kitao study the quantitative implications of three different proposals to reform (and improve the budget of) the Social Security system. This is done within a framework of a very detailed social security system that mimics the main features of the U.S. system. Either a reduction in benefits and payroll taxes or an increase in the normal retirement age will create a system surplus. In the first case, this occurs because of large increases in precautionary savings and hours worked. In the second case, it is due to the higher old-age labor participation and a modest increase in savings. Increasing the earliest retirement age, in contrast, results in current savings that are essentially canceled by later (and higher) benefit claims and thus has negligible effects on the Social Security budget.

REFERENCES

Bronars, Stephen G. and Deere, Donald R. "The Threat of Unionization, the Use of Debt, and the Preservation of Shareholder Wealth." *Quarterly Journal of Economics*, February 1991, *106*(1), pp. 231-54.

Burdett, Kenneth and Judd, Kenneth L. "Equilibrium Price Dispersion." *Econometrica*, July 1983, *51*(4), pp. 955-69.

Glover, Andrew; Heathcote, Jonathan; Krueger, Dirk and Ríos-Rull, José-Víctor. "Intergenerational Redistribution in the Great Recession." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/krueger.pdf.

Greenwood, Jeremy; Sanchez, Juan M. and Wang, Cheng. "Quantifying the Impact of Financial Development on Economic Development." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/greenwood.pdf.

Guler, Bulent; Guvenen, Fatih and Violante, Giovanni L. "Joint-Search Theory: New Opportunities and New Frictions." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/guvenen.pdf.

Hansen, G.D. "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked." *Journal of Applied Econometrics*, January-March 1993, *8*(1), pp. 71-80.

Head, Allen; Liu, Lucy Q.; Menzio, Guido and Wright, Randall. "Equilibrium Price Dispersion and Ridigity: A New Monetarist Approach." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/menzio.pdf.

Heathcote, Jonathan; Storesletten, Kjetil and Violante, Giovanni L. "The Macroeconomic Implications of Rising Wage Inequality in the United States." NBER Working Paper No. 14052, National Bureau of Economic Research, June 2008; www.nber.org/papers/w14052.pdf.

İmrohoroğlu, Selhattin and Kitao, Sagiri. "Social Security, Benefit Claiming, and Labor Force Participation: A Quantitative General Equilibrium Approach." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/imrohoroglu.pdf.

Jovanovic, Boyan and Menkveld, Albert J. "Middlemen in Limit-Order Markets." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/jovanovic.pdf.

Lagos, Ricardo and Wright, Randall. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy*, June 2005, *113*(3), pp. 463-84.

Matsa, David A. "Capital Structure as a Strategic Variable: Evidence from Collective Bargaining."' *Journal of Finance*, June 2010, *65*(3), pp. 1197-232.

McCall, John J. "Economics of Information and Job Search." *Quarterly Journal of Economic Theory*, February 1970, *84*(1), pp. 113-26.

Midrigan, Virgiliu and Xu, Daniel Y. "Finance and Misallocation: Evidence from Plant-Level Data." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/midrigan.pdf.

Milgrom, Paul and Stokey, Nancy L. "Information, Trade and Common Knowledge." *Journal of Economic Theory*, February 1982, *26*(1), pp. 17-27.

Monacelli, Tommaso and Quadrini, Vincenzo. "Financial Markets and Unemployment." Presented at the 35th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 21-22, 2010; http://research.stlouisfed.org/conferences/policyconf/papers2010/quadrini.pdf.

Mortensen, Dale T. "A Theory of Wage and Employment Dynamics," in Edmund S. Phelps, Armen A. Alchian, and Charles C. Holt, eds., *Microeconomic Foundations of Employment and Inflation Theory*. New York: W.W. Norton, 1970, pp. 167-211.

Myerson, Roger B. and Satterthwaite, Mark A. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory*, April 1983, *29*(2), pp. 265-81.

U.S. Social Security Administration Board of Trustees. "The 2009 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds." May 12, 2009; www.ssa.gov/OACT/TR/2009/tr09.pdf.

Federal Reserve Bank of St. Louis
P.O. Box 442
St. Louis, MO 63166-0442