Three Lessons for Monetary Policy from the Panic of 2008

James Bullard

This article is a modified version of a presentation given at the Federal Reserve Bank of Philadelphia's policy forum "Policy Lessons from the Economic and Financial Crisis," December 4, 2009. The presentation was made during a panel discussion that also included John Taylor and N. Gregory Mankiw.

Federal Reserve Bank of St. Louis *Review*, May/June 2010, *92*(3), pp. 155-63.

We have been wrestling with one of the most severe recessions in the post-World War II era; moreover, it has been accompanied by a widespread financial crisis. After unprecedented policy responses, there are signs of recovery on both fronts. So, it is not too early to take stock of our actions and attempt to learn lessons from our recent past—lessons for monetary policy, financial regulation, and other aspects of the crisis. My objective here is to focus on lessons for monetary policy alone and leave discussion of regulatory issues and financial markets for another day.

AN OVERVIEW OF THE CRISIS

First, let me explain the nature of the crisis as I see it. My narrative is a little bit different from what some people describe, so it's important that I establish it before I talk about any lessons to be learned. I think that history will assert that there was a panic in the autumn of 2008, and maybe that's fair. But this crisis actually began much earlier: The onset is usually dated as August 2007, with many key events along the way.

For instance, in October 2007, equity prices actually peaked. The initial reaction by policymakers, as well as markets, then, was to view the crisis as perhaps less severe than it actually turned out to be. In March 2008, Bear Stearns was purchased by J.P. Morgan with Fed assistance. The U.S. economy continued to grow through the second quarter of 2008. We did have a commodity price spike in the second quarter of 2008: $100 per barrel of West Texas intermediate crude oil in March. Then the price of oil went up another $45 from there, which is well above, in real terms, the 1980 peak in oil prices.

The economy began to slow down in the third quarter of 2008. And the contracting economy, both in the United States and abroad, intensified the financial crisis, which at that point had been roiling for an entire year. But because the economy started slowing, the crisis greatly worsened during the autumn; dozens of firms worldwide required assistance to avoid bankruptcy in the fourth quarter of 2008. In that quarter and the first quarter of 2009, major economies worldwide contracted.

James Bullard is president of the Federal Reserve Bank of St. Louis. The author appreciates the assistance and comments provided by colleagues at the Federal Reserve Bank of St. Louis, especially Chris Waller, senior vice president and director of research, and Bob Rasche, executive vice president and senior policy advisor. Marcela M. Williams, special research assistant to the president, provided assistance.

This is a little bit different from what you sometimes hear, but it is predicated on the idea that the crisis had been going on for a long time before the economy actually started to contract.[1]

The Monetary Policy Response

Basically, the Federal Reserve's monetary policy response to the crisis can be divided into three parts. The first part was a wide array of collateralized lending programs, which, in my discussion here, I am going to lump all together and call liquidity programs. After September 2008, these were funded by reserve creation—that is, by printing money. These programs are temporary in nature, and I don't view them as an inflationary threat.

The second part was to move the target policy interest rate toward zero—in fact, very close to zero. The Fed was actually very aggressive in lowering rates during the last part of 2007 into the first part of 2008.

The third part of the policy response was an aggressive asset purchase program. I'm going to put this topic under the title "quantitative easing." This response was also funded by reserve creation, like the liquidity programs, but in this case the balance sheet effects are far more persistent. I think that this program creates a medium-term inflation threat in a way that the liquidity programs do not. I will continue this discussion in the next section.

THREE LESSONS

I will focus on three lessons for monetary policy that have emerged from recent events: They have to do with understanding that (i) the role of lender of last resort can be carried out on a grand scale; (ii) quantitative easing can substitute for policy rate easing after the zero bound is encountered; and (iii) the connections between asset pricing and monetary policy must be a top priority going forward.

[1] This episode has been described as the "perfect storm": In sum, the financial crisis began in 2007—exacerbated by the spike in commodity prices in the first half of 2008—and the economic slowdown revealed itself in the autumn of 2008, at which time panic ensued worldwide.

Lesson 1: Lender of Last Resort on a Grand Scale

My first lesson is about the Fed's role as lender of last resort *on a grand scale*. So what is the lesson? It is that the Fed's ability to act decisively in a crisis through its lender-of-last-resort function far outstrips previous conventional wisdom. I think that the response has been more creative and much more substantial than people would have imagined: If you had a conference before the crisis to discuss the lender-of-last-resort function, I don't think the recent actions by the Fed would have been predicted.

Going forward, I think that these liquidity programs need to be carefully evaluated. That's a call for our research community to study them closely, and John Taylor has been a leader in analyzing the effectiveness and implications of these programs (Taylor and Williams, 2008, 2009). One concern is that the scale of these liquidity programs may unintentionally be setting up expectations of future intervention, and I believe we need to think carefully about that. How are markets expecting we're going to react in future crises? Is that something we desire or not? And how should we account for that?

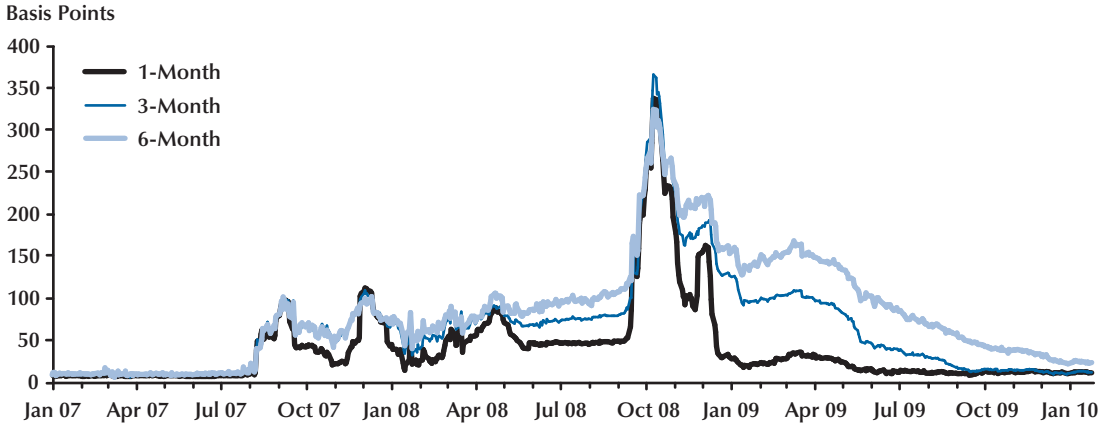
The lender-of-last-resort function—lending extensively in response to a crisis—has been an integral part of central banking for the past 200, maybe 300 years. It is all collateralized lending, and the basic premise is that it is necessary to provide a lot of liquidity to markets in the event of a crisis.

The Fed was very innovative in this area. We developed a wide array of liquidity programs in 2007-08 that were all designed to improve market functioning. These programs were always meant to be temporary in nature and are therefore priced in such a way that markets will not find them attractive once the crisis passes. And that is happening now on a large scale: As market functioning improves, these programs become less necessary.

Some of these programs are within the Fed's traditional purview; some were authorized under the so-called 13(3) provision in the Federal Reserve Act, which allows the Fed to lend to other parties

Figure 1

LIBOR-OIS Spread



NOTE: Data are through January 25, 2010.

SOURCE: *Financial Times* and Reuters.

in unusual and exigent circumstances. It's a wide variety, but all programs are intended to improve market functioning. Some may have worked better than others, of course, and so we should evaluate these programs carefully. As we do that, we should keep in mind that, simultaneous to the Fed's collateralized lending, many government guarantees were coming into play. As future research addresses this topic, it will be essential to evaluate the effect of the government guarantees in tandem with the effect of providing liquidity to the markets.

But regardless of how these individual programs perform, by many metrics they are considered an overall success. Although we're not completely back to pre-crisis levels, global financial markets are less strained than they were. Figure 1 shows a familiar picture: the LIBOR-OIS spread dating back to January 2007. Again, the crisis started in earnest in August 2007, the point at which this LIBOR-OIS spread jumps up. That jump was actually considered gigantic at the time, although it turned out to be relatively small, as events unfolded in the autumn of 2008.

These spreads have decreased substantially, even though they are not back to where they were during the first half of 2007.[2] This reversal is often attributed to the liquidity programs. You could argue about it, but I think the point is that the liquidity programs, as a pillar of the monetary policy response to the crisis, are certainly being used a lot less intensively today than they were even six months ago. The Fed's core idea is to let these programs continue to wind down naturally and to end the authority of the 13(3) provision for emergency programs in 2010.

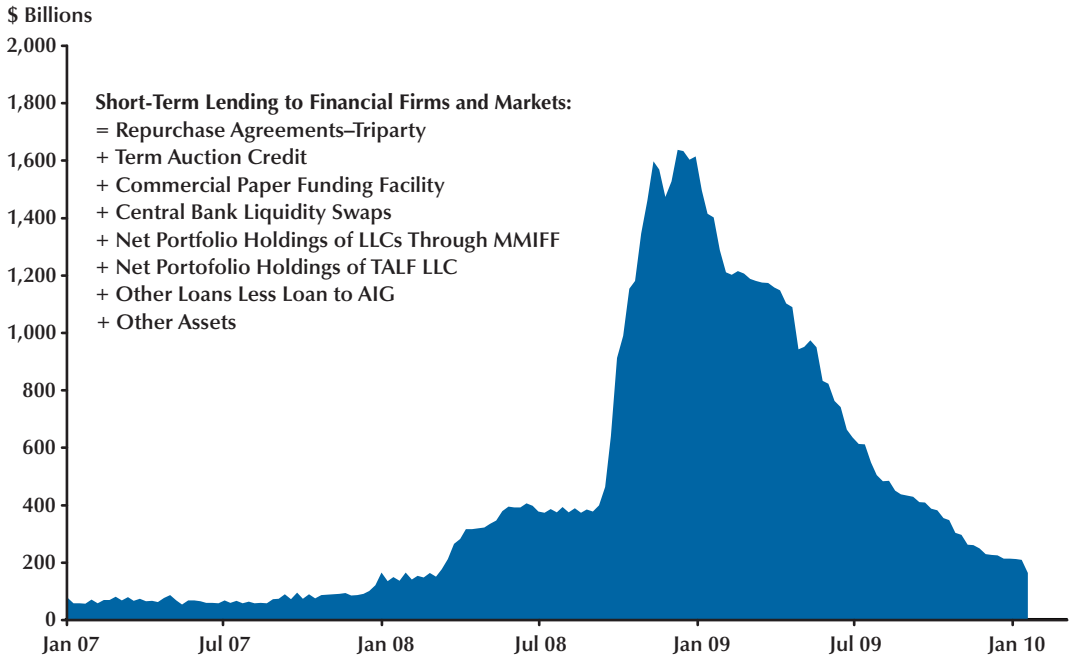
Figure 2 shows the volume of reserves supplied to financial firms and markets through liquidity facilities. The peak exceeds $1.6 trillion. This is what I meant by "lender of last resort on a grand scale." On September 11, 2001, we had about $40 billion of reserves in the system and we almost doubled that to about $75 billion.[3] At the time, we thought *that* was a gigantic sum. But this recent level of $1.6 trillion of lending to firms and markets in response to this crisis is truly a monumental attempt to alleviate liquidity constraints in the markets and to get past this crisis—and one metric for measuring how large this crisis really was.

[2] In fact, these spreads are quite close, with the exception of the 6-month spread.

[3] This figure is for total reserves adjusted for reserve requirements.

Figure 2

Short-Term Lending to Financial Firms and Markets



NOTE: Data are through January 10, 2010. MMIFF, Money Market Investor Funding Facility; TALF, Term Asset-Backed Securities Loan Facility.

SOURCE: Federal Reserve Board, H.4.1.

Again, these programs are priced to be unattractive in normal circumstances, which is why they have declined so rapidly from their peak of $1.6 trillion to under $300 billion. Firms do not want to use these programs unless they really need them. The expectation is that these programs will cease to exist by the first quarter of 2010 if financial conditions continue to improve; at that point, the Fed's short-term lending to ensure liquidity will return to a minimal level.

This lesson, then, is that these programs are much larger and more varied than could have been anticipated before the crisis. It is time to evaluate which ones worked and which ones didn't and to think much more carefully about the ramifications of the lender-of-last-resort policy, which has not often been as prominent a topic in the research world as other aspects of monetary policy. We also need to assess whether we have unwittingly set up expectations of future intervention that could be influencing markets today.

Lesson 2: The Several Faces of Monetary Policy

The United States has not had policy rates at (or essentially at) zero since the 1930s. Yet, even in this current environment, the Fed has not ceased to function. The second lesson is that monetary policy can be conducted by different means. Normally, we think of monetary policy as interest rate adjustment, but when you reach the zero bound in nominal interest rates, you have to adjust monetary policy in other dimensions and that's exactly what the Fed has done. There may have been some doubt—before this recent episode—about the ability of the Fed to conduct a business

cycle stabilization policy with policy rates near zero. But I think this episode clearly reveals that the Fed is perfectly capable of doing so.

In my opinion, analogous to interest rate policy, quantitative policy should be state contingent; that is, it should adjust according to incoming information on the state of the economy. And these types of policy do have some relation to each other. Although they are different, I think that any quantitative policy should be conducted in a manner that's *analogous* to interest rate policy. To me, that means adjusting the policy according to incoming information.

Since its March 17-18, 2009, meeting, the Federal Open Market Committee (FOMC) has explicitly stated that it will keep the federal funds rate target near zero for an "extended period." Any movement away from that position will be contingent on both inflation and real economic developments. But my question is this: How should the FOMC conduct business cycle stabilization policy during the period of near-zero policy rates? And the answer is that there are many interest rates and many assets that the Fed can influence.

The FOMC communicated its plans to make more than $1.7 trillion in outright asset purchases in a series of announcements beginning about December 2008. The purchases are agency debt (in this case, "agency" means Fannie Mae and Freddie Mac), agency mortgage-backed securities (MBS), and longer-term Treasury securities. The bulk of these purchases are agency MBS. Again, this is being financed by reserve creation, or printing money. Consequently, the monetary base has more than doubled, creating a medium-term inflation risk.

This point deserves some illumination. According to monetary theory, very large increases in the monetary base are inflationary. And the inflationary effects of very large increases in the monetary base depend on at least two factors: One factor is private sector expectations of the future level of the monetary base. As always, in macroeconomics and monetary theory, expectations are very important. With large increases that are expected to be temporary, as they are with the liquidity programs, I don't believe there is much of an inflationary threat. But large increases in the monetary base that are expected to be permanent—or at least more persistent—may indeed be inflationary. Increases in the monetary base that are associated with asset purchases fall into this more-persistent category, which is why this aspect of current policy poses a medium-term inflation risk.

The second factor that would affect the medium-term inflation risk is the speed with which the increases in the *monetary base* translate into increases in the *money supply*. The monetary base is not the money supply: Before monetary increases can affect inflation, the "money" must be incorporated into ordinary transactions. That is not occurring right now; the speed with which the monetary base is being translated into changes in the money supply is very slow at this point. This often happens when the economy slows down as rapidly as the U.S. economy has. But we can expect that this process will start to accelerate and that may affect the medium-term inflation risk.

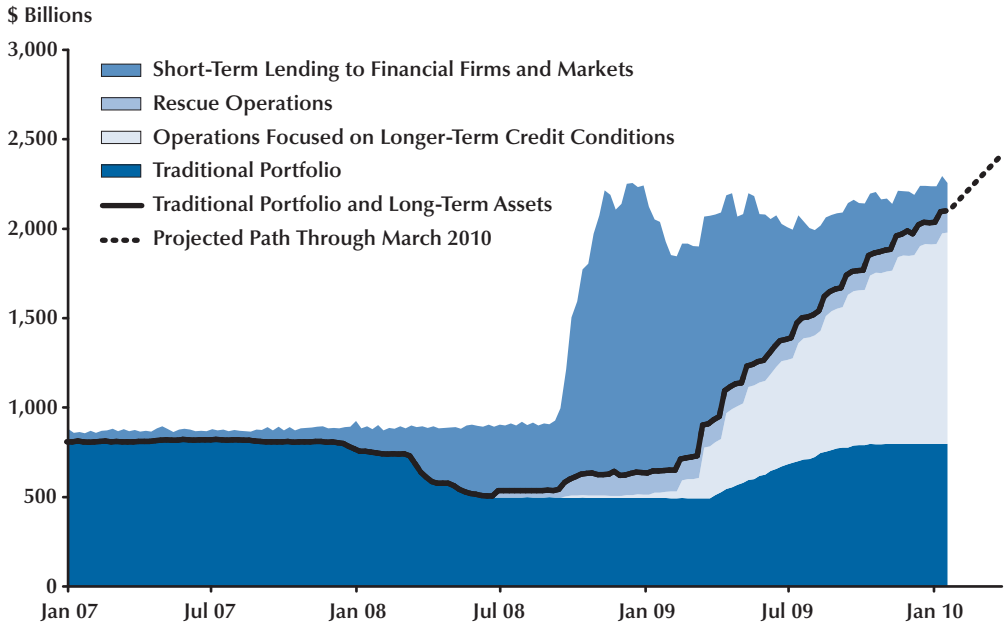
Figure 3 shows the Fed's balance sheet in a particularly instructive way. The area above the black line is a duplication of Figure 2: the total volume of the liquidity programs. As I said previously, that component of the balance sheet is not worrisome; it does not create a medium-term inflation risk.

The dark blue area below the black line shows the more traditional holdings of the Federal Reserve, such as Treasury securities, and the light blue area under the black line is the MBS purchase program, which has grown very large. The dotted black line projects how this part of the balance sheet will continue to grow through the first quarter of 2010. Under current conditions, the size of the balance sheet will increase to about $2.4 trillion. And again, unlike the liquidity programs, these purchases will not run off in a period of months because these assets have much longer maturities: seven to ten years. In this case, we're talking about mortgages. And so we would expect this expansion of the monetary base, then, to be much more persistent and not likely to dissipate in a timely fashion.

Moving on to the asset purchases as quantitative easing: Again, the FOMC moved its policy

Figure 3

Composition of the Federal Reserve's Balance Sheet and Projected Path Through March 2010



SOURCE: Federal Reserve Board, H.4.1.

rate toward zero in December of 2008, and right after that began this asset purchase program—in fact, in the first month of 2009. I think that this program has been regarded as successful, as it further eased monetary conditions after the zero bound was encountered. So in some sense, the asset purchase program substituted for additional easing that could not be done through the policy rates, since the policy rate had come very close to zero. So a natural way to run monetary policy going forward would be for the FOMC to continue to adjust the asset purchase program while the policy rate remains near zero.

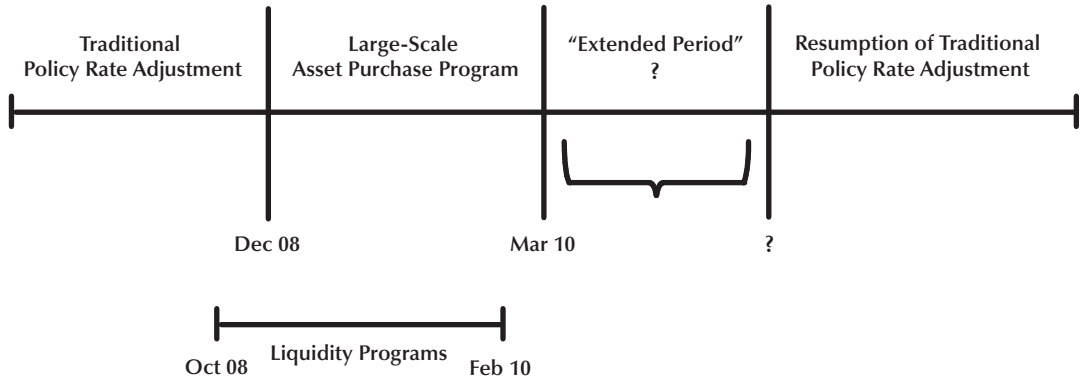
I like to think of asset purchases in terms of state-contingent policy. When we adjust interest rates, we always do so in response to economic conditions: readings on inflation and on the real economy. The famous Taylor rule is one example of that, but there are many other examples, and it's a natural way for the central bank to operate.

The asset purchase program that we have in place right now does not have this state-contingent character. What we on the FOMC did as a committee is simply announce that $1.725 billion of assets would be purchased by the first quarter of 2010. I don't see anything optimal about simply announcing a number and buying that amount of assets. It may be helpful for monetary policy going forward to think more in terms of adjusting this program as macroeconomic information arrives. That's what you would do with the Taylor rule. Although there's no guarantee at this point that current quantitative policy will become state-contingent, it seems to me if you're going to have two policy instruments in place, you should have them operate in the same fashion: both adjusted in response to incoming information.

And, with the policy rate near zero, the asset purchase program could very easily dominate policy for some time. In fact, I suggest staying active in the market for agency MBS. If encouraging information on the economy arrives after the first quarter of 2010, then we could consider

Figure 4

Timeline of Monetary Policy



removing some accommodation through asset sales. We certainly wouldn't want to do that in blanket fashion; this is all about adjustments at the margin, so you'd adjust a little bit through asset sales. On the other hand, if discouraging economic news came in, then you could consider additional asset purchases. This would allow monetary policy to remain active, responding to shocks during the period of near-zero interest rates.

Figure 4 shows a timeline. Traditional policy rate adjustment, as it's been practiced in the United States and around the world over the past 20 to 25 years and maybe quite a bit longer than that, is noted on the left side. The liquidity programs appear beneath and extend from October 2008 until the first quarter of 2010—February or March 2010. These liquidity programs were a response to the crisis, but they are set off to the side because they're not part of the traditional policy response or an attempt to run stabilization policy. Once the rate approached zero in December of 2008, we began our large-scale asset purchase program, which is shown as continuing through March 2010. The timeline then shows a period with some question marks that refers to the extended period language that the FOMC has adopted. And as I said, we will continue to keep rates low for an extended period, and what we do in the future will depend on how the data come in on the economy.

But during this period, you could also adjust your asset purchases in one direction or the other as information arrives, perhaps before you want to make a decision on the interest rate margin. And then at some point down the road—and I'm being very cagey here, by putting a question mark on that point—you'd make a decision on the interest rate; you'd return to a traditional policy rate adjustment and would go on from there. So this is just a suggestion about how to think about policy in 2010 and beyond.

In summary, the asset purchase program is very large. It is being financed by reserve creation—i.e., printing money. It is generally considered successful. And it has substituted for easing that cannot be accomplished through the policy rate. Longer-term interest rates generally fell as aspects of the program were communicated in Chairman Bernanke's announcements in late 2008 and then in further announcements in the first part of 2009. And I think that the FOMC could use the program to respond to incoming information on the economy during the period of near-zero interest rates.

Lesson 3: Bubbles

The third lesson is about asset price "bubbles." (I'm placing bubbles in quotation marks because I'm trying to exorcise the bubble language. I have been unable to do it so far.) It is a very serious

issue for monetary policy and has been debated extensively over the past 15 years. But now we are having a renewed and more intense debate. The main problem in thinking about this issue appears to me to be that it is hard to see what was wrong with the previous policy, given conventional ideas about what we attempt to accomplish with policy.

We have had two decades with two bubbles. The first one was in the 1990s, and the second one was in the current decade. In both cases, after the 1991 recession, and again after the 2001 recession, we had jobless recoveries; and it took a long time before the Fed decided to raise rates and come off their cyclical lows in either of those cases. In the 1990s we had the so-called "dot-com bubble." In the 2000s we had the "housing bubble," and the drag on the economy from the housing decline has been really severe since 2006.

Despite this, the monetary policy outcomes during the past two decades—up to the current recession—actually have been quite good. Unemployment hit lows of 3.8 percent and 4.4 percent, respectively. In the current decade, inflation has been low and stable. In terms of the conventional ideas about what monetary policy tries to accomplish, those years were quite good. Still, even without an increase in inflation, the asset price misalignment seemed to have caused significant problems for the macroeconomy, and that may mean that we should put more weight on asset prices going forward.

Now, there has been debate on this, and this is my final point. There is a policy debate on this topic and there is an academic debate. The policy debate has made good points: that it is difficult to identify asset price misalignments in real time and that interest rates are a blunt instrument to respond to asset price misalignments. I think not all bubbles are bad. Consider the "tech bubble" of the 1990s. A lot of good technology was developed then, even though it seemed to be an asset price misalignment. And this is what the policy debate has said.

There is also an academic literature on this issue, and it generally does not come into the policy discussion. The literature is about multiple equilibria: There is a set of expectations and a set of prices that will clear markets, but there is another set of expectations and another set of prices that will also clear markets. And, as described in the literature, the objective is to adopt a policy to quash these multiple equilibria so that you're left with only the fundamental equilibrium. At that point, the economy will bounce along according to the actual shocks that hit the economy.

In my mind, this is a more sophisticated way to think about "asset bubbles" and how to respond to them. According to this literature, one example of a policy that works fairly well is for monetary policy to react aggressively to shocks. If this is done, then multiple equilibria—specifically, those that are *not* based on fundamentals—are discarded and the economy is kept near its fundamental equilibrium. I interpret that as the best policy to follow to avoid problems with bubbles.

So if we approach this issue seriously and intensify this debate further, we may have to entertain these sorts of ideas and step up our analysis of this issue.

CONCLUSION

The recent crisis has been challenging, to say the least. But we have the opportunity to evaluate our responses and the effects of those responses. To sum up: The first lesson is that the lender-of-last-resort function has proven much more flexible and more powerful than previously believed. The second lesson is that the asset purchase program has shown that active stabilization policy is possible with the policy rate at zero. And the third lesson is that clearly the issue of "asset price bubbles" is a hard one for monetary policy and may require new and innovative analysis.

REFERENCES

Taylor, John B. and Williams, John C. "A Black Swan in the Money Market." *American Economic Journal: Macroeconomics*, January 2009, *1*(1), pp. 58-83.

Taylor, John B. and Williams, John C. "Further Results on a Black Swan in the Money Market." Unpublished manuscript, May 2008; www.stanford.edu/~johntayl/Taylor-Williams-Further%20Results%20on%20Black%20Swan.pdf.

Getting Back on Track: Macroeconomic Policy Lessons from the Financial Crisis

John B. Taylor

This article reviews the role of monetary and fiscal policy in the financial crisis and draws lessons for future macroeconomic policy. It shows that policy deviated from what had worked well in the previous two decades by becoming more interventionist, less rules-based, and less predictable. The policy implications are thus that policy should "get back on track." The article is a modified version of a presentation given at the Federal Reserve Bank of Philadelphia's policy forum "Policy Lessons from the Economic and Financial Crisis," December 4, 2009. The presentation was made during a panel discussion that also included James Bullard and N. Gregory Mankiw.

Federal Reserve Bank of St. Louis *Review*, May/June 2010, *92*(3), pp. 165-76.

I started doing research on the financial crisis in the spring and summer of 2007 just before the crisis flared up in August of that year. From the start, my approach has been empirical. I have not focused on who said what to whom when, however interesting and ultimately important that story is. Rather I looked at the timing of events and at data—at interest rates, stock prices, credit flows, money supply, housing starts, income, consumption—using statistical techniques and simple charts, concentrating on what is amenable to economic analysis. I also tried to use the discipline of "counterfactuals," or stating what alternative policies or events would have been and using economic models to examine the impacts. I looked at economic policy throughout the crisis, including the period leading up to the panic in the fall of 2008 and the year and a half since then. What I have found since the start of this research is that government interventions—many well-intentioned government interventions—did a great deal of harm. With these findings in mind, I wrote one of the first books on the crisis; it was based on a November 2008 speech in honor of David Dodge, the former governor of the Bank of Canada. I called the book *Getting Off Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis*. I think that events since that book was published have reinforced the title.

In these remarks I want to consider the macroeconomic—monetary and fiscal, as distinct from regulatory—policy implications of these findings. As I hope to show, once the findings are clearly laid out, the macroeconomic policy implications jump out at you and happen to be quite straightforward: *Get back on track. Return to what was working well before policy got off track.*

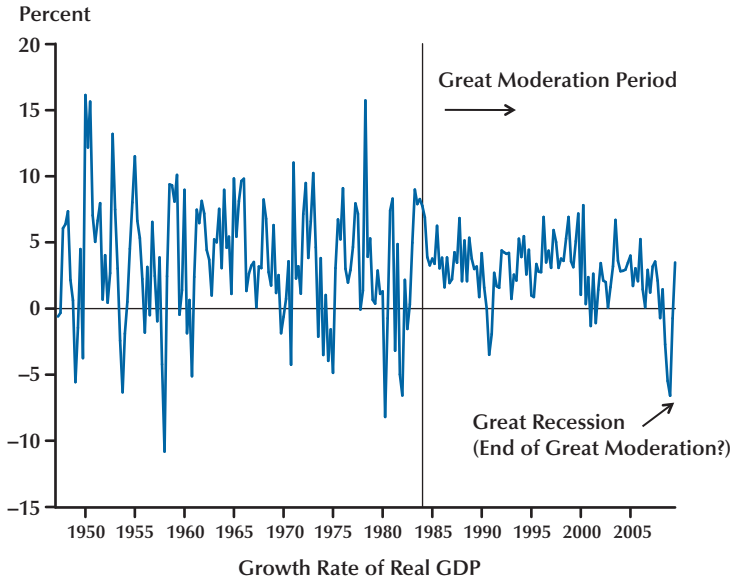
FROM THE GREAT MODERATION TO THE GREAT DEVIATION TO THE GREAT RECESSION

Figure 1 provides an illustration of what I have in mind. It shows the growth rate of real GDP in the United States, quarter by quarter, back to

John B. Taylor is the Mary and Robert Raymond Professor of Economics at Stanford University and the George P. Shultz Senior Fellow in Economics at the Hoover Institution.

Figure 1

From Great Moderation to Great Deviation to Great Recession



the late 1940s. The high volatility during the earlier decades—the 1950s through the 1970s—is clearly visible. During the latter part of this high-volatility period, I was a visiting scholar at the Federal Reserve Bank of Philadelphia; we were studying monetary policy decisions, trying to understand the reasons for the volatility and to find ways to reduce it.[1] As Figure 1 shows, the high volatility ended in the early 1980s. It is hard to say exactly when. Some economists say 1984. I say a little earlier, at the beginning of the expansion in 1982. However you date it, a Great Moderation—two or more decades of much less volatility—followed. And this Great Moderation, with its long expansions and short recessions and low inflation, continued until the recent financial crisis, when it apparently ended.

Why did the Great Moderation end? In my view, the answer is simple. The Great Moderation ended because of a "Great Deviation," in which economic policy deviated from what was working well during the Great Moderation. Compared with the Great Moderation, policy became more interventionist, less rules-based, and less predictable. When policy deviated from what was working well, economic performance deteriorated. And lo and behold, we had the Great Recession.

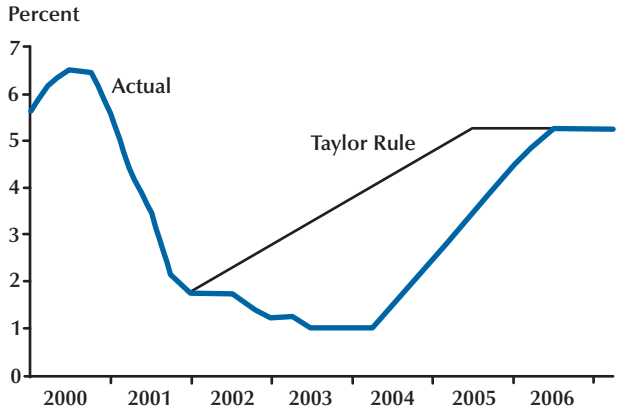
Monetary Excesses

A good illustration of policy decisions that fall under this Great Deviation rubric is shown in Figure 2. This chart, which appeared in *The Economist* magazine, October 18, 2007, plots the interest rate set by the Federal Reserve from 2000 to early 2007. I reproduced this chart in *Getting Off Track*; a version first appeared in the paper I prepared for the Kansas City Fed Symposium in the summer of 2007. Note how the interest rate came down in the recession of 2001, as it would be expected to do, but then became very low—falling below 2 percent and then down to 1 per-

[1] As a visiting scholar at the Federal Reserve Bank of Philadelphia in the late 1970s, I had the opportunity to participate in briefings for the president of the Bank before FOMC meetings. We discussed how policy should become more aggressive with respect to policy rate increases and how more attention should be paid to *real* interest rates. It is amazing how much things changed after that, but I worry that perhaps history is repeating itself.

Figure 2

Federal Funds Rate, Actual and Counterfactual



NOTE: From Taylor (2007).

cent—before rising back up again slowly. This is the period in which interest rates were too low according to the Taylor rule,[2] which is shown by the dark line in the figure representing what policy would have been had it followed the principles that worked well for the previous 20 years. That is, interest rates would not have reached such a low level and they would have returned much sooner to the neutral level. So in this sense there was a deviation from a more rules-based policy. The deviation was larger than in any period since the unstable decade before the Great Moderation.[3] One does not need to rely on the Taylor rule to come to the conclusion that rates were held too low. The real interest rate was negative for a very long period, similar to what happened in the 1970s.

So it should not be surprising that such an unusual policy led to some problems. According to my research, the low interest rates added fuel to the housing boom, which in turn led to risk taking in housing finance and eventually a sharp increase in foreclosures and balance sheet deterioration at many financial institutions. To test the connection with the housing boom, I built a simple model relating the federal funds rate to housing construction. I showed that a counterfactual higher federal funds rate would have avoided much of the boom as described in my 2007 Jackson Hole paper.[4]

I call this monetary policy decision a discretionary *intervention* by government because it was an intentional departure from the policies that were followed in the decades before. Some policymakers say the departure was undertaken to avoid downside risk, perhaps a Japanese-style deflation. I have no doubt that it was well-intentioned, an example of what used to be called discretionary fine tuning. The Fed's descriptions that rates would be low for a "prolonged period" and that rates would rise at a "measured pace" illustrate this fine tuning. Markets were generally aware of it and the departure from policy rules confirmed it. I think it is an example where the perfect can

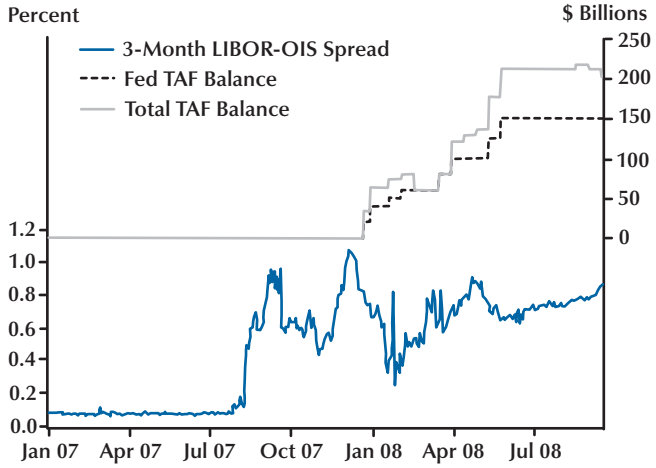
[2] I sometimes wish it were not called the Taylor rule, because I lose objectivity discussing it.

[3] Ben Bernanke (2010) replied to this criticism and I responded in Taylor (2010a).

[4] Alan Greenspan replied to this criticism (see Greenspan, 2010, for example) focusing on the long rate, not the short rate. My response is found in a "frequently asked questions" section of *Getting Off Track*.

Figure 3

The Term Auction Facility and the LIBOR-OIS Spread



become the enemy of the good. As Milton Friedman[5] once put it, "The attempt to do more than we can will itself be a disturbance that may increase rather than reduce instability."

This is not the whole government part of the story, of course. The government-sponsored enterprises, Fannie Mae and Freddie Mac, also encouraged the housing boom. But whether or not you include these on the list, the ultimate source of the extraordinary housing boom and the subsequent housing bust and financial distress was government policy. Capital inflows from abroad may have added to the problem, but the evidence is clear that monetary policy had deviated in the direction that would likely lead to poor policy performance. This is in contrast to the policy decisions during the Great Moderation.[6]

More Interventions

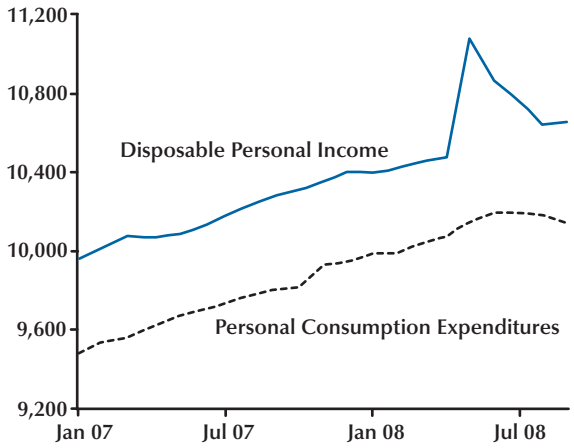
When the crisis became evident with the flare-up in the money markets in August 2007, a host of additional interventions were undertaken by government, but these had little positive impact. In my view the crisis was misdiagnosed as a liquidity problem rather than a counterparty risk problem in the banks; as a result, the policies did not address the problem. To illustrate this perspective, consider Figure 3, which shows the LIBOR-OIS spread through the summer of 2008 along with one of these interventions—the term auction facility (TAF). The LIBOR-OIS spread is the difference between the interest rate on 3-month unsecured loans between banks (LIBOR) and an estimate of what the federal funds rate will be, on average, over those same three months (OIS). The spread is a good measure of tension in the interbank market. The jump in the LIBOR-OIS spread in August 2007 is very clear in Figure 3. I first began researching that jump soon after it occurred, trying to determine what caused it. I enjoy following the federal funds market, and when I saw this jump I was naturally curious. Based on work with John Williams of the Federal Reserve Bank of San Francisco (Taylor and Williams, 2009), I concluded the jump in spreads was due to counterparty risk in the banking sector. We now know the banks were holding many toxic assets, but that was not clear to many at the time, and the problem was diagnosed as a liquidity

[5] Testimony to the Joint Economic Committee in 1958, quoted in Friedman and Heller (1969, p. 48).

[6] Many papers were written before this crisis on the effects of monetary policy in the Great Moderation: e.g., Ben Bernanke (2004) showed that policy rules made a substantial difference in, and were largely responsible for, the Great Moderation.

Figure 4

Consumption and the Jump in Personal Disposable Income Due to the Fiscal Stimulus: Part 1 (Monthly Data, Seasonally Adjusted, Annual Rates)



problem. John Williams and I called our paper "A Black Swan in the Money Market" because the event was so unusual.

As a result of the misdiagnosis, one of the policy interventions was to increase the supply of liquidity through the TAF, as shown in Figure 3, with some foreign central banks joining in. When these facilities were first enacted, in late December 2007, the LIBOR-OIS spread declined a bit. But this respite did not last, and as is clear in Figure 3 the spread rose again and remained high. I find no strong evidence that these liquidity facilities affected these rates. And the evidence remains lacking to the present. In fact, if you look at reasonable measures of risk in the banking sector, such as the spread between secured and unsecured interbank loans, you can explain the movements in LIBOR-OIS very well. In my view, this policy intervention prolonged the crisis because it did not address the balance sheet problem at the banks and other financial institutions.

Discretionary Countercyclical Fiscal Actions

Policy interventions also occurred on the fiscal side. Figure 4 illustrates one discretionary fiscal intervention—the fiscal stimulus of 2008. The chart shows that disposable personal income rose as checks were sent to people as part of the stimulus package. The intention was to provide temporary tax rebates so the recipients would spend that money and jump-start the economy. This action also was a deviation from policies that were working well for 20 years, a period when very few such discretionary policies were implemented.

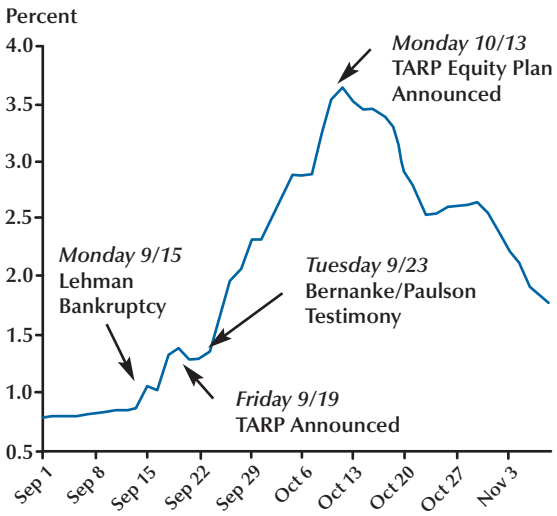
Figure 3 provides no evidence that the stimulus has had any impact in raising consumption. While disposable income increased dramatically as a result of the rebates, personal consumption expenditures did not increase. This is what economic theory—the permanent income theory or life cycle theory of consumption—would tell you. Again, the intervention did not address the toxic asset problem, and the crisis continued.

Interventions to Rescue the Creditors of Individual Financial Firms

The most unusual and significant actions were the government interventions to rescue financial firms and their creditors, culminating in the rollout of the Troubled Asset Relief Program

Figure 5

Event Study of the Worsening Crisis: LIBOR-OIS Spread, Fall 2008



(TARP) during the week of September 21, 2008. In my view, however, the rollout was part of a chaotic series of interventions going back to Bear Stearns in March 2008 and included the Fannie and Freddie interventions, the AIG intervention, and even the Lehman non-intervention, which I include because the decision not to intervene was a big surprise. Figure 5 shows the LIBOR-OIS spread during the panic period. Recall from Figure 4 that the LIBOR-OIS spread jumped in August 2007. But the spread increased by much more during the panic, by more than 350 basis points after hovering close to 100 basis points since August 2007. Figure 5 focuses on several key events, which are labeled on the graph. The Lehman bankruptcy occurred early Monday, September 15, after a long weekend during which a decision was made not to bail out Lehman and its creditors. Observe that the LIBOR-OIS spread increased slightly on September 15 and then fluctuated during the rest of the week. But these turned out to be relatively minor movements. The major movements in the spread occurred with the government's rollout of the TARP and the skeptical reaction in the Congress and much of the country to that TARP proposal. Note that Federal Reserve Board Chairman Bernanke and Secretary of the Treasury Hank Paulson gave testimony on Tuesday, September 23, to the Senate Banking Committee. The market turmoil significantly worsened in the following weeks. In the rollout of the TARP, people were warned by the government not only that "there is systemic risk" but also that "the Great Depression is coming." This scared people around the world and led to panic and a severe hit to the world economy.

Could it have been different? Could at least the chaotic pattern of these interventions been avoided? We can debate whether the intervention in the case of Bear Stearns was appropriate or not. I have my doubts, but let's put those doubts aside. The key question then pertains to the period *after* that intervention. It is not too difficult to imagine an environment in which the markets and the public in general would have been guided by a description by the Federal Reserve and the U.S. Treasury of the reasons behind the Bear Stearns intervention, as well as the direction and intentions of policy going forward. This sort of transparency would have given people some sense of policy actions to come. But no such description was provided.

Figure 6

Event Study of the Worsening Crisis: S&P 500 Index, Fall 2008

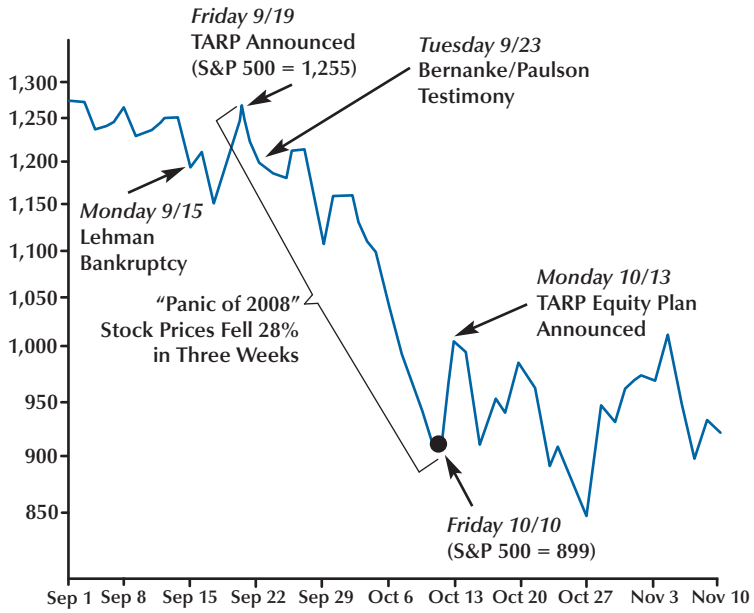


Table 1

Major Stock Price Indices During Fall 2008

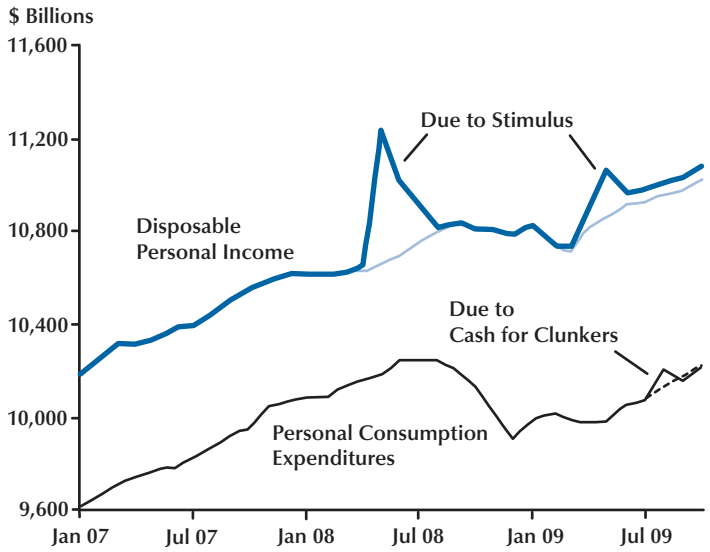
	S&P	FTSE	DAX	CAC	IBOVESPA	NIKKEI
September 12	1,252	5,417	6,235	4,332	52,393	12,215
September 15	1,192	5,204	6,064	4,169	48,419	11,609
September 19	1,255	5,311	6,134	4,324	53,055	11,921
October 10	899	3,821	4,544	3,176	40,829	8,276

NOTE: CAC, French stock market index (Cotation Assistée en Continu); DAX, German stock market index (Deutscher Aktien Index); FTSE, British stock market index (Financial Times Stock Exchange); IBOVESPA, Brazilian stock market index (Brazilian Índice Bovespa: Brazilian Bolsa de Valores do Estado de São Paulo [São Paulo Stock, Mercantile & Futures Exchange]); NIKKEI, stock market index for Tokyo Stock Exchange; S&P, Standard & Poor's U.S. stock index.

Figure 5 reveals something else that bolsters the case that uncertainty about the interventions made things worse. The turning point in the panic—measured by the LIBOR-OIS—occurred when uncertainty about the TARP was removed. Recall that the testimony on September 23, 2008, stated that the original purpose of the TARP was to buy up toxic assets on banks' balance sheets. People were skeptical about how that would work and government officials had difficulty explaining how it would work. Consequently, there was much uncertainty at the outset. The program itself was apparently not prepared very much in advance. But, after the TARP was changed and it was made clear on late Sunday, early Monday, October 13, 2008, that the funds would be used

Figure 7

Consumption and the Jump in Personal Disposable Income Due to the Fiscal Stimulus: Part 2 (Monthly Data, Seasonally Adjusted, Annual Rates)



to inject equity rather than buy toxic assets, conditions began to improve. You can see that this was the peak for the LIBOR-OIS spread, which continued to come down further.

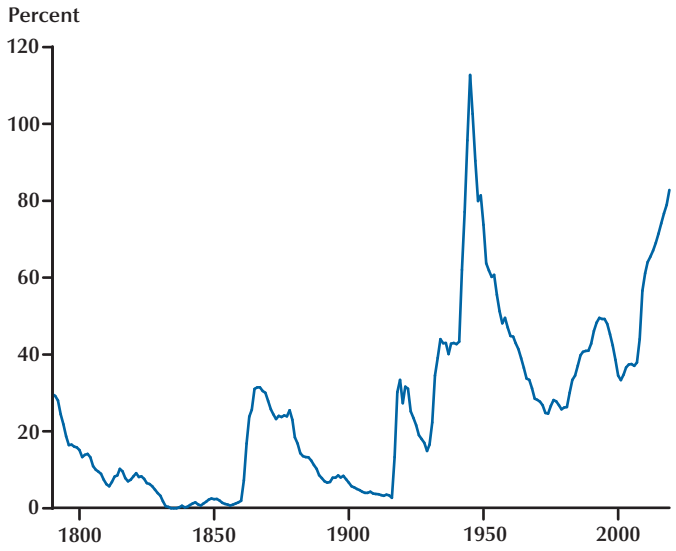
Other market measures show similar patterns. Figure 6 is the same type of event study as Figure 5 except it uses the S&P 500. Observe that the S&P 500 was higher the Friday after the Lehman bankruptcy than it was the Friday before. You can't prove causation with this timing of events, but it certainly suggests that the Lehman bankruptcy alone was not the cause of the panic. The sharp drop in the S&P 500 occurred much later. Moreover, the end of the panic in the stock market is on October 13, when the TARP equity plan was announced.

This panic quickly spread beyond the United States, as international data show. Table 1 shows major stock market indices around the world. The pattern is very similar to the United States. Equity prices came down on Monday, September 15, 2008, but were higher on Friday, September 19, after which they collapsed by 30 percent or so. Britain's FTSE behaves roughly this way and the story is the same for the German, French, and Japanese stock markets. It was a common story around the world. According to these data, the disruption does not seem to be as much due to the Lehman Brothers bankruptcy as it is to the series of policy responses.

What about other policy actions during the panic from late September into October? The panic is a complex period to analyze because many actions were taken at the same time, including the Fed's programs to assist money market mutual funds and the commercial paper market. These were intertwined with the FDIC bank debt guarantees and the clarification on October 13, after weeks of uncertainty, that the TARP would be used for equity injections. As discussed above, this clarification was a major reason for the halt in the panic in my view. Nevertheless, on the basis of conversations with traders and other market participants, the Fed's actions taken during the panic were also helpful in rebuilding confidence in money market mutual funds and the commercial paper market.

Figure 8

Federal Debt as a Percentage of GDP



Interventions After the Panic

Two other interventions were introduced by the Fed in the period following the panic: the program to purchase mortgage-backed securities (MBS) and the Term Asset-Backed Securities Loan Facility (TALF). The MBS program has turned out to be much larger, amounting to $1.25 trillion. My assessment of that program, based on research with Johannes Stroebel at Stanford, is that it had a rather small effect on mortgage rates once one controls for prepayment risk and default risk, but the estimates are uncertain. The TALF was very slow to start and it is still quite small.

On the fiscal side, interventions also continued. Figure 7 focuses on the impact of the second discretionary countercyclical stimulus package, which was passed in February 2009. Observe that the depiction here is simply an extension of Figure 4. Compared with the 2008 stimulus, the 2009 stimulus was larger, but the amount paid in checks was smaller and more drawn out. Nevertheless, there is still no noticeable effect on consumption. I also show the timing of the "Cash for Clunkers" program in Figure 7; it did encourage some consumption, but did not last and cannot be considered an effective method to stimulate the economy. In addition, my analysis of the government spending part of the stimulus is that it too had little positive impact.

The Legacy of the Interventions

Regardless of whether one thinks these interventions were bad or good, they have helped create huge legacies of debt, monetary overhang, and questionable policy precedents. First consider the deficit and the debt. To be sure, it is not only the crisis that has caused debt problems for the United States. Other powerful forces had been at work for some time. But the crisis has distracted us from efforts to deal with those forces.

Figure 8 shows the federal debt as a share of GDP, going back to the beginning of the United States of America. You can see the huge increase in the ratio during World War II, which fortunately was reversed in the years after the war. But we are heading up in that direction again. The CBO's projection through the next eight years shows a similarly huge increase in federal debt. This increase is partly due to the stimulus, partly due to the recession, and, more importantly now,

Figure 9

Federal Debt as a Percentage of GDP

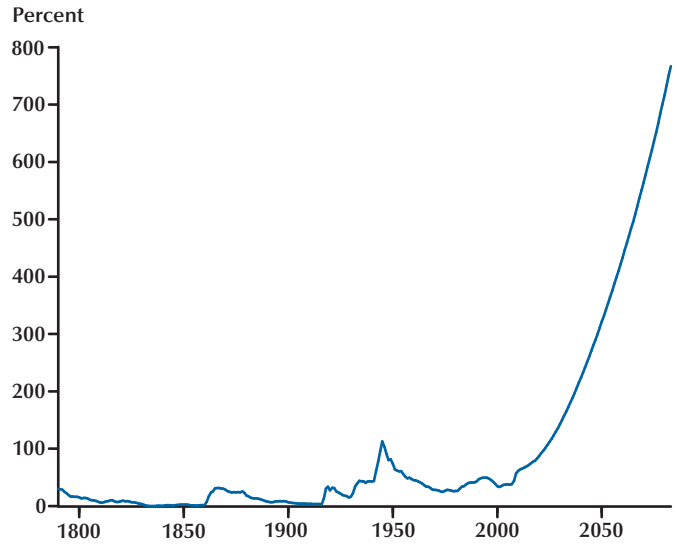
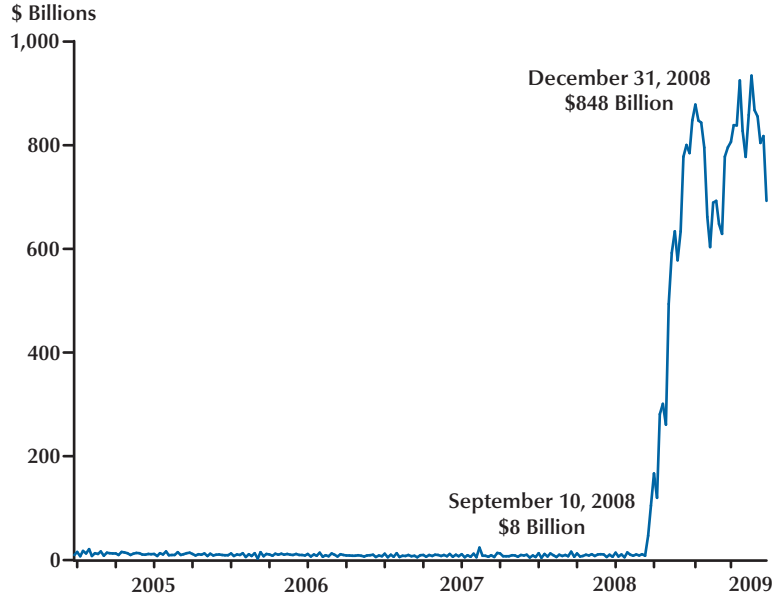


Figure 10

Reserve Balances of Depository Institutions at Federal Reserve Banks



partly due to the inability to rein in spending for our entitlement programs.

Figure 9 extends the data shown in Figure 8 for the next few decades. According to the CBO, the debt reaches 700 percent of GDP. This huge percentage dwarfs the debt the United States incurred after World War II. Of course, we all hope this is not going to happen; the United States of America would not be the United States of America if that were to happen. So something has got to give, and the question is *what*.

Whether one believes that the monetary policy actions worked or not, their consequences going forward are also negative. First, they raise questions about Fed independence. The programs are not monetary policy as conventionally defined, but rather fiscal policy or credit allocation policy financed through money creation and not by taxes or public borrowing. Why should such policies be run by an independent agency of government?

Second, unwinding the programs creates uncertainty. To wind down the programs in the current situation, the Fed must reduce the size of its MBS portfolio and reduce reserve balances. Figure 10 shows the huge size of reserve balances. The reserves rose because of the need to finance the Fed's interventions. Because there is uncertainty about how much impact the purchases have had on mortgage interest rates, there is uncertainty about how much mortgage interest rates will rise as the MBS are sold. There is also uncertainty about why banks are holding so many excess reserves. If the current level of reserves represents the amount banks desire to hold, then reducing reserves could cause a further reduction in bank lending.

Third, there is the risk of inflation. If the Fed is not able to reduce the size of the balance sheet as the economy recovers and as public debt increases, then inflationary pressures will undoubtedly increase.

POLICY IMPLICATIONS

In sum, this brief review of my research on the crisis shows that the government interventions taken before, during, and after the crisis did more harm than good. These interventions were a deviation from what was working well. We got off track. The policy implications are thus clear: Macroeconomic policy should get back on track.

For fiscal policy, this means avoiding further debt-increasing and wasteful discretionary stimulus packages, which do little to stimulate GDP. Ten years ago there was a near consensus that such programs were ineffective. Fiscal policy should focus on reducing the deficit and the growth of the debt-to-GDP ratio. Reforming existing entitlement programs to hold their growth down and limiting the creation of additional entitlement programs are essential.

For monetary policy, it means, as I testified at the House Financial Services Committee in March (Taylor, 2010b), returning to a policy with four basic characteristics: "First, the short-term interest rate (the federal funds rate) is determined by the forces of supply and demand in the money market. Second, the Fed adjusts the supply of money or reserves to bring about a desired target for the short-term interest rate; there is thus a link between the quantity of money or reserves and the interest rate. Third, the Fed adjusts the interest rate depending on economic conditions: The interest rate rises by a certain amount when inflation increases above its target and the interest rate falls by a certain amount when the economy goes into a recession. Fourth, to maintain its independence and focus on its main objectives of inflation control and macroeconomic stability, the Fed does not allocate credit or engage in fiscal policy by adjusting the composition of its portfolio toward or away from certain firms or sectors" (p. 4). Of course, this means we should exit from the MBS and other special programs as soon as possible. Obviously, we can't be draconian about this, but the sooner policymakers achieve this goal, the better future policy will be.

Some suggest that monetary policy has to do more things, such as taking actions to burst bubbles. Here let me say that I agree with the points made by James Bullard at the panel where these remarks were originally presented when he raised questions about whether policy can effectively pop bubbles and that it may do more harm than good. Our most successful past policy during the Great Moderation did not include such attempts to pop bubbles and the economy functioned very well.

REFERENCES

Bernanke, Ben S. "The Great Moderation." Remarks at the meetings of the Eastern Economic Association, Washington, DC, February 20, 2004.

Bernanke, Ben S. "Monetary Policy and the Housing Bubble." Annual Meeting of the American Economic Association, Atlanta, January 3, 2010.

Friedman, Milton and Heller, Walter. *Monetary vs. Fiscal Policy: A Dialogue*. New York: Norton, 1969.

Greenspan, Alan, "The Crisis." Paper prepared for the Brookings Panel on Economics Activity, March 9, 2010.

Taylor, John B. "Housing and Monetary Policy," in *Housing, Housing Finance, and Monetary Policy*, Federal Reserve Bank of Kansas City Symposium, Jackson Hole, WY, September 1, 2007.

Taylor, John B. "The Financial Crisis and the Policy Responses: An Empirical Analysis of What Went Wrong," in *A Festschrift in Honour of David Dodge's Contributions to Canadian Public Policy*. Ottawa: Bank of Canada, November 2008, pp. 1-18.

Taylor, John B. *Getting Off Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis*. Stanford: Hoover Press, 2009.

Taylor, John B. "The Fed and the Crisis: A Reply to Ben Bernanke." *Wall Street Journal*, January 11, 2010a.

Taylor, John B. "An Exit Rule for Monetary Policy." Testimony before the Committee on Financial Services, U.S. House of Representatives, March 25, 2010b.

Taylor, John B. and Williams, John C. "A Black Swan in the Money Market." *American Economic Journal: Macroeconomics*, January 2009, *1*(1), pp. 58-83.

Questions about Fiscal Policy: Implications from the Financial Crisis of 2008-2009

N. Gregory Mankiw

This article is a modified version of remarks given at the Federal Reserve Bank of Philadelphia's policy forum "Policy Lessons from the Economic and Financial Crisis," December 4, 2009. The presentation was made during a panel discussion that also included James Bullard and John Taylor.

Federal Reserve Bank of St. Louis *Review*, May/June 2010, *92*(3), pp. 177-83.

EVALUATING FISCAL POLICY

My objective here is to offer some reflections about recent fiscal policy. But I want to begin with an example, and I hope a useful analogy, about medicine.

Imagine you are a physician and a patient comes to you with some adverse symptoms. He's in pretty bad shape. You have never treated a condition quite like this before, and the causes of his ailments are not at all clear. You remember reading about a similar case in medical school, and so you try to recall as much as possible and come up with a theory as to what is making this patient sick. Then you choose the medicine that you hope will make the patient better.

What you would prefer to do is run a controlled experiment. You'd like to assemble 100 patients with similar conditions, give 50 of them the medicine you think might work and 50 of them a placebo, and then see whether the patients receiving the medicine recover. But you do not have 100 patients; you have only one. What do you do?

Well, you take your best shot. You decide what you believe to be the most likely cause of the patient's trouble and the most likely remedy to improve his health. Then you administer the medicine.

But the patient returns a few weeks later, and his symptoms are worse. What do you conclude? You might decide that you gave him the wrong medicine. Or you might decide that the patient is even sicker than you thought he was when he first came in and that you should increase the dosage.

If you consider the epistemological question, you realize that either of those conclusions is possible. The patient could have been sicker than you realized, and maybe you do have the right medicine. Or maybe you're on the wrong track completely. Because you have only one patient, you don't have any way to conduct that controlled experiment. So, you can't be sure.

This analogy describes the problem facing the Obama administration right now. They entered office when our economy was very sick. They concluded that, if they did nothing, the unemployment rate would reach 9 percent. (That was their forecast in the documents they released shortly after the election.) But they had a plan to stimulate the economy by spending a great deal of money, and with that stimulus—according to their estimates—the unemployment rate would not exceed 8 percent.

At the time of this writing, unemployment is more than 10 percent. So what do we make of that?

N. Gregory Mankiw is the Robert M. Beren Professor of Economics at Harvard University.

The administration has concluded that the patient must be sicker than they thought.

I should emphasize that the Obama administration has been guided by a standard theory in the economics profession—textbook Keynesian theory, to be specific. They thought that, because of the credit crisis, people were not able to obtain loans; and, because people were not able to obtain loans, there was insufficient aggregate demand. Looking to the standard textbook theory that fiscal policy can prop up aggregate demand, they implemented their stimulus plan.

In the forecast the administration used to describe the effects of the stimulus, they provided the multipliers that guided their actions. Their government-purchases multiplier was 1.57, and their tax multiplier was 0.99. Because 1.57 is larger than 0.99, they concluded it was better to spend money than to cut taxes. They obtained those numbers from a simulation of a standard macroeconometric model—the kind of model people have been building for years. It is the kind of model that Robert Lucas famously critiqued, but many economists have continued to build such models and use them for policy analysis and forecasting.

The question, ultimately, is whether the model is right, whether it accurately describes how the world works. If you could be sure the model was right—if God told you, "Yes, this model of the economy that you're simulating, it is truth"—then, when your stimulus plan was followed by 10 percent unemployment, you would know what to conclude: The patient was sicker than you thought, because, after all, the stimulus worked; my model said so. And if the stimulus worked, then unemployment would surely have been 11 percent if not for the stimulus. That is the position the administration is taking now.

Of course, we do not know if it is the right model. Macroeconomists have to be extremely humble. There is a lot we do not know.

I teach the *Principles of Economics* course at Harvard. It is a full-year course, and I start with what we economists are confident is true and then move on to material that is less and less certain as the year goes on. We look first at supply and demand, the theory of comparative advantage, profit maximization, marginal revenue equals marginal cost—the premises we agree on. Eventually, as the course goes on, we move to macroeconomics. We examine classical monetary theory, growth theory, and at the very end of the course the theory of business cycles, which is the topic we understand least of all.

I am actually a believer in Keynesian theory; much of my research is in that field. But even as a believer in many aspects of Keynesian theory, I appreciate that you cannot approach this subject matter without showing some humility about what we, as economists, can truly be confident about.

In an attempt to "know" as much as possible, the Obama administration is compiling data to measure the effect of the stimulus. (See www.recovery.gov, where you can find state-level job creation "data" reported to two decimals of accuracy!) This effort is, I think, the least credible part of the whole endeavor. The reporting errors are tremendous because no one accurately fills out these questionnaires with the true number of jobs they are creating.

My favorite story is about a firm that was selling boots to the U.S. Army; their managers decided that they were creating one job for every pair of boots they sold, because, after all, a soldier could not go to work without a pair of boots. This anecdote received some attention only because a reporter looked through the job-creation numbers and discovered that this particular organization was the most efficient job creator in the country: For every $100 spent, they created a job!

Putting aside these absurd reporting errors, even if the reporting were perfectly correct, the whole activity still makes no sense. When we talk about the effects of government purchases on aggregate demand, and therefore on job creation, there is an array of general equilibrium effects ("knock-on effects") that are tremendously important—some positive, some negative. These job-creation surveys cannot possibly capture these effects.

The positive effects are those that arise from the conventional Keynesian fiscal-policy multipliers. Higher government spending leads to higher income, which causes higher consumption and therefore higher income yet again. But

recovery.gov cannot possibly account for all the multiplier effects involved.

There are also growth-retarding effects associated with government spending. If people observe substantial government debt being issued, they may anticipate higher future taxes and therefore cut back on their current consumption. These retarding effects are also absent from recovery.gov. In addition, there are crowding-out effects working through financial markets. Increased government borrowing drives up long-term interest rates and reduces spending today. But, given its limited scope, recovery.gov cannot take that into account either.

So even if the administration could accurately measure what they set out to measure, the data they created would not accurately describe, from a macroeconomist's standpoint, how many jobs were created.

TAXING LESS VERSUS SPENDING MORE

The larger question for me is this: Would we have been better off focusing on the tax side of fiscal policy rather than on the government spending side? I think there are several issues here to consider.

The first question is whether the government can spend large sums of money quickly and wisely. To consider the question, I'll offer a personal story. I live in a suburb of Boston called Wellesley. Coincidentally, my town has been debating building a new high school, which is now being constructed just a few blocks from where I live. The project is creating many jobs and, in fact, seems to have been planned and implemented in an intelligent and prudent way. Participating in such a process makes you realize how much time it takes to accomplish something this substantial. The town spent many months debating whether to build a new high school or renovate the old high school. Once they decided that issue, they spent many months designing the new school and determining everything they wanted in a school. They spent many more months selecting the site to build on and which houses to take over to obtain the land. It was a years-long process.

Now, what if someone had provided all the funds for a new Wellesley high school but demanded that it be built immediately? Presumably, quality would have been sacrificed for speed. It would have been built more quickly because, presumably, if people are told they have to spend money immediately, they will find some way to do it. But in such an environment—town planning or national fiscal policy—you wonder whether money can be spent both quickly and wisely.

The second question, which is more academic in nature, is about the size of the relevant multipliers. The textbook Keynesian model tells us that government purchases multipliers are larger than tax multipliers. And, again, the Obama administration's economic team consulted these standard models in reaching their conclusions. However, there is a variety of evidence that calls those conclusions into question.

Somewhat ironically, one piece of evidence against this preference for government spending over cutting taxes comes from Christina Romer, the current chair of the Council of Economic Advisors for President Obama. About six months before she took that job, she wrote a paper with her husband David Romer about the impact of tax policy on the economy (Romer and Romer, 2009). The Romers wanted to measure the influence of taxes on gross domestic product (GDP), ensuring that they identified the exogenous movements in taxes and separating those from the endogenous movements, where taxes were responding directly to the economy.

Their conclusion was that the tax multiplier was 3—that is, every dollar spent on tax cuts would raise GDP by $3. That is roughly three times the size that the Obama administration assumed with their policy simulations. To be clear, I don't blame the Obama administration for relying on more conventional multiplier estimates. Nor is it reasonable to assume that simply because they appointed Christina Romer to chair the Council that they must use her research in measuring the multiplier. But the Romers' research does suggest that conventional results about tax policy on GDP are understated.

Of course, it could be the case—and this is in fact the Obama administration's interpretation—

that *all* multipliers may be larger than previously measured. Fiscal policy may be so potent that, if the tax multiplier is 3, the government's spending multiplier might be 4 or 5. The Romers did not analyze government spending multipliers in their recent study, but only tax multipliers. Clearly, it is still an open question. Yet, there has been a variety of research on government spending multipliers using techniques similar to those used by the Romers. This research has tried to uncover exogenous movements in government spending, and no one finds government spending multipliers to be especially large.

The best work on this topic, I think, is from Valerie Ramey at the University of California at San Diego (Ramey, 2009). Ramey finds government spending multipliers of about 1.4, which is not very different from what the Obama administration assumed, but much smaller than the Romers found for taxes. Similarly, Andrew Mountford and Harold Uhlig, using vector autoregression techniques, have also found that taxes have a more potent effect than government spending (Mountford and Uhlig, 2009).

The piece of evidence I want to draw your attention to in particular, though, is some very recent work by my colleagues Alberto Alesina and Silvia Ardagna at Harvard. They used data from the Organisation for Economic Co-operation and Development (OECD) to identify every major fiscal stimulus in those countries (Alesina and Ardagna, 2009). They then separated out the successful policies—those that in fact were followed by robust economic growth—from the unsuccessful ones and compared their characteristics. They found that the successful stimulus packages cut business and income taxes and the unsuccessful stimulus packages increased government spending and transfer payments.

The data in the Alesina-Ardagna study are mostly European, with only a small portion from the United States. But they lead to conclusions that are very similar to those from Mountford and Uhlig's work using U.S. data in vector autoregressions. These conclusions are also consistent with the work from Ramey and the Romers that looks at the historical record to identify multipliers. A growing body of evidence seems to suggest that taxes may be a better tool for fiscal stimulus than conventional models have indicated.

TAXES AS FISCAL POLICY TOOL

What, then, is behind these conclusions that taxes have a more potent effect on the economy than spending? The answer is not clear-cut, but it is easy to speculate why this could be true. Most obviously, there are effects on the supply side. Tax rates influence work incentives, for instance. But even if you believe that aggregate demand drives the economy in the short run, as many Keynesians do, you might consider that taxes affect aggregate demand in ways that are not included in the textbook Keynesian model.

When we change taxes, we typically do not just write checks to taxpayers. Usually, we change marginal tax rates: We change corporate income taxes, change personal income taxes, and maybe even institute an investment tax credit. These measures have more complicated and nuanced effects on aggregate demand than what the textbook Keynesian model assumes. It is not simply a change in cash flow; it is actually a change in marginal incentives and can even be a direct encouragement to spend. One example is creating tax incentives to invest.

The Cash for Clunkers program involved that type of incentive. I was not much in favor of that specific micromanagement of how people should spend their money. Nonetheless, the fact that people responded to a tax incentive like Cash for Clunkers does suggest that a more comprehensive program (such as an investment tax credit) might have stimulated spending even more broadly.

Many other tax policies have been discussed recently. One in particular has received some attention: a tax cut for new hires. As many have pointed out, the premise behind this policy is that, because unemployment is so high even as we proceed through the recovery, we should create incentives for businesses to hire new workers.

There is a case to be made for a payroll tax cut. In fact, at one point I advocated an immediate and permanent payroll tax cut financed by a gradual increase in gasoline taxes over time. But

a tax cut for new hires is probably not a good idea. The basic problem is that we do not know how to properly define—or enforce a definition of—a "new hire." Presumably, we do not want a business to hire Peter by firing Paul and call Peter a new hire; that would cause a great deal of inefficient churning in the labor force.

Usually, when tax credits for new hires are proposed, the idea is to establish some baseline employment and give credit to businesses that meet or exceed that baseline. But even establishing a baseline has its limitations. Consider an industry hit particularly hard by a recession—say, construction—in which employment is well below the baseline established for the tax breaks. Because a few new hires would not make these firms eligible for the tax breaks, these firms would have no marginal incentive to hire additional workers. Conversely, industries that have been expanding would be rewarded for hires they may have made even without the tax incentives. This policy, then, would likely create tremendous disparities across industries that could be both inequitable and inefficient.

There is also the problem of new firms. New firms are always a large part of economic growth and the overall dynamics of the economy. By definition, all employees of a new firm are "new hires." But if there is a tax credit for new hires at new firms, then that provides all sorts of incentives for existing firms to, say, lay off the janitorial staff and hire instead an independent janitorial contractor that just started up as a new firm. The bottom line is that it is very difficult to implement a tax credit for new hires, attractive as the idea seems at first.

THE LONG-RUN FISCAL SITUATION AND THE HEALTHCARE CHALLENGE

Any discussion of fiscal policy has to be couched in terms of the long-run fiscal picture. I know that some economists have made the argument that we should not be concerned about the long-run fiscal picture when we consider short-run fiscal policy: Once the economy begins to grow again, they contend, tax revenue will flow in and the longer-term fiscal picture will be improved. That would be true if these policy multipliers were very large, but they are probably not so large, in my judgment, that we can ignore the long-run problems created by short-run policies.

We now face a very dire long-run fiscal picture, which is being driven by a couple of factors. One is the aging of the population, with the first wave of Baby Boomers beginning to retire. I explain this scenario to my students at Harvard this way: "My generation has promised ourselves generous retirement benefits in the form of Social Security and Medicare, and we promise you're going to pay for it. How do you feel about that?"

The bills for those benefits are large partly because of aging and partly because of higher healthcare costs. The latter is one of the motivations, allegedly, for the current healthcare legislation. I am personally skeptical that this legislation will reduce healthcare spending substantially—if at all. So on that score, we are not making true progress on the long-run fiscal picture. To address the issue, I would raise the age of eligibility for Social Security and Medicare. Economists seem to like this idea, but in polls of the general public it is much less popular, so I don't expect this idea to be implemented.

I should note that I am not optimistic that any proposed measure will reduce healthcare costs significantly. People talk about "bending the curve" and squeezing out waste, fraud, abuse, and so on. My reading of the evidence brings me to the conclusion that healthcare is growing more expensive over time mainly because technologies are improving. That is a good thing, but it is also expensive, and so we must find a way to pay for it. I am not sure what the right answer is, but I don't think that rooting out waste, fraud, and abuse is going to save much money. I am sure there is some waste, fraud, and abuse—there is in all systems—but it is not likely a primary driver of healthcare costs.

One of the classic hypothetical questions economists ask when referring to healthcare costs is, "Would you rather go back to 1950s medical care and 1950s prices?" If that option were offered at your place of work, my guess is that you would not take it. What that means is, in some real sense,

healthcare is cheaper today if you adjust prices properly to account for quality improvements. A dollar of healthcare today has more value than a dollar of healthcare in 1950.

I posed another hypothetical question in a *New York Times* column in September 2009: Imagine a Dorian Gray pill has been invented and that taking the pill every day will keep you the same age for as long as you take the pill. It is perfect healthcare: You would not age or become sick or die. The problem is that these pills are expensive—let's say, it costs $1,000 to manufacture each daily dose. Again, it is perfect healthcare but at a very, very expensive price. How would we, as a society, deal with that? I don't think we have the answer to that question, but in some sense we are moving in that direction already, with health technology continually improving but also becoming more and more expensive. As a society, we have not figured out how we are going to eventually say no to people or have people say no for themselves.

THE COMING VALUE-ADDED TAX?

I have my own normative conclusions about addressing some of these problems, such as raising the retirement age. What I believe is more likely to occur, however, is that taxes will be raised to very high levels. I thought it was very interesting when, in late 2009, Nancy Pelosi suggested the idea of a value-added tax.

A value-added tax (VAT) is an efficient tax from an economic standpoint. It is basically a flat consumption tax, so it tends to be an efficient way to raise revenue. But it is also a fairly well hidden tax, and there is some debate among proponents and opponents about that aspect.

In particular, people look at European countries and see the connection between their large governments and their VATs and conclude that the problem with the VAT is that it makes government grow too much. That is one possibility. Another possibility is that governments grow, and, when they do so, they look for ways to raise revenue efficiently. In many cases, they turn to a VAT. My guess is that the latter is probably the correct direction of causation and that it is also probably the direction in which we're heading: larger government and higher taxes. If we use the tax system we have now, though, the revenue raised will likely fall short; hence, we are probably going to move in the direction of a VAT.

Such a large change in our tax policy would inevitably incite comparisons with Europe. There is a literature about how and why the European workforce differs from the American workforce—specifically, why Europeans enjoy spending more time at the café than Americans do and why we work harder than they do. There are many hypotheses out there. Olivier Blanchard says that it stems from cultural tastes (Blanchard, 2004): Europeans have more joie de vivre than Americans, and therefore they want to enjoy their high productivity by spending more time enjoying leisure. My colleagues Alberto Alesina and Ed Glaeser, as well as Bruce Sacerdote from Dartmouth (Alesina, Glaeser, and Sacerdote, 2005), say that it is the presence and scope of powerful labor unions in Europe that have negotiated shorter workweeks, more vacation days, and so on. But Ed Prescott tells us it is the high tax rates in Europe (Prescott, 2004), and I actually find this argument the most compelling. What that means is, if we are heading toward higher tax rates, my children will enjoy a lot more leisure than I do.

REFERENCES

Alesina, Alberto and Ardagna, Silvia. "Large Changes in Fiscal Policy: Taxes versus Spending." Working paper, Harvard University, October 2009.

Alesina, Alberto; Glaeser, Edward L. and Sacerdote, Bruce. "Work and Leisure in the United States and Europe: Why So Different?" *NBER Macroeconomics Annual*. Volume 20. Chicago: University of Chicago Press, 2005, pp. 1-64.

Blanchard, Olivier. "The Economic Future of Europe." *Journal of Economic Perspectives*, Fall 2004, *18*(4), pp. 3-26.

Mountford, Andrew and Uhlig, Harald. "What Are the Effects of Fiscal Policy Shocks?" NBER Working Paper No. 14551, National Bureau of Economic Research, December 2008.

Prescott, Edward C. "Why Do Americans Work So Much More than Europeans?" Federal Reserve Bank of Minneapolis *Quarterly Review*, July 2004, *28*(1), pp. 2-13; http://minneapolisfed.org/research/qr/qr2811.pdf.

Ramey, Valerie. "Identifying Government Spending Shocks: It's All in the Timing." Working paper, University of California at San Diego; NBER Working Paper No. 15464, National Bureau of Economic Research, October 2009.

Romer, Christina D. and Romer, David H. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." Working paper, University of California at Berkeley, April 2009 (forthcoming in *American Economic Review*); http://emlab.berkeley.edu/users/dromer/papers/RomerandRomer.pdf.

Nonlinear Effects of School Quality on House Prices

Abbigail J. Chiodo, Rubén Hernández-Murillo, and Michael T. Owyang

We reexamine the relationship between quality of public schools and house prices and find it to be nonlinear. Unlike most studies in the literature, we find that the price premium parents must pay to buy a house in an area associated with a better school increases as school quality increases. This is true even after controlling for neighborhood characteristics, such as the racial composition of neighborhoods, which is also capitalized into house prices. In contrast to previous studies that use the boundary discontinuity approach, we find that the price premium from school quality remains substantially large, particularly for neighborhoods associated with high-quality schools. (JEL C21, I20, R21)

Federal Reserve Bank of St. Louis *Review*, May/June 2010, *92*(3), pp. 185-204.

The relationship between house prices and local public goods and services has been widely studied in the literature, dating back to Oates's (1969) seminal paper, in which he studied the effect of property tax rates and public school expenditures per pupil on house prices. Oates conjectured that if, according to the Tiebout (1956) model, individuals consider the quality of local public services in making locational decisions, an increase in expenditures per pupil should result in higher property values, whereas an increase in property tax rates would result in a decline in property values, holding other things equal across communities. Oates suggested that the variation in expenditures per pupil partially reflected the variation in the quality of public schools.

In the analysis of school quality, researchers have often applied the hedonic pricing model developed by Rosen (1974). In this model, the implicit price of a house is a function of its comparable characteristics, as well as measures of school quality and a set of neighborhood characteristics. A house's comparable characteristics include the number of bedrooms, square footage, and so on. The estimated coefficients from the regression represent the capitalization of the different components into house values.

In an influential study, Black (1999) argued that previous research estimating hedonic pricing functions introduced an upward bias from neighborhood quality effects that are unaccounted for in the data.[1] Specifically, she noted that better schools may be associated with better neighborhoods, which could independently contribute to higher house prices. Black circumvented this problem by estimating a linear hedonic pricing function using a restricted sample of data from

[1] By neighborhood *quality* we refer to the availability of mass transit and thoroughfares, proximity to commercial and industrial areas, and other such amenities, in addition to sociodemographic characteristics.

Abbigail J. Chiodo is a former research analyst at the Federal Reserve Bank of St. Louis. Rubén Hernández-Murillo is a senior economist and Michael T. Owyang is a research officer at the Federal Reserve Bank of St. Louis. Jeremy Bixby, Katie Caldwell, Kristie M. Engemann, Christopher Martinek, Mark L. Opitz, and Deborah Roisman provided research assistance. The authors acknowledge First American (Real Estate Solutions) for house price data and technical support.

houses along the boundaries of school attendance zones.[2] She rationalized that, while test scores make a discrete jump at attendance boundaries, changes in neighborhoods are smoother.[3] The linear specification of the hedonic approach, including Black's (1999) variation, presupposes that the marginal valuation of below-average schools is equal to the valuation of above-average schools and results in a constant premium on school quality.[4]

In this paper, we argue that the relationship between school quality and house prices in the boundary discontinuity framework is better characterized as a *non*linear relationship. We formulate motivating hypotheses for the presence of nonlinear effects of school quality on house prices based on heterogeneous parent valuations of school quality and competition in the housing market. We then test for nonlinear effects estimating a nonlinear pricing function in the St. Louis, Missouri, metropolitan area, using standardized state math test scores as the measure of education quality. To control for neighborhood quality, we measure education capitalization by using Black's method of considering only houses located near attendance zone boundaries. We find that the effect of school quality is indeed best characterized as a nonlinear function.

We find, as did Black (1999), that controlling for unobserved neighborhood characteristics with boundary fixed effects reduces the premium estimates from test scores relative to the hedonic regression with the full sample of observations. We also find, however, that the linear specification for test scores underestimates the premium at high levels of school quality and overestimates the premium at low levels of school quality. In contrast to Black (1999) and many subsequent studies in the literature, we find that the effects of school quality on housing prices remain substantially large even after controlling for neighborhood demographics, such as the racial composition of neighborhoods, in addition to boundary fixed effects. We also find that the racial composition of neighborhoods has a statistically significant effect on house prices.

This paper is organized as follows. The next section presents a survey of the recent literature. We then describe the hypotheses and the econometric model. Our data description is followed by the empirical results.

LITERATURE REVIEW

Ross and Yinger (1999) and Gibbons and Machin (2008) provide surveys of the literature on capitalization of local public goods and services. Examples of the traditional full-sample hedonic regression approach include papers by Haurin and Brasington (1996), Bogart and Cromwell (1997), Hayes and Taylor (1996), Weimer and Wolkoff (2001), and Cheshire and Sheppard (2002). Additional works are surveyed in Sheppard (1999).

Various studies in the hedonic analysis tradition have used so-called input-based measures of education quality, such as per-pupil spending. Hanushek (1986, 1997) found that school inputs have no apparent impact on student achievement and are therefore inappropriate as measures of school quality. His insights have led to the more prevalent use of *output-based* measures, such as standardized test scores.[5] The research on education production functions also has made the case that *value-added* measures of achievement—often measured as the marginal improvement in a particular cohort's performance over a period of time—would be more appropriate as measures of quality in capitalization studies. However, con-

[2] A school's attendance zone delimits the geographic area around the public school the residents' children would attend. In this text, we often refer interchangeably to a school's attendance zone as the *school*, but this term should not be confused with *school district*, which is an administrative unit in the public school system often comprising several schools.

[3] Black's (1999) boundary discontinuity approach is part of the more general *regression discontinuity design* surveyed by Imbens and Lemieux (2008).

[4] Nonlinear effects are nevertheless routinely allowed among some house characteristics, such as the number of bathrooms and the age of the building.

[5] Some authors, however, have expressed concerns about the potential endogeneity of school quality when it is measured by indicators of student performance. Gibbons and Machin (2003), for example, argue that better school performance in neighborhoods with high house prices may reflect that wealthy parents buy bigger houses with more amenities and therefore devote more resources to their children.

structing value-added measures requires tracking groups of students over time and implies more sophistication in the decisionmaking process of potential buyers, as value-added measures are not commonly available to the public. Brasington (1999), Downes and Zabel (2002), and Brasington and Haurin (2006) found little support for using value-added school quality measures in the capitalization model; they argued that home buyers favor, in contrast, more traditional measures of school quality in their housing valuations.

A prevalent concern of capitalization studies is the possibility of omitted variable bias, induced by failing to account for the correlation between school quality and unobserved neighborhood characteristics, as better schools tend to be located in better neighborhoods. As mentioned previously, Black (1999) tackled this problem by restricting the sample to houses near the boundaries between school attendance zones and controlling for neighborhood characteristics with boundary fixed effects. A rudimentary precursor of this idea was analyzed by Gill (1983), who studied a sample of houses in Columbus, Ohio, restricting observations to neighborhoods with similar characteristics. Also, Cushing (1984) analyzed house price differentials between adjacent blocks at the border of two jurisdictions in the Detroit, Michigan, metropolitan area. Recent examples of this approach include studies by Leech and Campos (2003), Kane, Staiger, and Samms (2003), Kane, Staiger, and Riegg (2005), Gibbons and Machin (2003, 2006), Fack and Grenet (2007), and Davidoff and Leigh (2007).

The boundary discontinuity approach has been criticized in some recent studies motivated primarily by concerns about the successful removal of any remaining omitted spatial fixed effects (Cheshire and Sheppard, 2004) or the possibility of discontinuous changes in neighborhood characteristics, which also depends on the definition of "neighborhood" that is adopted (Kane, Staiger, and Riegg, 2003; Bayer, Ferreira, and McMillan, 2007). However, barring the availability of repeat sales data or information on boundary redistricting or policy changes to supply the exogenous variation required for identification, in the case of stable boundary definitions and cross-sectional data, the boundary discontinuity approach remains a useful methodology. In addition to boundary discontinuities, recent studies have used various methods of addressing the omitted variables and endogeneity issues, including time variation (Bogart and Cromwell, 2000; Downes and Zabel, 2002; Figlio and Lucas, 2004; Reback, 2005, among others), natural experiments (Bogart and Cromwell, 2000, and Kane, Staiger, and Riegg, 2005), spatial statistics (Gibbons and Machin, 2003, and Brasington and Haurin, 2006), or instrumental variables (Rosenthal, 2003, and Bayer, Ferreira, and McMillan, 2007).

In this paper, we measure school quality at the individual school level and we regress house prices on their physical characteristics and a full set of pairwise boundary dummies to control for unobserved neighborhood characteristics. Additionally, in response to the criticisms of the boundary discontinuity approach, we augment the estimation by controlling for a set of demographic characteristics defined at the Census-block level (as opposed to the larger block groups or tracts). Many papers that do not use the boundary discontinuity approach measure education quality at the school-district level, as opposed to considering schools individually. These studies also face the challenge of devising appropriate definitions of neighborhoods to match the geographic level at which school quality is measured. For example, Clapp, Nanda, and Ross (2008) measure school quality at the school-district level and use Census-tract fixed effects to control for omitted neighborhood characteristics. Brasington and Haurin (2006) also measure school quality at the school-district level but use spatial statistics rather than fixed effects to control for neighborhood characteristics.

To the best of our knowledge, nonlinear hedonics from school quality have been explored only by Cheshire and Sheppard (2004) in a study of primary and secondary schools in the United Kingdom. They estimate a full-sample, standard hedonic regression modified to include Box-Cox transformations of house prices, house characteristics, and measures of school quality. Their evidence suggests that the price-quality relationship is highly nonlinear. Although Cheshire and

Sheppard include a wide variety of local neighborhood characteristics as controls, their approach also suffers from the possibility of omitted variable bias present in traditional hedonic models.

A previous study of house prices in the St. Louis metropolitan area by Ridker and Henning (1967) found no evidence of education capitalization in St. Louis house prices. Although their main concern was to determine the negative effect of air pollution on housing prices, they included a dummy variable that indicated residents' attitudes about the quality of the schools (above average, average, and below average). Ridker and Henning (1967) acknowledged, however, that their study may suffer from small-sample bias that could explain this seemingly contradictory finding. Kain and Quigley (1970) also conducted an early study of the components of a hedonic price index for housing in the St. Louis metropolitan area, but it does not consider measures of school quality.

THE MODEL

In this section, we discuss three motivating hypotheses that can generate nonlinear effects from school quality on house prices. We argue that the nonlinearity with respect to school quality illustrates two aspects of the market for public education that are reflected in the housing market. Although developing a full theoretical model is beyond the scope of our paper, interested readers are referred to a previous working paper version in which we sketch a search model of the housing market in the spirit of Wheaton (1990) and Williams (1995) that can motivate these features.

Three Arguments for Nonlinear Effects

First, in an environment in which potential buyers are heterogeneous in the intensity of their preferences for school quality and neighborhood characteristics, buyers with a stronger preference for education quality may concentrate their buying search for a house in the highest-quality attendance zones. As school quality increases, competition from other buyers creates an increasingly tight housing market, because the housing supply in these areas is often very inelastic, as most metropolitan areas have a fixed housing stock in the short run.

This argument is similar to that proposed by Hilber and Mayer (2009). They argue that scarcity of land confounds identification of the education premium. Brasington (2002) and Hilber and Mayer (2009) have also noted that the extent of capitalization in a hedonic framework may vary depending on whether houses are located near the interior or the edge of an urban area. They find that capitalization is weaker toward the edge, where housing supply elasticities and developer activity are greater.

Second, alternative schooling arrangements (e.g., private schools, home schooling, magnet schools) can provide home buyers with high-quality education even if they choose to live in lower-quality public school attendance zones, allowing for a reduced price premium in these neighborhoods. The existence of these options underlies our belief that a constant premium across the range of school quality is not realistic.

The previous two hypotheses rely on the heterogeneity of preferences for school quality and neighborhood characteristics among the population of prospective home buyers, a feature widely documented in the literature. Bayer, Ferreira, and McMillan (2007), for example, argue that there is a considerable degree of heterogeneity in homeowners' preferences for schools and racial composition of neighborhoods.

Finally, an alternative hypothesis that can generate nonlinearities is that school quality can be considered a luxury good; therefore, at higher-quality schools (and therefore richer neighborhoods), people would be willing to pay more for the same marginal increase in school quality.

The Econometric Model

We now estimate a model of house prices. Specifically, we estimate the dollar value difference in home prices for a quantified increase in school quality. We discuss three alternative specifications that include two different identification techniques to disentangle neighborhood quality from school quality.

Pure Hedonic Pricing Model. As a benchmark, we introduce a hedonic pricing equation

in which the sale price is described as a function of the characteristics of the house and its location-specific attributes, including the quality of the school associated with it. The basic hedonic function can be described as follows:

$$(1) \qquad \ln\left(p_{iaj}\right) = \kappa + \mathbf{X}_{\mathrm{i}}'\boldsymbol{\beta} + \mathbf{Z}_{\mathrm{j}}'\boldsymbol{\delta} + \mu_a\psi^H + \varepsilon_{iaj},$$

where p_{iaj} is the price of house i in attendance zone a in neighborhood j. The vector $\mathbf{X}_{\mathrm{i}}$ represents the comparable aspects of house i (e.g., the number of bedrooms, bathrooms, and so on) and vector $\mathbf{Z}_{\mathrm{j}}$ represents local characteristics. The value μ_a is the quality of the school in attendance zone a. In this paper, we measure school quality with an index constructed from test scores, defined at the school level and expressed in standard deviations (SDs) from the mean. The quantity of interest ψ^H is the education capitalization premium and represents the percentage increment in house prices from increasing school test scores by 1 SD.

Thus, the house price reflects all relevant attributes; that is, the physical and location-specific characteristics of the home are capitalized into the house value even if they are not directly consumable by the current tenants (because of their effects on the resale value of the house).[6] One potential problem with this specification is that the comparable house characteristics, $\mathbf{X}_{\mathrm{i}}$, do not fully capture the quality of the house (updates, condition, landscaping, layout, and so on), the quality of the surrounding neighborhood, and various other factors. The hedonic pricing function attempts to capture these factors with the inclusion of the $\mathbf{Z}_{\mathrm{j}}$ vector. The success with which the model captures these unobserved factors often depends on how coarsely the geographic area encompassed by $\mathbf{Z}_{\mathrm{j}}$ is defined (i.e., for how small a vicinity around the house $\mathbf{Z}_{\mathrm{j}}$ provides variation).

Linear Boundary Fixed Effects Model. As discussed earlier, the methodology of adding the location characteristics vector, $\mathbf{Z}_{\mathrm{j}}$, may reduce but not entirely account for all of the variation that can be introduced on a neighborhood level. Suppose that the neighborhood characteristics gradient is large in absolute value. This implies that houses a few blocks away from each other can vary a great deal in "atmosphere" and, therefore, in price. This variation can be related to distance to amenities, mass transit, and thoroughfares (i.e., highway access), proximity to commercial and industrial zoning, single-family housing density, and so on. The vector $\mathbf{Z}_{\mathrm{j}}$ may be unable to account for all the unobserved neighborhood variation that confounds the estimate of the capitalization premium because of the potential correlation with school quality. Much of this variation (though admittedly not all) can be corrected for by analyzing houses that are geographically close.

The boundary discontinuities refinement considers only houses that are geographically close to school attendance zone boundaries and replaces the vector of local characteristics with a full set of pairwise boundary dummies. Each house in this reduced sample is associated with the nearest, and hence unique, attendance zone boundary. This yields the following:

$$(2) \qquad \ln\left(p_{iab}\right) = \kappa + \mathbf{X}_{\mathrm{i}}'\boldsymbol{\beta} + \mathbf{K}_{\mathrm{b}}'\boldsymbol{\varphi} + \mu_a\psi^L + \varepsilon_{iab},$$

where $\mathbf{K}_{\mathrm{b}}$ is the vector of boundary dummies and the subscript b indexes the set of boundaries. The resulting education premium calculated with the linear boundary fixed effects model is ψ^L. Equation (2), then, is equivalent to calculating differences in house prices on opposite sides of attendance boundaries while controlling for house characteristics and relating the premium to test-score information.

The boundary dummies allow us to account for unobserved neighborhood characteristics of houses on either side of an attendance boundary because two homes next to each other generally would have the same atmosphere. For this approach to be successful, particular care must be taken to exclude from the sample attendance zones whose boundaries coincide with administrative boundaries, rivers, parks, highways, or other landmarks that clearly divide neighborhoods, as neighborhood characteristics in these cases would be expected to vary discontinuously at the boundary.

[6] For example, if the current tenants have no school-aged children.

Nonlinear Boundary Fixed Effects Models. As an alternative to the linear model, we consider the possibility that the capitalization premium is not constant over the range of school qualities. This is accomplished by testing whether the education capitalization term enters nonlinearly. Consider the following pricing equation:

$$\ln(p_{iab}) = \kappa + \mathbf{X}_i'\boldsymbol{\beta} + \mathbf{K}_b'\boldsymbol{\varphi} + f(\mu_a) + \varepsilon_{iab}, \quad (3)$$

where $f(\mu_a)$ represents a potentially nonlinear function of school quality. For simplicity, suppose the function $f(\mu_a)$ is composed of a linear polynomial term and higher-order polynomial terms in school quality. That is,

$$f(\mu_a) = \psi_1\mu_a + \psi_2\mu_a^2 + \psi_3\mu_a^3, \quad (4)$$

where ψ_m, $m = 1,2,3$, are scalar parameters. We then rewrite equation (3) as

$$\ln(p_{iab}) = \kappa + \mathbf{X}_i'\boldsymbol{\beta} + \mathbf{K}_b'\boldsymbol{\varphi} + \psi_1\mu_a + \psi_2\mu_a^2 + \psi_3\mu_a^3 + \varepsilon_{iab}. \quad (5)$$

Specification (5) offers several advantages over the linear form (equation (2)). First, the rate at which the nominal premium varies across the range of school quality is not fixed. This allows us to differentiate the incremental effects on house prices of low- versus high-quality school attendance zones. Second, with a constant premium the linear model *penalizes* houses in low-quality school attendance zones by valuing them below what would be predicted by their comparable attributes.[7] Moreover, the penalty increases as the school quality worsens. This scenario is unappealing because, as mentioned before, potential buyers who value education quality often can find substitute arrangements outside the public school system. Our prediction is that houses in lower-quality attendance zones command a smaller premium; in other words, the price function should be flatter for areas with lower test scores and steeper for those with higher test scores. This possibility is explicitly excluded in the linear model.

[7] We adopt the convention that an increase in school quality induces a *premium* on house prices, whereas a decrease in school quality imposes a *penalty* on house prices.

A Note on the Estimation. We estimated regression equations (1), (2), and (5) with ordinary least squares. In all cases, we computed robust standard errors (SEs) clustered at the school level. For completeness, the "Results" section also presents the estimation of the nonlinear models using the full sample. We included boundary dummies in the regression equation and estimated the coefficients for these variables directly.

In an attempt to reduce any remaining bias from omitted characteristics, some recent studies, such as that by Bayer, Ferreira, and McMillan (2007), have supplemented their analysis by including demographic controls in the regressions. We therefore present results of the boundary fixed effects regressions in which the vector $\mathbf{Z}_j$ of neighborhood characteristics has been reinserted in the estimation. In particular, we control for the racial composition of neighborhoods. Studies that specifically consider the racial composition of neighborhoods include those by Bogart and Cromwell (2000), Downes and Zabel (2002), Cheshire and Sheppard (2004), Kane, Staiger, and Riegg (2005), Reback (2005), Clapp, Nanda, and Ross (2007), and Bayer, Ferreira, and McMillan (2007).

DATA

In this analysis, we restrict our attention to single-family residences and elementary school attendance zones. Each observation corresponds to a house and is described by variables reflecting its physical characteristics, the quality of the local public elementary school that children in the household would attend, and the characteristics of the neighborhood in which the house is located—namely, demographic indicators measured at the Census-block level and property tax rates measured at the school-district level.

Real Estate Prices and Housing Characteristics

We obtained house price and house characteristics data from First American Real Estate Solutions. The observations selected correspond to a cross section of single-family residences sold during the 1998-2001 period in the St. Louis,

Missouri, metropolitan area. The data are from transactions as recorded in county property records. After eliminating from the original dataset observations with missing or outlier house prices (outside a bound of 3.5 SDs from the mean unadjusted house price), our sample includes 38,656 single-family residences.

We deflated house prices to 1998 dollars with the Office of Federal Housing Enterprise Oversight repeat-sales price index for the entire St. Louis metropolitan area.[8] In the full sample the resulting adjusted house price has a mean of $148,082 and an SD of $161,397. House characteristics include the total number of rooms, number of bedrooms, number of bathrooms, lot size, internal square footage, age of the structure, and number of stories in the house.

Attendance Zones

For the boundary discontinuity analysis, we obtained the definitions of 121 attendance zones for elementary schools in 15 school districts in St. Louis County. Most of these were obtained by contacting the school districts directly. Boundaries were variously provided as listings of streets, maps, and in some isolated cases as geocoded files. We, in turn, geocoded all the attendance zones and determined the boundary for every pair of adjacent schools, as in Black's paper (1999). We also geocoded each house in our sample using the street address. We then selected houses within a 0.1-mile buffer of the boundaries and assigned them to the nearest (and therefore unique) pairwise boundary.[9] We also eliminated from the boundary sample observations in St. Louis County that were associated with the boundaries of St. Louis City schools because the City property records contained no house price information. The final boundary sample consisted of 10,190 single-family residences.

Neighborhood Characteristics

Houses were also matched to Census blocks as the geographic unit at which we measured neighborhood demographics. We used the publicly available population tables at the block level from the Census 2000 Summary File 1, which includes counts by age, sex, and race, to construct the following measures: percent of females, percent of school-aged children (between 5 and 14 years of age), and percent of nonwhite population (defined as the total population count minus the count of white people).[10]

Additionally, we include as neighborhood controls the property tax rates defined at the school-district level for the years 1998 through 2001. In this case, each house was matched to the tax rate prevailing during the year of sale in its associated school district.[11] Table 1 presents summary statistics for house prices and characteristics with neighborhood characteristics for both the full and boundary samples.

Test Scores

As the measure of school quality, we use a school-level index generated by the Missouri Department of Elementary and Secondary Education. This index is computed from test score data from the Missouri Assessment Program (MAP); annual MAP testing is a statewide mandate for public schools. The MAP test includes a *Mathematics* section, a *Communication Arts* section (which includes a *Reading* portion), a *Science* section, and a *Social Studies* section.

Neither individual student scores nor school-level averages of these scores are publicly available. Instead, for each content area, the publicly available data provide the overall school-level MAP index. This index is obtained with a state-defined formula as the weighted sum of the percentages of students in each of five performance categories (Advanced, Proficient, Nearing

[8] House prices were deflated using the average price index corresponding to the quarter of the sale. The results were qualitatively unaffected if the National Association of Realtors price index was used instead.

[9] Black considers a number of different boundary width ranges and finds no significant differences. Our sample does not permit wider boundaries as these would encompass some attendance zones almost entirely.

[10] Our choice of demographic variables was limited by the availability of information at the block level in the public data files. Alternative measures such as median household income or share of households with a female head of household are not available at the block level.

[11] The analysis was not affected qualitatively if an average over the period was used instead.

Table 1

Summary Statistics (House and Neighborhood Characteristics)

	Full sample (*N* = 38,656)		Boundary sample (*N* = 10,190)	
House variables	Mean	SD	Mean	SD
Sale price (1998 US$)	148,081.67	161,397.24	142,033.42	176,191.20
Log of sale price	11.62	0.73	11.56	0.75
Number of bedrooms	2.96	0.84	2.9	0.84
Number of bathrooms	2.01	0.95	1.95	0.93
Number of bathrooms (squared)	4.97	5.05	4.66	5.04
Age of building	38.91	20.63	40.72	21.27
Age of building (squared)	1,939.38	1,922.87	2,110.15	2,028.41
Lot area (1,000s of sq. ft.)	14.75	38.35	13.61	39.20
Living area (1,000s of sq. ft.)	1.16	0.44	1.13	0.42
Number of stories	1.24	0.42	1.23	0.41
Total number of rooms	6.38	1.6	6.26	1.57
	Full sample (*N* = 6,360 blocks)		**Boundary sample (*N* = 2,560 blocks)**	
Census variables	Mean	SD	Mean	SD
Percent female population	51.17	11.22	51.34	11.33
Percent nonwhite population	20.43	29.29	22.42	30.67
Percent population 5 to 14 years of age	9.34	9.58	9.98	9.38

Table 2

Summary Statistics (Test Scores and Property Tax)

Variable	Mean	SD	Minimum	Maximum
Test scores (*N* = 121 schools)				
Math MAP score	211.45	19.44	168.14	250.18
Science MAP score	211.88	22.56	100.00	242.61
Reading MAP score	200.73	20.15	100.00	228.94
Property tax (*N* = 15 school districts)				
Property tax rate ($1/$1,000 of assessed house value)	4.23	0.91	2.60	5.74

Table 3

Correlation Table

Variable	Log house price	Math score	Math score (squared)	Math score (cubed)	Number of bedrooms	Number of bathrooms	Number of bathrooms (squared)	Age of building	Age of building (squared)	Lot area (1,000s of sq. ft.)	Living area (1,000s of sq. ft.)	Number of stories	Number of rooms	Census block: Percent female	Census block: Percent nonwhite	Census block: Percent people 5 to 14 years of age	Property tax rate
Log house price	1.00																
Math score	0.66	1.00															
Math score (squared)	–0.14	–0.35	1.00														
Math score (cubed)	0.50	0.87	–0.44	1.00													
Number of bedrooms	0.57	0.34	–0.09	0.25	1.00												
Number of bathrooms	0.68	0.50	–0.07	0.37	0.64	1.00											
Number of bathrooms (squared)	0.63	0.43	–0.01	0.32	0.58	0.94	1.00										
Age of building	–0.32	–0.38	0.17	–0.27	–0.29	–0.48	–0.39	1.00									
Age of building (squared)	–0.21	–0.29	0.16	–0.21	–0.21	–0.36	–0.28	0.94	1.00								
Lot area (1,000s of sq. ft.)	0.23	0.13	0.02	0.10	0.14	0.17	0.19	–0.03	–0.01	1.00							
Living area (1,000s of sq. ft.)	0.51	0.36	–0.04	0.26	0.39	0.45	0.43	–0.30	–0.27	0.20	1.00						
Number of stories	0.46	0.31	–0.02	0.22	0.47	0.56	0.52	–0.21	–0.07	0.06	–0.13	1.00					
Number of rooms	0.65	0.37	–0.04	0.26	0.82	0.71	0.67	–0.22	–0.14	–0.18	0.47	0.49	1.00				
Census block: Percent female	–0.11	–0.09	0.05	–0.07	0.11	–0.09	–0.08	0.04	0.02	–0.07	–0.08	–0.04	–0.11	1.00			
Census block: Percent nonwhite	–0.49	–0.69	0.48	–0.60	–0.25	–0.35	–0.28	0.30	0.24	–0.09	–0.24	–0.21	–0.26	0.16	1.00		
Census block: Percent people 5-14 yrs. of age	0.01	–0.07	0.12	–0.07	0.14	0.08	0.08	–0.13	–0.09	0.02	0.03	0.09	0.11	0.04	0.17	1.00	
Property tax rate	–0.47	–0.68	0.26	–0.56	–0.20	–0.35	–0.29	0.27	0.25	–0.07	–0.29	–0.14	–0.22	0.05	0.56	0.14	1.00

Table 4

Education Regressions: Full Sample

	Log house price		
Variable	(1)	(2)	(3)
Math score	0.21734***	0.22192***	0.31693***
	(7.79)	(7.13)	(7.70)
Math score (squared)		0.03002	0.01555
		(1.48)	(0.76)
Math score (cubed)			–0.03606**
			(–2.60)
Number of bedrooms	0.01062	0.01502	0.01575
	(1.09)	(1.52)	(1.62)
Number of bathrooms	0.14086***	0.14413***	0.13458***
	(4.75)	(4.93)	(4.44)
Number of bathrooms (squared)	–0.00612	–0.00740	–0.00501
	(–1.14)	(–1.37)	(–0.89)
Age of building	0.00065	0.00057	0.00123
	(0.37)	(0.31)	(0.67)
Age of building (squared)	0.00002	0.00002	0.00002
	(1.35)	(1.31)	(1.03)
Lot area (1,000s of sq. ft.)	0.00123***	0.00120***	0.00119***
	(4.21)	(4.27)	(4.17)
Living area (1,000s of sq. ft.)	0.45365***	0.44475***	0.43526***
	(20.02)	(17.35)	(19.05)
Number of stories	0.39693***	0.38775***	0.37835***
	(11.29)	(10.58)	(10.87)
Number of rooms	0.07484***	0.07421***	0.07245***
	(10.10)	(10.21)	(10.11)
Census block: Percent female	–0.00061	–0.00050	–0.00053
	(–0.88)	(–0.73)	(–0.79)
Census block: Percent nonwhite	–0.00221***	0.00277***	–0.00257***
	(–3.62)	(–5.06)	(–4.57)
Census block: Percent people 5 to 14 years of age	–0.00017	–0.00033	–0.00021
	(–0.19)	(–0.38)	(–0.24)
Property tax rate	–0.04636	–0.04457	–0.03562
	(–1.65)	(–1.51)	(–1.28)
Constant	10.00143***	9.99065***	9.96337***
	(59.89)	(57.55)	(58.13)
N	38,656	38,656	38,656
R^2	0.697	0.699	0.702
Adjusted R^2	0.697	0.698	0.702

NOTE: *t*-Statistics are listed in parentheses. ***Significant at the 1 percent level.

Proficient, Progressing, and Step 1). The formula is MAP index = (percent in Step 1) × 1 + (percent in Progressing) × 1.5 + (percent in Nearing Proficient) × 2 + (percent in Proficient) × 2.5 + (percent in Advanced) × 3. The weights are exogenously determined by the Missouri Department of Elementary and Secondary Education.[12]

For our study we chose the math MAP index for elementary schools only (fourth grade) as our measure of school quality.[13] This measure was then averaged over the 1998-2001 period to remove any year-to-year noise in the component variables (as in Bayer, Ferreira, and McMillan, 2007). Because our housing data are essentially cross sectional, this procedure provides one consistent score for each school in the sample.

Table 2 presents summary statistics for MAP indices along with property tax rates among the schools and school districts included in the sample. Table 3 presents the correlation matrix for the variables used in the analysis.

EMPIRICAL RESULTS

Standard Hedonic Regression

Table 4 presents the regression results using the full sample, which includes neighborhood demographic controls but excludes the boundary fixed effects. In addition to the traditional linear model, we include the quadratic and cubic specifications in test scores for completeness.

The housing characteristics enter the pricing equation with the expected sign. Increases in living area, lot size, and the total number of rooms increase the price of a house on average. Similarly, the number of bathrooms and the number of stories have a positive and statistically significant effect. The number of bedrooms, the number of bathrooms squared, the age of the building, and its square do not seem to have a statistically significant effect in the full sample.

Among the neighborhood demographics only the percent of the nonwhite population (measured at the block level) is capitalized into house prices with a negative and statistically significant effect. The estimated coefficients indicate that an increase of 1 percentage point in the proportion of the nonwhite population decreases house prices by about 22 (in the linear model) to 27 (in the quadratic model) basis points. The property tax rate does not have a statistically significant effect.

As expected, the regressions illustrate a strong relationship between school quality and house prices. The coefficient of 0.21734 in the traditional linear model (column 1) reveals that an increase in school test scores of a half SD results in a house premium of about 11 percent (0.21734/2 = 10.867 percent) or about $16,000 at the mean price. A half-SD increase is equivalent to an increase of 4.6 percent in the math MAP index.

The quadratic and cubic models in columns 2 and 3 of Table 4, respectively, also indicate a large and positive linear coefficient of school quality on house prices. The coefficient for the square of the math score is, however, not statistically significant in columns 2 and 3. Interestingly, the cubic coefficient in column 3 is statistically significant, but it enters with a negative sign, which indicates that the house price premium does not monotonically increase over the range of school quality. In any case, these models suggest that nonlinearities are relevant. This is confirmed by a battery of Wald specification tests (Table 5). These tests reject the null hypothesis of a model with a constant education premium. We find that the restriction of not including a quadratic or cubic term ($\psi_2 = \psi_3 = 0$) is rejected at the 1 percent level, while not including a cubic term ($\psi_3 = 0$) is rejected at the 5 percent level. However, the restriction of no quadratic term ($\psi_2 = 0$) is not rejected. Thus, the evidence indicates that the preferred specification for the education premium in the full sample is the cubic model.

Boundary Discontinuity Models

Table 6 presents the results for the restricted boundary sample (omitting the estimated coeffi-

[12] This formula was updated in 2007 when the number of performance categories was reduced to four.

[13] We consider the math score to be a measure of school quality superior to the reading or science measures. First, the math scores are arguably the most objective measure. Second, the distribution of the school math MAP index among the schools was contained almost entirely within 2 SDs of the mean. In contrast, the reading and science indices contained a large number of outliers, particularly in the lower tail. We did not consider the social sciences scores.

Table 5

Specification Tests: Full Sample with Neighborhood Controls

	Premium Model $f(Y) = \psi_1 Y + \psi_2 Y^2 + \psi_3 Y^3$		
	Linear	Quadratic	Cubic
Null hypothesis	$\psi_1 = 0$	$\psi_1 = \psi_2 = 0$	$\psi_1 = \psi_2 = \psi_3 = 0$
Wald *F*-statistic	60.757***	27.686***	30.665***
Null hypothesis		$\psi_2 = 0$	$\psi_2 = \psi_3 = 0$
Wald *F*-statistic		2.192	7.446***
Null hypothesis			$\psi_3 = 0$
Wald *F*-statistic			6.754**

NOTE: **Significant at the 5 percent level; ***significant at the 1 percent level.

cients for the boundary fixed effects). As in the full sample, house characteristics are statistically significant and with the expected sign. In contrast to the full sample results, the age of the building and its square, along with the square of the number of bathrooms, are statistically significant. Compared with the full sample results, the estimated coefficients for house characteristics are smaller in magnitude but very stable across specifications.

In the linear model in column 1, school quality is a statistically significant contributor to house prices and enters with the expected positive sign. Compared with the results from the full sample regression, the estimated coefficient declines in magnitude by a factor of about four. The estimate of the education premium implies that a half-SD increase (equivalent to an increase of 4.6 percent) in the average school score leads to an increase of about 3.2 percent in house prices, or about $4,766 evaluated at the full sample mean price. This value is only slightly higher than that estimated by Black (1999). She reports a 2.1 percent increase (or $3,948 at her sample mean) in house prices for a 5 percent increase in test scores.

The two specifications of the nonlinear boundary fixed effects models in columns 2 and 3 indicate that the quadratic coefficient of school quality is statistically significant, but the cubic coefficient is not. The positive sign of the quadratic coefficient indicates that the capitalization effect of school quality is increasing over the range of test scores.

Specifications 1, 2, and 3 do not include additional controls for neighborhood quality other than the boundary fixed effects. As mentioned previously, some authors have raised concerns about whether the boundary discontinuity approach fails to control for omitted neighborhood characteristics and suggest that explicit additional controls be included in the estimation. We therefore include the same demographic controls as in the full sample regression—namely, the percent of female population, the percent of nonwhite population, and the percent of school-aged children, all measured at the block level. We also include the school-district property tax rate.

Columns 4, 5, and 6 in Table 6 show that these additional variables are directly capitalized into house prices. The percent of the nonwhite population is statistically significant and enters with a negative sign as in the full sample results. The magnitude of the effect is similar to the full sample results and indicates a decline of about 22 basis points in house prices for a 1-percentage-point increase in the proportion of the nonwhite population. We interpret the significance of this variable, as in other papers, as evidence of preferences about the racial composition of neighborhoods.

In contrast to the full sample results, the percent of school-aged children is statistically significant and indicates an increase in house prices of about 15 basis points for a 1-percentage-point

Table 6

Education Regressions: Restricted Boundary Sample

	Log-adjusted price					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Math score	0.06437**	0.06274***	0.04659	0.03227*	0.03579*	0.03172
	(2.58)	(2.90)	(1.64)	(1.78)	(1.93)	(1.20)
Math score (squared)		0.02656**	0.02909**		0.02209**	0.02284**
		(2.47)	(2.47)		(2.48)	(2.40)
Math score (cubed)			0.00514			0.00137
			(0.73)			(0.21)
Number of bedrooms	0.03726***	0.03730***	0.03749***	0.03816***	0.03805***	0.03809***
	(3.88)	(3.89)	(3.90)	(4.02)	(4.01)	(4.00)
Number of bathrooms	0.10834***	0.10785***	0.10792***	0.10349***	0.10318***	0.10320***
	(5.78)	(5.80)	(5.82)	(5.81)	(5.82)	(5.83)
Number of bathrooms (squared)	–0.00529*	–0.00533*	–0.00535*	–0.00488	–0.00491	–0.00491
	(–1.68)	(–1.70)	(–1.71)	(–1.58)	(–1.60)	(–1.60)
Age of building	–0.00408***	–0.00411***	–0.00412***	–0.00453***	–0.00454***	–0.00454***
	(–2.73)	(–2.75)	(–2.76)	(–3.11)	(–3.13)	(–3.14)
Age of building (squared)	0.00004***	0.00004***	0.00004***	0.00004***	0.00004***	0.00004***
	(2.89)	(2.91)	(2.92)	(3.15)	(3.16)	(3.17)
Lot area (1,000s of sq. ft.)	0.00089**	0.00089**	0.00089**	0.00088**	0.00088**	0.00088**
	(2.41)	(2.41)	(2.41)	(2.39)	(2.40)	(2.39)
Living area (1,000s of sq. ft.)	0.35315***	0.35228***	0.35236***	0.34332***	0.34297***	0.34301***
	(15.43)	(15.29)	(15.29)	(15.52)	(15.49)	(15.49)
Number of stories	0.27574***	0.27559***	0.27558***	0.26621***	0.26625***	0.26626***
	(9.30)	(9.30)	(9.31)	(9.55)	(9.57)	(9.57)
Number of rooms	0.05974***	0.05952***	0.05945***	0.05902***	0.05893***	0.05891***
	(7.38)	(7.33)	(7.31)	(7.43)	(7.40)	(7.39)
Census block: Percent female				–0.00044	–0.00039	–0.00039
				(–0.66)	(–0.59)	(–0.59)
Census block: Percent nonwhite				–0.00219***	–0.00223***	–0.00222***
				(–3.50)	(–3.56)	(–3.55)
Census block: Percent people 5 to 14 years of age				0.00154**	0.00153**	0.00154**
				(2.25)	(2.24)	(2.25)
Property tax rate				–0.06787***	–0.05526***	–0.05465***
				(–3.21)	(–2.88)	(–2.73)
Constant	11.13260***	11.12998***	11.13935***	8.86314***	8.72871***	8.72454***
	(32.85)	(32.96)	(32.97)	(59.31)	(62.61)	(60.54)
N	10,190	10,190	10,190	10,182	10,182	10,182
R^2	0.769	0.77	0.77	0.772	0.772	0.772
Adjusted R^2	0.763	0.763	0.763	0.766	0.766	0.766
Boundary fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

NOTE: *t*-Statistics are listed in parentheses. *Significant at the 10 percent level; **significant at the 5 percent level; ***significant at the 1 percent level.

Table 7

Specification Tests: Boundary Sample

	Premium Model $f(Y) = \psi_1 Y + \psi_2 Y^2 + \psi_3 Y^3$		
	Linear	Quadratic	Cubic
Without neighborhood controls			
Null hypothesis	$\psi_1 = 0$	$\psi_1 = \psi_2 = 0$	$\psi_1 = \psi_2 = \psi_3 = 0$
Wald *F*-statistic	6.632**	4.658**	3.130**
Null hypothesis		$\psi_2 = 0$	$\psi_2 = \psi_3 = 0$
Wald *F*-statistic		6.115**	3.114**
Null hypothesis			$\psi_3 = 0$
Wald *F*-statistic			0.527
With neighborhood controls			
Null hypothesis	$\psi_1 = 0$	$\psi_1 = \psi_2 = 0$	$\psi_1 = \psi_2 = \psi_3 = 0$
Wald *F*-statistic	3.178*	3.581**	2.381*
Null hypothesis		$\psi_2 = 0$	$\psi_2 = \psi_3 = 0$
Wald *F*-statistic		6.166**	3.102**
Null hypothesis			$\psi_3 = 0$
Wald *F*-statistic			0.043

NOTE: **Significant at the 5 percent level; ***significant at the 1 percent level.

increase in the proportion of children between 5 and 14 years of age. The property tax rate is also statistically significant and enters with a negative sign.

The inclusion of explicit neighborhood controls does not affect the magnitude of the coefficients of the housing characteristics, but it decreases the magnitude of the linear test score coefficient by almost half. The quadratic coefficient declines only slightly. The linear coefficient on school quality remains, nevertheless, statistically significant, and the results suggest that the magnitude of the effect of school quality on house prices remains substantially large.

Wald specification tests (Table 7) confirm that, with or without the inclusion of additional neighborhood controls, the preferred specification is the quadratic model. These tests also reject, as in the full sample regressions, the null hypothesis of a model with a constant education premium. We find that the restriction of not including a quadratic or cubic term ($\psi_2 = \psi_3 = 0$) is rejected at the 5 percent level. However, the restriction of no cubic term ($\psi_3 = 0$) is not rejected.

Implicit Housing Premia

Figure 1 illustrates the preferred specification for the house pricing function with the more conservative model with boundary fixed effects resulting from the inclusion of additional neighborhood controls. The plot includes 1-SE bands.[14] We argued earlier that competition in the housing market generates increasing tightness in areas associated with higher school quality, but that competition is not as prevalent in areas associated with lower school quality. The pricing function in Figure 1 confirms our argument.

The premium from school quality on housing prices is better illustrated in Figure 2. This figure is constructed from the pricing function of specification 5 in Table 6 and represents the percentage increase in house prices in response to a half-SD increase in math test scores plotted

[14] The asymptotic variance of the price function was computed using the delta method as

$$\text{AsyVar}\left(f(\mu;\beta)\right) = \frac{\partial f(\mu;\beta)}{\partial \beta'} \text{AsyVar}(\beta) \left(\frac{\partial f(\mu;\beta)}{\partial \beta'}\right)'.$$

Figure 1

Implied Price Function (with Neighborhood Controls)

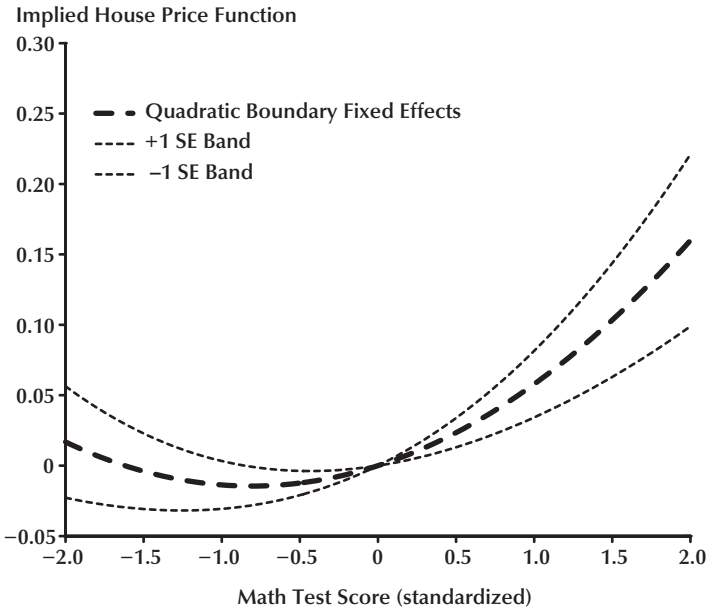
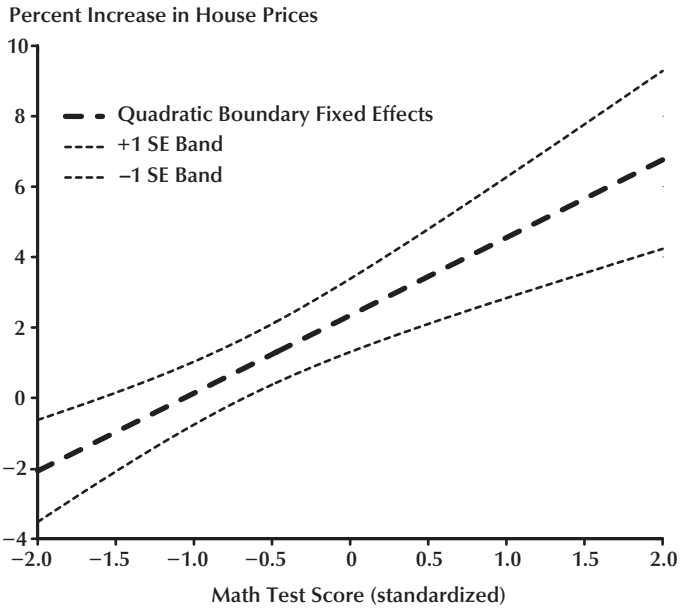


Figure 2

Implied Premium Function (with Neighborhood Controls)



NOTE: The plots show the response to a half-SD increase in math test scores.

Table 8

Implied House Price Premia from School Quality

Regression model	Full sample with neighborhood controls			Boundary sample: Without neighborhood controls			Boundary sample: With neighborhood controls		
Linear coefficient	0.21734	0.22192	0.31693	0.06437	0.06274	0.04659	0.03227	0.03579	0.03172
Quadratic coefficient	—	0.03002	0.01555	—	0.02656	0.02909	—	0.02209	0.02284
Cubic coefficient	—	—	–0.03606	—	—	0.00514	—	—	0.00137
Case 1 (mean score minus 1 SD)									
Percent increase in house price	10.87	8.84	11.53	3.22	1.15	0.6	1.61	0.13	–0.01
Dollar value at mean (full sample)	16,092	13,097	17,066	4,766	1,696	885	2,389	197	–11
Dollar value at mean (boundary sample)	15,435	12,562	16,369	4,571	1,626	849	2,292	189	–10
Case 2 (mean score)									
Percent increase in house price	10.87	11.85	15.78	3.22	3.80	3.12	1.61	2.34	2.17
Dollar value at mean (full sample)	16,092	17,542	23,374	4,766	5,629	4,622	2,389	3,468	3,219
Dollar value at mean (boundary sample)	15,435	16,826	22,419	4,571	5,399	4,433	2,292	3,326	3,088
Case 3 (mean score plus 1 SD)									
Percent increase in house price	10.87	14.85	9.23	3.22	6.46	7.19	1.61	4.55	4.77
Dollar value at mean (full sample)	16,092	21,988	13,662	4,766	9,562	10,642	2,389	6,739	7,058
Dollar value at mean (boundary sample)	15,435	21,090	13,104	4,571	9,171	10,207	2,292	6,464	6,770

NOTE: The table presents the premium in house prices evaluated at different math scores resulting from a change in math score of 0.5 SD (equivalent to 4.6 percent of the mean score). The premium is computed from the logarithm specification $\Delta p/p = \Delta \ln(p) = \Delta f(\mu)$, so the percent change in house prices is given by $\Delta f(\mu) = f(\mu_1) - f(\mu_0)$ and the premium at the mean price is $\Delta f(\mu) \times \bar{p}$.

along the range of school scores within 2 SDs of the mean.

The plotted function reveals a monotonically increasing premium across the spectrum of school quality. The plot indicates that, even with the most conservative estimates, the premium for houses in areas associated with high-quality schools remains substantially large. The plot also reveals a much smaller premium for houses in areas associated with low-quality schools, where house prices seem to be driven almost entirely by housing and neighborhood characteristics other than public school quality.

Table 8 summarizes the implied school quality premia from school quality for all models and provides the dollar equivalent of the implied percentage increase in house prices relative to the mean house prices in the full and boundary samples that results from a half-SD increase in test scores.

The linear model with the full sample regression results in a constant premium of 10.87 percent or about $16,000 at the mean house price. The cubic model in the full sample, which the specification tests suggest is the preferred model, illustrates a nonmonotonic premium that ranges from 11.53 percent for houses in areas where school quality is 1 SD below the mean to 15.78 percent in areas where school quality coincides with the average, and finally to 9.23 percent in areas where school quality is 1 SD above the mean.

The boundary sample models with and without additional neighborhood controls indicate that the premium is severely overestimated in the traditional hedonic regressions, even accounting for nonlinearities. Nevertheless, even in the most conservative estimates, the premium remains substantially large, especially for areas associated with very high-quality schools. Table 8 also shows two characteristics in the quadratic equation—the middle column of the third panel: The premium is very small in areas where test scores are 1 SD below the mean (about 0.13 percent or less than $200) and monotonically increases in areas with higher test scores (about 2.34 percent or $3,468 in areas with average test scores [Case 2] and 4.55 percent or $6,739 in areas with test scores 1 SD above the mean [Case 3]).

CONCLUSION

Traditional empirical models of the capitalization of education quality on house prices have established that the quality of primary school education is positively correlated with house prices. Recent capitalization studies have used various approaches to address concerns about omitted variable bias induced by failing to account for the correlation between school quality and unobserved neighborhood characteristics. Most of these variations on the traditional hedonic approach (including the boundary discontinuity regression) have assumed that the house price premium is constant because in all these models the contribution from school quality on house prices is constrained to be linear.

In this paper, we propose an alternative formulation that allows for nonlinear effects of school quality. We show that this formulation is preferred by the data over a baseline linear boundary fixed effects model and that the rate at which the house price premium rises increases over the range of school quality. In other words, the standard linear specification for test scores overestimates the premium at low levels of school quality and underestimates the premium at high levels of school quality.

In the St. Louis metropolitan area, houses associated with a school ranked at 1 SD below the mean are essentially priced on physical characteristics only. In contrast, houses associated with higher-quality schools command a much higher price premium.

Interestingly, and in contrast to many studies in the literature, the price premium remains substantially large, especially for houses associated with above-average schools. This is true even in our most conservative estimates, which complement the boundary discontinuity approach by explicitly controlling for neighborhood demographics. These estimates also reveal that the racial composition of neighborhoods is capitalized directly into house prices.

REFERENCES

Bayer, Patrick; Ferreira, Fernando and McMillan, Robert. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy*, August 2007, *115*(4), pp. 588-638.

Black, Sandra E. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, May 1999, *114*(2), pp. 577-99.

Bogart, William T. and Cromwell, Brian A. "How Much More Is a Good School District Worth?" *National Tax Journal*, June 1997, *50*(2), pp. 215-32.

Bogart, William T. and Cromwell, Brian A. "How Much Is a Neighborhood School Worth?" *Journal of Urban Economics*, March 2000, *47*(2), pp. 280-305.

Brasington, David. "Which Measures of School Quality Does the Housing Market Value?" *Journal of Real Estate Research*, 1999, *18*(3), pp. 395-413.

Brasington, David. "Edge Versus Center: Finding Common Ground in the Capitalization Debate." *Journal of Urban Economics*, November 2002, *52*(3), pp. 524-41.

Brasington, David and Haurin, Donald R. "Educational Outcomes and House Values: A Test of the Value Added Approach." *Journal of Regional Science*, May 2006, *46*(2), pp. 245-68.

Cheshire, Paul and Sheppard, Stephen. "The Welfare Economics of Land Use Planning." *Journal of Urban Economics*, September 2002, *52*(2), pp. 242-69.

Cheshire, Paul and Sheppard, Stephen. "Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty." *Economic Journal*, November 2004, *114*(499), pp. F397-F424.

Clapp, John M.; Nanda, Anupam and Ross, Stephen L. "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition of Property Values." *Journal of Urban Economics*, March 2008, *63*(2), pp. 451-66.

Cushing, Brian J. "Capitalization of Interjurisdictional Fiscal Differentials: An Alternative Approach." *Journal of Urban Economics*, May 1984, *15*(3), pp. 317-26.

Davidoff, Ian and Leigh, Andrew. "How Much Do Public Schools Really Cost? Estimating the Relationship Between House Prices and School Quality." Discussion Paper No. 558, Australian National University, Centre for Economic Policy Research, July 2007.

Downes, Thomas A. and Zabel, Jeffrey E. "The Impact of School Characteristics on House Prices: Chicago 1987-1991." *Journal of Urban Economics*, July 2002, *52*(1), pp. 1-25.

Fack, Gabrielle and Grenet, Julien. "Do Better Schools Raise Housing Prices? Evidence from Paris School Zoning." Unpublished manuscript, Paris École Normale Supériere, October 2007.

Figlio, David N. and Lucas, Maurice E. "What's in a Grade? School Report Cards and House Prices." *American Economic Review*, June 2004, *94*(3), pp. 591-604.

Gibbons, Stephen and Machin, Stephen. "Valuing English Primary Schools." *Journal of Urban Economics*, March 2003, *53*(2), pp. 197-219.

Gibbons, Stephen and Machin, Stephen. "Paying for Primary Schools: Admission Constraints, School Popularity, or Congestion?" *Economic Journal*, March 2006, *116*(510), pp. C77-C92.

Gibbons, Stephen and Machin, Stephen. "Valuing School Quality, Better Transport, and Lower Crime: Evidence from House Prices." *Oxford Review of Economic Policy*, Spring 2008, *24*(1), pp. 99-119.

Gill, H. Leroy. "Changes in City and Suburban House Prices During a Period of Expected School Desegregation." *Southern Economic Journal*, July 1983, *50*(1), pp. 169-84.

Hanushek, Eric A. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, September 1986, *24*(3), pp. 1141-77.

Hanushek, Eric A. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis*, Summer 1997, *19*(2), pp. 141-64.

Haurin, Donald R. and Brasington, David. "School Quality and Real House Prices: Inter- and Intrametropolitan Effects." *Journal of Housing Economics*, December 1996, *5*(4), pp. 351-68.

Hayes, Kathy J. and Taylor, Lori L. "Neighborhood School Characteristics: What Signals Quality to Homebuyers?" Federal Reserve Bank of Dallas *Economic Review*, Fourth Quarter 1996, pp. 2-9; www.dallasfed.org/research/er/1996/er9604a.pdf.

Hilber, Christian A.L. and Mayer, Christopher J. "Why Do Households Without Children Support Local Public Schools? Linking House Price Capitalization to School Spending." *Journal of Urban Economics*, January 2009, *65*(1), pp. 74-90.

Imbens, Guido W. and Lemieux, Thomas. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 2008, *142*(2), pp. 615-35.

Kain, John F. and Quigley, John M. "Evaluating the Quality of the Residential Environment." *Environment and Planning*, 1970, *2*(1), pp. 23-32.

Kane, Thomas J.; Staiger, Douglas O. and Reigg, Stephanie K. "Changing School Assignments and Housing Values." Working paper, UCLA Department of Policy Studies, 2003.

Kane, Thomas J.; Staiger, Douglas O. and Reigg, Stephanie K. "School Quality, Neighborhoods and Housing Prices: The Impacts of School Desegregation." NBER Working Paper No. 11347, National Bureau of Economic Research, May 2005; www.nber.org/papers/w11347.pdf.

Kane, Thomas J.; Staiger, Douglas O. and Samms, Gavin. "School Accountability Ratings and Housing Values," in William Gale and Janet Pack, eds., *Brookings-Wharton Papers on Urban Affairs*. Washington, DC: Brookings Institution, 2003, pp. 83-137.

Leech, Dennis and Campos, Erick. "Is Comprehensive Education Really Free? A Case-Study of the Effects of Secondary School Admissions Policies on House Prices in One Local Area." *Journal of the Royal Statistical Society Series A*, 2003, *166*(1), pp. 135-54.

Oates, Wallace E. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy*, November/December 1969, *77*(6), pp. 957-71.

Reback, Randall. "House Prices and the Provision of Local Public Services: Capitalization Under School Choice Programs." *Journal of Urban Economics*, March 2005, *57*(2), pp. 275-301.

Ridker, Ronald G. and Henning, John A. "The Determinants of Residential Property Values with Special Reference to Air Pollution." *Review of Economics and Statistics*, May 1967, *49*(2), pp. 246-57.

Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, January/February 1974, *82*(1), pp. 34-55.

Rosenthal, Leslie. "The Value of Secondary School Quality." *Oxford Bulletin of Economics and Statistics*, July 2003, *65*(3), pp. 329-55.

Ross, Stephen and Yinger, John. "Sorting and Voting: A Review of the Literature on Urban Public Finance," in Paul Cheshire and Edwin Mills, eds., *Handbook of Regional and Urban Economics. Volume 3: Applied Urban Economics*. Chap. 47. Amsterdam: North Holland, 1999, pp. 2003-060.

Sheppard, Stephen. "Hedonic Analysis of Housing Markets," in Paul Cheshire and Edwin Mills, eds., *Handbook of Regional and Urban Economics. Volume 3: Applied Urban Economics*. Chap. 41. Amsterdam: North Holland, 1999, pp. 1595-635.

Thorsnes, Paul and Reifel, John W. "Tiebout Dynamics: Neighborhood Response to a Central-City/Suburban House-Price Differential." *Journal of Regional Science*, October 2007, *47*(4), pp. 693-719.

Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, October 1956, *64*(5), pp. 416-24.

Weimer, David L. and Wolkoff, Michael J. "School Performance and Housing Values: Using Non-Contiguous District and Incorporation Boundaries to Identify School Effects." *National Tax Journal*, June 2001, *54*(2), pp. 231-53.

Wheaton, William C. "Vacancy, Search, and Prices in a Housing Market Matching Model." *Journal of Political Economy*, December 1990, *98*(6), pp. 1270-92.

Williams, Joseph T. "Pricing Real Assets with Costly Search." *Review of Financial Studies*, Spring 1995, *8*(1), pp. 55-90.

Institutional Causes of Output Volatility

Levon Barseghyan and Riccardo DiCecio

The authors investigate the relationship between the quality of institutions and output volatility. Using instrumental variable regressions, they address whether higher entry barriers and lower property rights protection lead to higher volatility. They find that a 1-standard-deviation increase in entry costs increases the standard deviation of output growth by roughly 40 percent of its average value in the sample. In contrast, property rights protection has no statistically significant effect on volatility. (JEL O11, O17, O43)

Federal Reserve Bank of St. Louis *Review*, May/June 2010, *92*(3), pp. 205-23.

Poor macroeconomic policies in less-developed countries have been blamed for the negative relationship between growth and macroeconomic volatility as measured by the volatility of the growth rate of output per worker. Acemoglu et al. (2003) offer a different explanation: Volatility is related to institutional quality; once institutions are controlled for, macroeconomic policies (i.e., fiscal, monetary, and exchange rate policy) have only a minor effect on volatility. This finding raises the question of how institutions affect output volatility—more precisely, which institutional features are most responsible for the relationship documented by Acemoglu et al. (2003). We use instrumental variable (IV) regressions to disentangle the effect of two distinct types of institutions: entry barriers and property rights protection. We find that higher entry barriers lead to higher output volatility. In contrast, property rights protection appears to have no effect on output volatility.

Entry barriers and property rights protection are correlated in the data, although their economic effects, both empirically and theoretically, are quite different. Barseghyan (2008) shows that worse property rights protection leads to lower educational attainment and a lower capital-to-output ratio: A lack of property rights enforcement discourages investment in all types of capital. The effect of property rights on total factor productivity (TFP) is much weaker and is mostly statistically insignificant. On the other hand, entry costs have no effect on the capital-to-output ratio but do have a strong effect on TFP. According to prevalent theories of industry structure (e.g., Hopenhayn, 1992), this is exactly what should be expected: Higher entry barriers reduce entry, protect incumbent firms, and allow those with lower productivity to survive. Thus, the results of our paper suggest that differences in output volatility are driven by industry structure, which, in turn, is significantly affected by entry barriers. This is consistent with the findings of Acemoglu et al. (2003) that a significant part of the effect of institutions on economic outcomes occurs through microeconomic channels.

In a related paper, we explore the link between entry costs and cross-country output and TFP differences through the lenses of general equilib-

Levon Barseghyan is an assistant professor of economics at Cornell University. Riccardo DiCecio is an economist at the Federal Reserve Bank of St. Louis. Charles S. Gascon provided research assistance.

rium models. Barseghyan and DiCecio (2010) construct a model with endogenous entry and operation decisions by firms and calibrate it to match the U.S. distribution of firms by size. Higher entry costs lead to greater misallocation of productive factors and lower steady-state TFP and output. As in the data, higher entry costs are associated with a larger informal sector and overall number of operating firms, a smaller number of legally registered firms, and a higher concentration of employment in the smallest and largest firms.

In our investigation, we use a measure of entry costs originally constructed by Djankov et al. (2002) and later expanded by the World Bank (2007). Unlike most measures of institutional quality, this is a continuous variable that captures the precise quantitative value of the object of interest. We control for property rights by considering five proxies for property rights institutions: the rate of debt recovery from a "going-out-of-business" borrower, three indices of property rights protection, and a social infrastructure measure. Sources of exogenous variation in entry costs and the property rights measures are given by the following instruments: geographic latitude, fraction of the population speaking a major European language, the country's legal origin, European settler mortality in early stages of colonization, and indigenous population density in the early sixteenth century, the use of which is further explained below.

The IV regressions reveal that entry costs have a statistically significant effect on output volatility. The economic effect of entry costs is worth emphasizing. A 1-standard-deviation (SD) increase in entry costs is estimated to increase the SD of the growth rate of output per worker by 41 percent of its average value in our sample. Also, a 1-SD increase in entry costs increases the magnitude of the worst output drop by 60 percent of its sample average. We perform a variety of checks to ensure that the estimated strong effect of entry costs on volatility is robust. Notably, we entertain the possibility raised by Glaeser et al. (2004) that the defining characteristic of a successful European settlement was an increase in human capital. We include human capital as an endogenous variable in the IV regressions. The robustness exercises confirm that entry costs are an important determinant of output volatility. Moreover, the magnitude of this effect is close to the one estimated in the benchmark regressions. The effect of property rights on volatility remains insignificant throughout robustness analysis.

This paper belongs to the empirical literature on institutions and growth, such as Hall and Jones (1999); Acemoglu, Johnson, and Robinson (2002); Acemoglu et al. (2003); Dollar and Kraay (2003); Easterly and Levine (2003); Rodrik, Subramanian, and Trebbi (2004); and earlier contributions by Knack and Keefer (1995) and Mauro (1995). The empirical strategy used in the paper is closest to that of Acemoglu and Johnson (2005) and Barseghyan (2008). As in these papers, our analysis hinges on the availability of a set of instruments that affect current economic outcomes only through institutions and are capable of separating the effects of various institutional features.

Our findings suggest that entry costs, by affecting the composition of the pool of firms, have an impact on volatility. Comparably, Koren and Tenreyro (2007) highlight the importance of the sectorial composition in understanding the relationship between development and volatility. Kraay and Ventura (2007) argue that comparative advantage determines differences in the composition of firms between rich and poor countries, making least-developed countries more volatile.

In the next section, we present the data and methodology used in the empirical investigation. We present the results of the empirical investigation in the following section and discuss their robustness in the final section. The appendix provides data sources and definitions.

DATA AND METHODOLOGY

Output Volatility

The benchmark measure of volatility is constructed using purchasing power–adjusted gross domestic product (GDP) per worker annual data from the Penn World Table 6.2 constructed by Heston, Summers, and Aten (2006). We consider only countries for which (i) the data for output per worker are available for at least 20 years and

(ii) entry costs data are available.[1] Our benchmark measure of volatility is the SD of the growth rate of output per worker. To assess the robustness of our results, we also consider the worst output drop (i.e., the minimum growth rate of output per worker). For comparison, we also construct the average growth rate for each country and report descriptive statistics for it.

ENTRY COSTS, PROPERTY RIGHTS, AND SOCIAL INFRASTRUCTURE

Entry costs are from the World Bank's *Doing Business* dataset and are available for 132 countries.[2] They include all official fees and dues that an entrepreneur must pay in the process of completing legal procedures for starting a new firm. They are constructed for a "standardized" firm. Although this standardized firm is relatively small, it is quite representative of a typical firm because smaller production units have a large share of aggregate employment.[3]

In most developed countries, entry costs are not a significant burden on entrepreneurs: For example, in Canada entrepreneurs pay less than 1 percent of gross national income (GNI) per capita in entry costs, whereas the cross-country average is 79 percent of GNI per capita. Higher entry costs are associated with worse macroeconomic conditions along several dimensions, as shown in Table 1 and Figures 1 through 3. Entry costs are positively correlated with volatility and negatively correlated with average growth. Also, higher entry costs are associated with more severe economic crises, measured by the worst output drop.

Finding a suitable proxy for property rights protection is more challenging. The first variable we use is the rate of debt recovery from a "going-out-of-business" borrower. This is, to our knowledge, the only available quantitative measure that can proxy property rights protection. The second variable, "constraint on executive power," refers to "the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities" (Jaggers and Marshall, 2000). It can be used as a proxy for the protection of private citizens and businesses against government expropriation. However, it may ignore the risk of expropriation by other agents. The third variable is the property rights protection index constructed by the Heritage Foundation (2006). The fourth variable is the "expropriation risk" constructed by the Political Risk Services (1999). It measures the risk of expropriation of private foreign investment by the government.[4] Finally, we consider the social infrastructure measure proposed by Hall and Jones (1999). It was constructed as the average between the government anti-diversion policy index and the openness to international trade measure of Sachs and Warner (1995). All property rights measures and social infrastructure are strongly positively correlated with each other and are negatively correlated with output growth volatility and entry costs (see Table 1).

ECONOMETRIC MODEL

The target is to identify and estimate the following relationship:

$$Y_i = \gamma_0 + \gamma_E E_i + \gamma_0 O_i + Z_i' \gamma_Z + \varepsilon_i,$$

where Y_i is the volatility of output growth for country i, E_i is the measure of entry costs, O_i is the proxy for other institutions, Z_i is the vector

[1] Notice that for different countries the volatility, average growth, and worst output drop are computed for different time periods. Our results are robust to the use of the same sample for all countries (e.g., 1961-2003).

[2] We consider only countries for which both volatility and entry costs data are available.

[3] In a sample of countries in the Organisation for Economic Co-operation and Development, for which harmonized firm-level data are available, the employment share of the firms with fewer than 50 workers is substantial—about one-third of the total. In less-developed and developing countries, which constitute a large part of our sample, the employment share of smaller establishments is much larger than in developed countries—typically more than 60 percent of the total (see Tybout, 2000).

[4] Acemoglu and Johnson (2005) use constraint on executive power, the Heritage Foundation index, and expropriation risk to proxy for property rights. Their preferred measure is constraint on executive power because it conceptually refers to constraints directly imposed on government actions. The other variables are equilibrium outcomes driven by policies that may result from such constraints.

Table 1

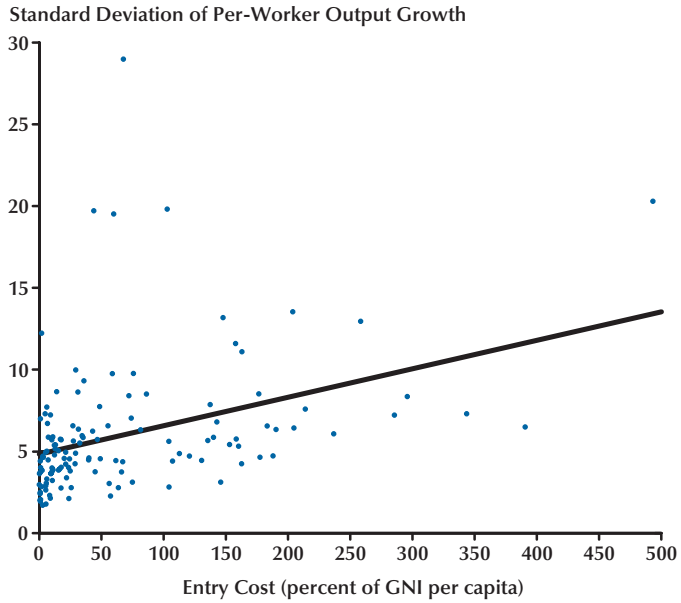
Descriptive Statistics and Correlations of Macro Variables, Institutions, and Moments of the Distribution of Firms

				Correlations										
Variables	Observations	Sample average	SD	SD growth rate	Worst output drop	Average growth rate	Entry costs	Debt recovery rate	Constraint on executive power	Heritage Foundation index	Expropriation risk	Social infrastructure	Average firm size	Variance of firm size
SD of growth rate	132	6.02	4.01	1.00										
Worst output drop	132	–13.23	9.64	–0.90	1.00									
Average growth rate	132	1.58	1.85	–0.30	0.43	1.00								
Entry costs	132	79.45	133.51	0.26	–0.32	–0.48	1.00							
Debt recovery rate	132	31.97	26.59	–0.44	0.45	0.34	–0.41	1.00						
Constraint on executive power	132	4.83	1.99	–0.52	0.51	0.35	–0.40	0.51	1.00					
Heritage Foundation index	112	3.20	1.11	–0.30	0.41	0.38	–0.46	0.79	0.58	1.00				
Expropriation risk	59	6.52	1.51	–0.22	0.33	0.50	–0.32	0.59	0.36	0.73	1.00			
Social infrastructure	117	0.51	0.49	–0.16	0.20	0.15	–0.22	0.37	0.36	0.85	0.73	1.00		
Average firm size	79	3.07	1.38	0.42	–0.39	–0.30	0.23	–0.73	–0.53	–0.75	–0.64	–0.79	1.00	
Variance of firm size	79	2.47	1.17	0.41	–0.41	–0.26	0.14	–0.61	–0.56	–0.65	–0.39	–0.68	0.84	1.00

NOTE: See the appendix for data sources and definitions.

Figure 1

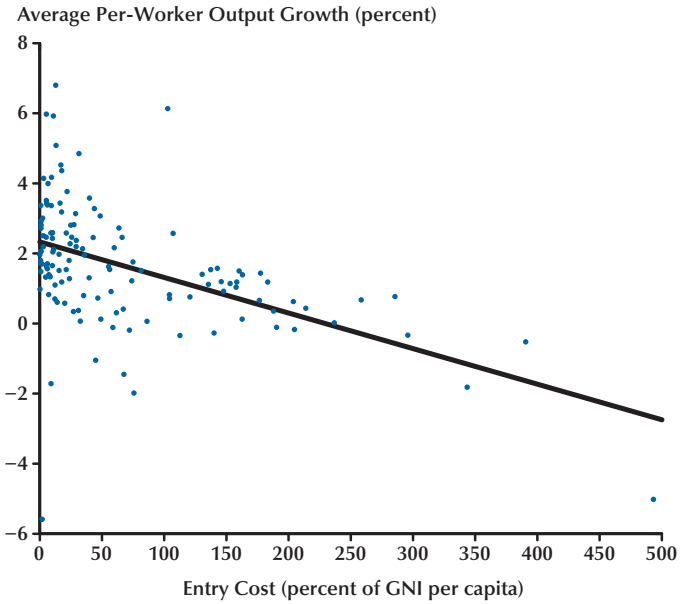
Volatility and Entry Costs: Data and Linear Fit Line (slope 1.74, *p*-value 0.000)



NOTE: Two outliers (entry costs > 500 percent of GNI) excluded.

Figure 2

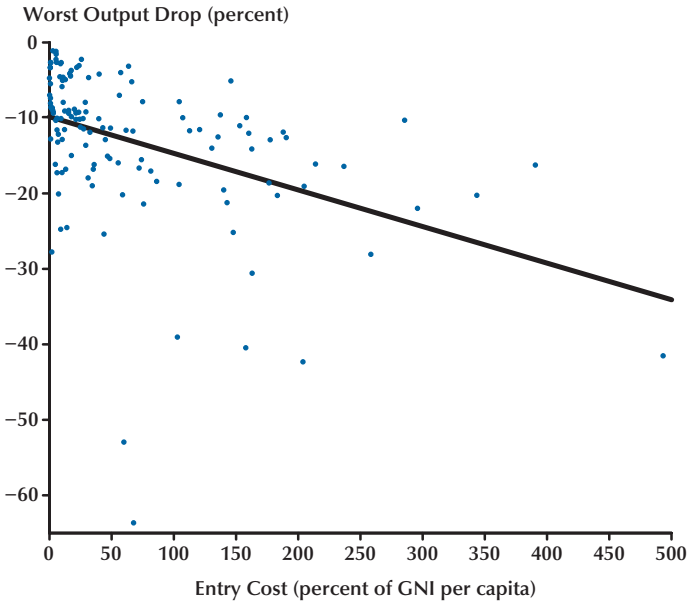
Growth and Entry Costs: Data and Linear Fit Line (slope –1.01, *p*-value 0.000)



NOTE: Two outliers (entry costs > 500 percent of GNI) excluded.

Figure 3

Crises and Entry Costs: Data and Linear Fit Line (slope –4.85, *p*-value 0.000)



NOTE: Two outliers (entry costs > 500 percent of GNI) excluded.

of additional controls, and ε_i is the error term. An IV procedure is implemented because of potential endogeneity, omitted variable bias, and measurement error. The following two assumptions must be satisfied for an IV regression to be valid:

(A1) The instruments must satisfy the rank condition

$$rank\ \mathbb{E}\left(\left[1\ I\ Z\right]' \left[1\ E\ O\ Z\right] \right) = (3+z),$$

where I denotes the vector of instruments and z is the number of additional controls.

(A2) The instruments must be uncorrelated with the error term, ε_i.

As discussed in our results below, we test whether these two assumptions are satisfied in the data to corroborate our IV regression analysis.

INSTRUMENTS

From the set of instruments available in the literature, we use geographic latitude, the fraction of the population speaking a major European language, legal origin, and, for a subsample of former colonies, European settler mortality and indigenous population density.

The first two instruments are those used by Hall and Jones (1999), who argued that geographic characteristics and the extent to which major European languages have been adopted in a country are correlated with the quality of the country's institutions. This is true because (i) Europeans were more likely to settle and establish Western institutions where the geographic characteristics were more similar to those in their countries of origin and (ii) the extent to which European culture and, consequently, European institutions have spread in a country is likely to be correlated with the adoption of European languages.

Legal origin (La Porta et al., 1999) has a strong effect on various institutional features related to property rights, most notably on the degree of legal formalism, which is associated with judicial transparency and fairness, safeguards against corruption, and enforceability of contracts.

Settler mortality and population density, introduced by Acemoglu, Johnson, and Robinson (2002) and Acemoglu et al. (2003), can be used as instruments because of their lasting effects on countries' institutional development. Early European settlements were negatively affected by high mortality rates. In places where Europeans were settling in large numbers, it was in their interest to promote free entrepreneurship, provide property rights protection, and so on. Higher indigenous population density, on the other hand, provided Europeans with an opportunity to capture and exploit local labor, giving rise to extractive institutions and, therefore, poor property rights protection. Higher population density should not necessarily lead to higher entry barriers. In fact, as shown in the next section, the data reveal the opposite: Population density has a negative effect on entry costs.

We do not use the fraction of population speaking English or the predicted measure of trade shares (Frankel and Romer, 1999), which have been used by Hall and Jones (1999). Once the five instruments previously described are controlled for, these instruments have no predictive power for entry costs or property rights measures. Therefore, they are not relevant to our analysis.

Because of data availability, our regressions rely on samples of different sizes. The largest sample consists of 123 countries.

MOMENTS OF THE DISTRIBUTION OF FIRMS BY SIZE

In Table 1, we also report statistics for the mean and the variance of the distribution of firms by size, based on Alfaro, Charlton, and Kanczuk (2009). Higher volatility is associated with a lower density of firms (i.e., a larger average firm size) and more heterogeneity in firm size (i.e., a higher variance of the distribution of firms by size). The first two moments of the distribution of firms by size are negatively related to measures of institutional quality and positively correlated with entry costs (Figures 4 and 5).

RESULTS

Endogenous Regressors and Instruments

As a starting point, we identify the minimum number of instruments that allow us to separately identify the effect of entry costs and the effect of property rights on output volatility. Table 2 presents the results of the ordinary least squares regressions of the endogenous regressors on all available instruments. In column 1, entry costs is the dependent variable. The regressors in columns 2 through 5 are the proxies for property rights protection. In column 6, social infrastructure is the dependent variable.

The table shows the correlation patterns of institutional variables with instruments; the differences guide our initial choice of instruments. The European languages variable has an effect on entry costs, but no statistically significant effect on the debt recovery rate, the Heritage Foundation index, expropriation risk, or social infrastructure. Legal origin has no effect on entry costs, but has an effect on the debt recovery rate, the Heritage Foundation index, expropriation risk, and social infrastructure. This suggests that IV regressions that use only the legal origin and European languages variables as instruments might achieve identification. A natural advantage of these regressions is that they do not involve population density or settler mortality and therefore can be implemented on the full sample rather than the subsample of former colonies.

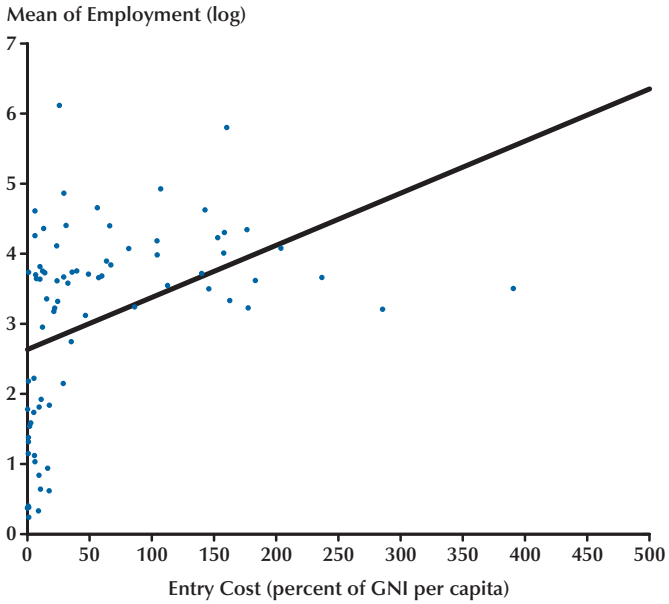
Population density has the expected negative effect on property rights measures and social infrastructure, but its effect on entry costs has the wrong sign.[5] Settler mortality has the expected negative effect on all endogenous regressors. Neither of these variables has a statistically significant effect on constraint on executive power. Because the latter is correlated with the European languages variable and latitude, we consider IV regressions that use population density, settler mortality, and the European languages variables (or latitude) as instruments.

In each of the following IV regressions, we formally test whether the rank condition (A1) is

[5] That is, higher population density implies lower entry barriers.

Figure 4

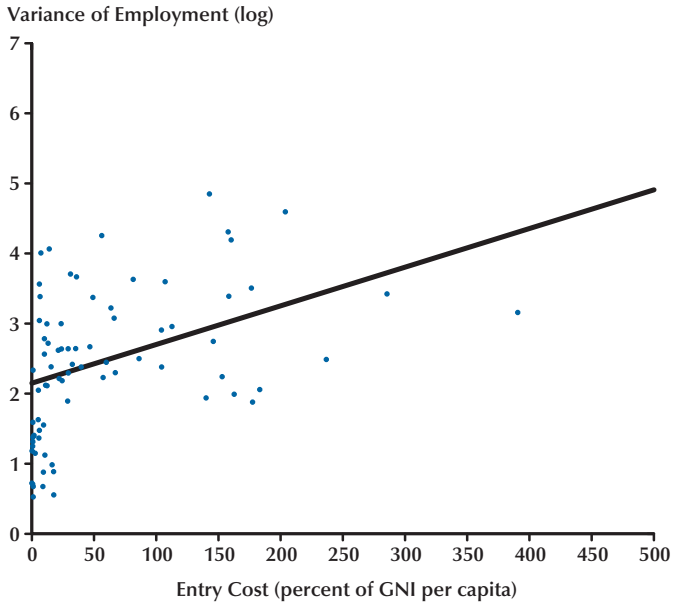
Average Firm Size and Entry Costs: Data and Linear Fit Line (slope 0.74, *p*-value 0.001)



NOTE: Two outliers (entry costs > 500 percent of GNI) excluded.

Figure 5

Variance of Firm Size and Entry Costs: Data and Linear Fit Line (slope 0.55, *p*-value 0.001)



NOTE: Two outliers (entry costs > 500 percent of GNI) excluded.

Table 2

Assessing Instruments: Ordinary Least Squares Regressions of Endogenous Regressors on Instruments

	Endogenous regressors					
Instruments	Entry costs (2)	Debt recovery rate (2)	Constraint on executive power (3)	Heritage Foundation index (4)	Expropriation risk (5)	Social infrastructure (6)
Latitude	–126.13	31.03	2.77	1.74	1.71	0.02
	(99.63)	(20.03)	(1.19)	(0.70)	(1.20)	(0.21)
	0.21	0.13	0.02	0.02	0.16	0.93
European languages	–115.55	–2.20	2.55	0.29	0.42	0.00
	(40.97)	(9.43)	(0.50)	(0.28)	(0.38)	(0.06)
	0.01	0.82	0.00	0.31	0.27	0.96
British legal origin	0.43	11.59	0.31	0.63	0.77	0.07
	(65.94)	(5.14)	(0.35)	(0.18)	(0.34)	(0.04)
	0.99	0.03	0.38	0.00	0.03	0.10
Log population density	–19.44	–2.81	–0.05	–0.18	–0.19	–0.05
	(9.12)	(1.99)	(0.14)	(0.07)	(0.11)	(0.01)
	0.04	0.16	0.74	0.02	0.09	0.00
Log settler mortality	53.19	–6.69	–0.11	–0.15	–0.30	–0.06
	(26.92)	(2.07)	(0.17)	(0.08)	(0.16)	(0.02)
	–0.05	0.00	0.54	0.08	0.07	0.00
Observations	61	61	60	58	61	61
R^2	0.25	0.46	0.56	0.60	0.41	0.52

NOTE: Robust standard errors are reported in parentheses with corresponding *p*-values listed below. See the appendix for data sources and definitions.

satisfied. In addition, when the number of instruments exceeds the number of endogenous regressors, we perform a test for overidentifying restrictions.

RESULTS

Our preliminary regressions are carried out with two instruments: the legal origin and European languages variables. The results of these regressions are reported in columns 1 through 3 of Table 3. In the regressions reported in column 1, property rights are proxied by the debt recovery rate, in column 2 by the Heritage Foundation index, and in column 3 by social infrastructure.

We report three numbers for each instrumented variable: the coefficient, the heteroskedasticity robust standard error (SE), and the corresponding *p*-value. (To save space, the intercept is not reported.) We also report the *p*-value of the Cragg-Donald insufficient rank test (see Cragg and Donald, 1993). The null of this test is that the rank is insufficient. The rejection of the test provides confidence that the rank condition (A1) is satisfied. The number of observations is reported last.

As columns 1 through 3 show, entry costs have a statistically significant adverse effect on volatility. However, the null of the Cragg-Donald

Table 3

Instrumental Variable Regressions of Standard Deviation of Growth Rate of Output Per Worker on Entry Costs and a Measure of Property Rights

	Dependent variable: Standard deviation of growth rate					
	Instruments: Legal origin and European languages			Instruments: Legal origin and latitude		
Independent variable	Debt recovery rate (1)	Heritage Foundation index (2)	Social infrastructure (3)	Debt recovery rate (4)	Heritage Foundation index (5)	Social infrastructure (6)
Entry costs	1.64	1.68	1.53	1.56	1.10	1.14
	(0.58)	(0.77)	(1.53)	(0.94)	(1.00)	(0.62)
	0.00	0.03	0.01	0.10	0.27	0.07
Measure of property rights	–0.03	–0.56	–2.95	–0.03	–0.59	–3.56
	(0.01)	(0.40)	(1.33)	(0.02)	(0.54)	1.33
	0.05	0.16	0.03	0.19	0.28	0.01
Insufficient rank	0.07	0.20	0.23	0.36	0.35	0.29
Observations	121	110	113	123	111	113

NOTE: Heteroskedastic robust standard errors are reported in parentheses with corresponding *p*-values listed below. See the appendix for data sources and definitions.

test is rejected in only one regression,[6] implying that the instruments are not well suited to separately identify the effect of entry costs and property rights. For robustness, we also report the results of these regressions when latitude is used as an instrument instead of European languages (columns 4 through 6). While the results are similar to those reported in columns 1 through 3, the *p*-values of the entry costs coefficient and of the Cragg-Donald test are larger. This is expected given that neither legal origin nor latitude is strongly correlated with entry costs.

Our benchmark regressions use three instruments: settler mortality, population density, and European languages. Columns 1 through 5 of Table 4 show the results for all five proxies of property rights protection. The effect of entry costs in all these regressions is negative and statistically significant. Its magnitude is close to that reported in Table 3. Neither property rights nor social infrastructure has a statistically significant effect in any of these regressions. The null of the Cragg-Donald test is rejected once at the 1 percent level, twice at the 5 percent level, and twice at the 10 percent level. The null of the Hansen-Sargan overidentification test,[7] which is that the exclusion restriction (A2) holds, is not rejected in any of these regressions. This lends credibility to the validity of the instruments. Columns 6 through 10 of Table 4 repeat these regressions but use latitude rather than European languages as an instrument. The results of these regressions are similar to those in columns 1 through 5, but as indicated by the *p*-values of the Cragg-Donald test, this set of instruments is weaker.

[6] The regressions with constraint on executive power and expropriation risk are not reported, because their *p*-values of Cragg and Donald's test are very high.

[7] The Economic Significance of Entry Barriers. See Sargan (1958) and Hansen (1982); see Hayashi (2000) for a textbook treatment.

Table 4

Instrumental Variable Regressions of Standard Deviation of Growth Rate of Output Per Worker on Entry Costs and a Measure of Property Rights

	Dependent variable: Standard deviation of growth rate									
	Instruments: Settler mortality, population density, and European languages					Instruments: Settler mortality, population density, and latitude				
	Debt recovery rate (1)	Constraint on executive power (2)	Heritage Foundation index (3)	Expropriation risk (4)	Social infrastructure (5)	Debt recovery rate (6)	Constraint on executive power (7)	Heritage Foundation index (8)	Expropriation risk (9)	Social infrastructure (10)
Entry costs	1.47	2.03	1.64	1.62	1.61	2.17	1.93	1.91	2.07	2.04
	(0.62)	(0.85)	(0.71)	(0.66)	(0.63)	(0.94)	(0.79)	(0.89)	(0.85)	(0.78)
	0.02	0.02	0.02	0.01	0.01	0.02	0.01	0.03	0.02	0.01
Measure of property rights	–0.01	0.29	–0.02	–0.04	–0.30	0.02	0.07	0.21	0.25	1.51
	(0.03)	(0.28)	(0.45)	(0.40)	(1.89)	(0.04)	(0.26)	(0.47)	(0.46)	(2.29)
	0.60	0.29	0.96	0.92	0.87	0.60	0.80	0.66	0.60	0.51
Insufficient rank	0.04	0.06	0.03	0.10	0.01	0.43	0.08	0.11	0.24	0.05
Overidentification	0.25	0.57	0.85	0.23	0.23	0.62	0.61	0.98	0.61	0.72
Observations	59	58	56	59	59	59	58	56	59	59

NOTE: Heteroskedastic robust standard errors are reported in parentheses with corresponding *p*-values listed below. See the appendix for data sources and definitions.

THE ECONOMIC SIGNIFICANCE OF ENTRY BARRIERS

The results previously described suggest that entry barriers have a significant effect on output volatility. The average value of the entry costs coefficient in the 10 regressions in Table 4 is 1.85. This implies that a 1-SD increase in entry costs in our sample results in a 2.5-percentage-point increase in the SD of the growth rate of output, which is roughly 41 percent of its mean value in our sample.

Entry Costs and Industry Structure

A structural interpretation of our results relies on the seminal work of Hopenhayn (1992). Costlier entry leads to less competition and a lower number of operating firms. With the protection from potential entrants afforded by high entry costs, low-productivity firms can survive and operate. This implies that operating firms are more heterogeneous—that is, a higher dispersion of firm productivity.[8] This mechanism magnifies the volatility stemming from aggregate uncertainty. In the data, the lower density of operating firms and the higher heterogeneity in firm size are associated with higher macroeconomic volatility (Figures 4 and 5). Unfortunately, the paucity of data prevents us from analyzing directly the empirical relationship between entry costs and industry structure in this paper. We leave this task for future research.

Robustness

The effect of entry costs on output volatility is statistically and economically significant, and this result is not driven by an omission of human capital, corruption, or business regulation from the regressions. Moreover, the instruments do not have an independent effect on output volatility, especially those correlated with entry costs. Once entry costs are controlled for, property rights appear to have no effect on output volatility.[9]

While we found no indication that an omitted endogenous regressor biases the results, it is possible that entry costs and property rights capture the effect of other institutions that are correlated with the instruments and affect output volatility. If this were the case, our results should be interpreted as strong evidence for the existence of a set of institutions that are distinct from those related to property rights and that affect output volatility. Entry costs should be viewed as a good proxy for this set of institutions.

Other Volatility Measures

We also investigate whether entry costs affect the magnitude of economic downturns.[10] In Table 5, we perform regressions identical to those in Table 4, except the outcome of interest is the worst drop in output, which is computed as the minimum growth rate of output per worker.

The results of these regressions are in accord with our previous findings: Entry costs have a strong effect on the severity of economic crises in all regressions; property rights protection does not have a significant effect in any regression. The null of the Cragg-Donald test is not rejected (at the 10 percent level) in 7 of the 10 regressions. The null of the overidentification test is not rejected in any regression.

The magnitude of the effect of entry costs on the severity of an economic crisis is very large. The average value of the entry costs coefficient in the 10 regressions in Table 5 is 5.93. This implies that a 1-SD increase in entry costs increases the magnitude of the worst output drop by about 60 percent of its mean value in our sample.

CONCLUSION

Understanding the reasons behind cross-country differences in economic outcomes remains a primary goal of economics. Although recent advances in the literature have identified institutions as major determinants of economic outcomes, little is known about the role and relative

[8] See Barseghyan and DiCecio (2010) for a derivation of this result in a general equilibrium setting.

[9] Barseghyan and DiCecio (2009) report the regression tables (Tables 3C-3E and 4-7B) of these robustness checks.

[10] Our results are also robust to the use of the range of the growth rate of output per worker as a measure of volatility. The corresponding regression table is available from the authors upon request.

Table 5

Instrumental Variable Regressions of Largest Drop of Output Per Worker on Entry Costs and a Measure of Property Rights

	Dependent variable: Standard deviation of growth rate									
	Instruments: Settler mortality, population density, and European languages					Instruments: Settler mortality, population density, and latitude				
	Debt recovery rate (1)	Constraint on executive power (2)	Heritage Foundation index (3)	Expropriation risk (4)	Social infrastructure (5)	Debt recovery rate (6)	Constraint on executive power (7)	Heritage Foundation index (8)	Expropriation risk (9)	Social infrastructure (10)
Entry costs	–4.21	–6.57	–5.08	–4.89	–4.82	–7.35	–6.23	–6.76	–6.82	–6.59
	(1.94)	(2.78)	(2.29)	(2.16)	(2.01)	(3.42)	(2.63)	(3.04)	(3.08)	(2.64)
	0.03	0.02	0.03	0.02	0.02	0.03	0.02	0.03	0.03	0.01
Measure of property rights	0.05	–1.30	0.15	0.01	0.56	–0.10	–0.54	–1.31	–1.24	–6.86
	(0.08)	(1.02)	(1.56)	(1.31)	(6.14)	(0.14)	(1.01)	(1.82)	(1.79)	(8.23)
	0.51	0.20	0.93	1.00	0.93	0.47	0.59	0.47	0.49	0.40
Insufficient rank	0.04	0.06	0.03	0.10	0.01	0.43	0.08	0.11	0.24	0.05
Overidentification	0.18	0.44	0.10	0.13	0.13	0.66	0.56	0.97	0.65	0.80
Observations	59	58	56	59	59	59	58	56	59	59

NOTE: Heteroskedastic robust standard errors are reported in parentheses with corresponding *p*-values listed below. See the appendix for data sources and definitions.

importance of specific institutions. We find that entry regulation is an important determinant of output volatility, while property rights protection is not. These results strengthen the view that entry costs are an important institutional feature and that the effect of institutions on the economy occurs through their impact on industry structure (see, e.g., Nickell, 1996; Acemoglu et al., 2003; Nicoletti and Scarpetta, 2003; Bastos and Nasir, 2004; Sivadasan, 2009; Alesina et al., 2005; Bruhn, 2008; Djankov, Ganser et al., 2010; Barseghyan, 2008).

For policymakers seeking well-defined strategies to stabilize the economies of less-developed countries, our paper provides an additional argument for the elimination of entry barriers: The estimated effect of such a policy is a sizable decrease in output volatility.

REFERENCES

Acemoglu, Daron and Johnson, Simon. "Unbundling Institutions." *Journal of Political Economy*, October 2005, *113*(5), pp. 949-95.

Acemoglu, Daron; Johnson, Simon and Robinson, James A. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, December 2001, *91*(5), pp. 1369-401.

Acemoglu, Daron; Johnson, Simon and Robinson, James A. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, November 2002, *117*(4), pp. 1231-94.

Acemoglu, Daron; Johnson, Simon; Robinson, James A. and Thaicharoen, Yunyong. "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth." *Journal of Monetary Economics*, January 2003, *50*(1), pp. 49-123.

Alesina, Alberto; Ardagna, Silvia; Nicoletti, Giuseppe and Schiantarelli, Fabio. "Regulation and Investment." *Journal of the European Economic Association*, June 2005, *3*(4), pp. 791-825.

Alfaro, Laura; Charlton, Andrew and Kanczuk, Fabio. "Plant-Size Distribution and Cross-Country Income Differences," in Jeffrey Frankel and Christopher Pissarides, eds., *NBER International Seminar on Macroeconomics 2008*. Volume 5. Chicago: University of Chicago Press, 2009.

Barseghyan, Levon. "Entry Costs and Cross-Country Differences in Productivity and Output." *Journal of Economic Growth*, June 2008, *13*(2), pp. 145-67.

Barseghyan, Levon and DiCecio, Riccardo. "Entry Costs, Industry Structure, and Cross-Country Income and TFP Differences." Working Paper No. 2009-005C, Federal Reserve Bank of St. Louis, February 2010; http://research.stlouisfed.org/wp/2009/2009-005.pdf.

Barseghyan, Levon and DiCecio, Riccardo. "Institutional Causes of Macroeconomic Volatility." Working Paper No. 2008-021C, Federal Reserve Bank of St. Louis, May 2009; http://research.stlouisfed.org/wp/2008/2008-021.pdf.

Bastos, Fabiano and Nasir, John. "Productivity and Investment Climate: What Matters Most?" Policy Research Working Paper No. 3335, World Bank, June 2004.

Bruhn, Miriam. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico." Policy Research Working Paper No. 4538, World Bank, January 2008.

Cragg, John G. and Donald, Stephen G. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory*, June 1993, *9*(2), pp. 222-40.

Djankov, Simeon; Ganser, Tim; McLiesh, Caralee; Ramalho, Rita and Shleifer, Andrei. "The Effect of Corporate Taxes on Investment and Entrepreneurship." *American Economic Journal: Macroeconomics*, 2010, forthcoming.

Djankov, Simeon; La Porta, Rafael; Lopez de Silanes, Florencio and Shleifer, Andrei. "The Regulation of Entry." *Quarterly Journal of Economics*, February 2002, *117*(1), pp. 1-37.

Dollar, David and Kraay, Aart. "Institutions, Trade, and Growth." *Journal of Monetary Economics*, January 2003, *50*(1), pp. 133-62.

Easterly, William and Levine, Ross. "Tropics, Germs, and Crops: How Endowments Influence Economic Development." *Journal of Monetary Economics*, January 2003, *50*(1), pp. 3-39.

Frankel, Jeffrey A. and Romer, David. "Does Trade Cause Growth?" *American Economic Review*, June 1999, *89*(3), pp. 379-99.

Glaeser, Edward L.; La Porta, Rafael; Lopez de Silanes, Florencio and Shleifer, Andrei. "Do Institutions Cause Growth?" *Journal of Economic Growth*, September 2004, *9*(3), pp. 271-303.

Gunnemark, Erik V. *Countries, Peoples, and Their Languages: The Geolinguistic Handbook*. Dallas: Summer Institute of Linguistics, Inc., 1991.

Hall, Robert E. and Jones, Charles I. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, 1999, *114*(1), pp. 83-116.

Hansen, Lars P. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, February 1982, *50*(4), pp. 1029-54.

Hayashi, Fumio. *Econometrics*. Princeton, NJ: Princeton University Press, 2000.

The Heritage Foundation. *2006 Index of Economic Freedom*. Washington, DC, and New York: The Heritage Foundation and *The Wall Street Journal*, 2006.

Heston, Alan; Summers, Robert and Aten, Bettina. "Penn World Table Version 6.2." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Hopenhayn, Hugo A. "Entry, Exit, and Firm Dynamics in Long-Run Equilibrium." *Econometrica*, September 1992, *60*(5), pp. 1127-50.

Hunter, Barbara F., ed. *Ethnologue: Languages of the World*. Twelfth Edition. Gothenburg: Lanstryckeriet, 1992.

Jaggers, Keith and Marshall, Monty G. *Polity IV Project*. Center for International Development and Conflict Management, University of Maryland, 2000; www.systemicpeace.org/polity/polity4.htm.

Knack, Stephen and Keefer, Philip. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures." *Economics and Politics*, November 1995, *7*(3), pp. 207-27.

Koren, Miklós and Tenreyro, Silvana. "Volatility and Development." *Quarterly Journal of Economics*, February 2007, *122*(1), pp. 243-87.

Kraay, Aart and Ventura, Jaume. "Comparative Advantage and the Cross-Section of Business Cycles." *Journal of the European Economic Association*, December 2007, *5*(6), pp. 1300-33.

La Porta, Rafael; Lopez de Silanes, Florencio; Shleifer, Andrei and Vishny, Robert. "The Quality of Government." *Journal of Law, Economics, and Organization*, March 1999, *15*(1), pp. 222-79.

Mauro, Paolo. "Corruption and Growth." *Quarterly Journal of Economics*, August 1995, *110*(3), pp. 681-712.

Nickell, Stephen J. "Competition and Corporate Performance." *Journal of Political Economy*, August 1996, *104*(4), pp. 724-46.

Nicoletti, Giuseppe and Scarpetta, Stefano. "Regulation, Productivity and Growth: OECD Evidence." *Economic Policy*, April 2003, *18*(36), pp. 9-72.

Political Risk Services. *Political Risk Yearbook*. East Syracuse, NY: PRS Group, 1999.

Rodrik, Dani; Subramanian, Arvind and Trebbi, Francesco. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth*, June 2004, *9*(2), pp. 131-65.

Sachs, Jeffrey D. and Warner, Andrew M. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity*, 1995, *26*(1), pp. 1-95.

Sargan, John D. "The Estimation of Economic Relationships Using Instrumental Variables." *Econometrica*, April 1958, *26*(3), pp. 393-415.

Sivadasan, Jagadeesh. "Barriers to Competition and Productivity: Evidence from India." *B.E. Journal of Economic Analysis & Policy*, 2009, *9*(1), Advances, Article 42.

Tybout, James R. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, March 2000, *38*(1), pp. 11-44.

World Bank. *Doing Business in 2004—Understanding Regulation*. Washington, DC: International Bank for Reconstruction and Development/World Bank, 2004.

World Bank. *Doing Business in 2005—Removing Obstacles to Growth*. Washington, DC: International Bank for Reconstruction and Development/World Bank, 2005.

World Bank. *Doing Business in 2006—Creating Jobs*. Washington, DC: International Bank for Reconstruction and Development/World Bank, 2006a.

World Bank. *Doing Business in 2007—How to Reform*. Washington, DC: International Bank for Reconstruction and Development/World Bank, 2006b.

World Bank. *Doing Business in 2008*. Washington, DC: The International Bank for Reconstruction and Development/World Bank, 2007.

APPENDIX: DATA SOURCES AND DEFINITIONS

Entry Costs (The World Bank, 2004, 2005, 2006a,b, 2007)[11]

Entry costs are constructed for "a 'standardized' firm which has the following characteristics: (1) it performs general industrial or commercial activities, it operates in the largest city (by population), (2) it is exempt from industry-specific requirements (including environmental ones), it does not participate in foreign trade and does not trade in goods that are subject to excise taxes (e.g., liquor, tobacco, gas), it is a domestically-owned limited liability company, (3) its capital is subscribed in cash (not in-kind contributions) and is the higher of (i) 10 times GDP per capita in 1999 or (ii) the minimum capital requirement for the particular type of business entity, it rents (i.e., does not own) land and business premises, it has between 5 and 50 employees one month after the commencement of operations, all of whom are nationals, it has turnover of up to 10 times its start-up capital, and it does not qualify for investment incentives."

Debt Recovery Rate (The World Bank, 2004, 2005, 2006a,b, 2007)

The recovery rate is recorded as cents on the dollar recovered by claimants' creditors, tax authorities, and employees through the bankruptcy proceedings. The calculation takes into account whether the business is kept as a going concern during the proceedings, as well as bankruptcy costs and the loss in value due to the time spent closing down.

Purchasing Power Parity–Adjusted GDP Per Worker (Center for International Comparisons of Production, Income and Prices, University of Pennsylvania—Penn World Table 6.2)[12]

Constraint on Executive Power (*Polity IV Project*, Jaggers and Marshall, 2000)[13]

This variable "refers to the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities," and takes values from 1 to 7, where 1 = unlimited authority; 3 = slight to moderate limitations; 5 = substantial limitations; and 7 = executive parity (between the executive(s) and accountability groups) or subordination. For more details, see the *Polity IV Project* manual.

Property Rights Protection Index (Based on the Heritage Foundation's 2006 Index of Economic Freedom dataset)[14]

From 1 to 5 (in the regressions, the scale is reversed, e.g., 5 = 1 and 1 = 5):

1. Private property guaranteed by government; court system efficiently enforces contracts; justice system punishes those who unlawfully confiscate private property; corruption nearly nonexistent, and expropriation highly unlikely.
2. Private property guaranteed by government; court system suffers delays and is lax in enforcing contracts; corruption possible but rare; expropriation unlikely.
3. Court system inefficient and subject to delays; corruption may be present; judiciary may be influenced by other branches of government; expropriation possible but rare.

---

[11] Available at www.doingbusiness.org/.

[12] Available at http://pwt.econ.upenn.edu/.

[13] Available at www.systemicpeace.org/polity/polity4.htm.

[14] Available at www.heritage.org/Index/.

4. Property ownership weakly protected; court system inefficient; corruption present; judiciary influenced by other branches of government; expropriation possible.
5. Private property outlawed or not protected; almost all property belongs to the state; country in such chaos (for example, because of ongoing war) that property protection nonexistent; judiciary so corrupt that property not effectively protected; expropriation frequent.

The index is constructed based on the following factors: (i) freedom from government influence over the judicial system; (ii) commercial code defining contracts; (iii) sanctioning of foreign arbitration of contract disputes; (iv) government expropriation of property; (v) corruption within the judiciary; (vi) delays in receiving judicial decisions and/or enforcement; and (vii) legally granted and protected private property.

Protection Against Expropriation Risk (Acemoglu, Johnson, and Robinson, 2001)

Risk of expropriation of private foreign investment, from 0 to 10. Higher score means less risk. Original source: Political Risk Services (September 1999).

Social Infrastructure (Hall and Jones, 1999)

Hall and Jones constructed this measure as an average of the openness to trade index and the Government Anti-Diversion Policies (GADP) index. The openness to trade index was taken from Sachs and Warner (1995). The GADP index is an equal-weighted average of five indices: (i) law and order, (ii) bureaucratic quality, (iii) corruption, (iv) risk of expropriation, and (v) government repudiation of contracts. All of these were taken from Political Risk Services.

European Settler Mortality (Acemoglu, Johnson, and Robinson, 2001)

Estimated mortality for European settlers during the early period of European colonization (before 1850).

Population Density in 1500 (Acemoglu, Johnson, and Robinson, 2002)

Indigenous population density in 1500, indicated as inhabitants per square kilometer.

Fraction of Population Speaking a Major European Language (Hall and Jones, 1999, based on Gunnemark, 1991, and Hunter, 1992)

Latitude (La Porta et al., 1999)

The absolute value of the latitude of the country, scaled to values between 0 and 1. Original source: *CIA World Factbook*.

Government Corruption Variable (La Porta et al., 1999)

"Low ratings indicate 'high government officials are likely to demand special payments' and 'illegal payments are generally expected through lower levels of government' in the form of bribes connected with import and export licenses, exchange controls, tax assessment, policy protection, or loans." Scale 0 to 10. Average value over 1972-95. Original source: *International Country Risk Guide*, produced by Political Risk Services; www.prsgroup.com/icrg.aspx.

Business Regulation (La Porta et al., 1999)

This index ranges from 1 to 5. The index is constructed based on the following factors: (i) licensing requirements to operate a business; (ii) ease of obtaining a business license; (iii) corruption within the

bureaucracy; (iv) labor regulations, such as established workweeks, paid vacations, and parental leave, as well as selected labor regulations; (v) environmental, consumer safety, and worker health regulations; and (vi) regulations that impose a burden on business. Original source: The Heritage Foundation's Index of Economic Freedom dataset (2006).

Moments of the Distribution of Employment by Size Class Across Countries (Alfaro, Charlton, and Kanczuk, 2009)

These data are constructed from microdata collected in Dun & Bradstreet's *WorldBase*. The unit of observation is the plant.

For our cross-sectional study, only one observation is needed for each of the variables above. For entry costs and the debt recovery rate, we take the average over the five years (2004-08) for which data are available. For the constraint on executive power variable and the property rights index, we average over the last 10 years in which they were reported: 1994-2003 and 1996-2005, respectively. For the expropriation risk variable, we use the average over 1985-1995.

Ideally, the averages over the same period of time for all variables would be used. Unfortunately, this is not possible because of data limitations. For some countries data for one or more years might be missing. We ignore these years when constructing averages.[15]

[15] When constructing the averages for constraint on executive power, interregnum and transitional periods are ignored, except for the Democratic Republic of the Congo (Kinshasa). Because all years between 1994 and 2003 were classified as interregnum or transitional for this country, we use the value for year 1991, the last year for which constraint on executive power was recorded.