FEDERAL RESERVE BANK OF ST. LOUIS

REVIEW

THIRD QUARTER 2016
VOLUME 98 | NUMBER 3

REVIEW

Volume 98 • Number 3

Review

Review is published four times per year by the Research Division of the Federal Reserve Bank of St. Louis. Complimentary print subscriptions are available to U.S. addresses only. Full online access is available to all, free of charge.

Online Access to Current and Past Issues

The current issue and past issues dating back to 1967 may be accessed through our Research Division website: **http://research.stlouisfed.org/publications/review**. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* also are available to our readers on this website.

Review articles published before 1967 may be accessed through our digital archive, FRASER: **http://fraser.stlouisfed.org/publication/?pid=820**.

Review is indexed in Fed in Print, the catalog of Federal Reserve publications (**http://www.fedinprint.org/**), and in IDEAS/RePEc, the free online bibliography hosted by the Research Division (**http://ideas.repec.org/**).

Authorship and Disclaimer

The majority of research published in *Review* is authored by economists on staff at the Federal Reserve Bank of St. Louis. Visiting scholars and others affiliated with the St. Louis Fed or the Federal Reserve System occasionally provide content as well. *Review* does not accept unsolicited manuscripts for publication.

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

Subscriptions and Alerts

Single-copy subscriptions (U.S. addresses only) are available free of charge. Subscribe here: **https://research.stlouisfed.org/publications/review/subscribe/**.

Our monthly email newsletter keeps you informed when new issues of *Review*, *Economic Synopses*, *The Regional Economist*, and other publications become available; it also alerts you to new or enhanced data and information services provided by the St. Louis Fed. Subscribe to the newsletter here: **http://research.stlouisfed.org/newsletter-subscribe.html**.

Economic Data

General economic data can be obtained through FRED® (Federal Reserve Economic Data), our free database with about 300,000 national, international, and regional data series, including data for our own Eighth Federal Reserve District. You may access FRED through our website: **http://research.stlouisfed.org/fred2**.

ISSN 0014-9187

Sales of Distressed Residential Property: What Have We Learned from Recent Research?

Jeffrey P. Cohen, Cletus C. Coughlin, and Vincent W. Yao

During the housing bust many homeowners found themselves "underwater"—the amount they owed on their mortgages exceeded the value of the associated property—and they either could not or possibly chose not to stay current on their mortgage payments. As a consequence, sales of so-called distressed properties, often after a foreclosure, became commonplace. This spurred numerous research papers on various related issues. The authors' review summarizes the research findings on three topics: the impact of changes in housing prices on foreclosures; the impact of foreclosure on the sales price of the foreclosed house; and the impact of foreclosure on the sales prices of nearby houses. Not surprisingly, declining housing prices are associated with increasing foreclosure rates; however, various other factors, such as a job loss or expected housing prices, can also play an important role. This review highlights various theoretical and econometric issues that have raised doubts about the accuracy of estimated price impacts of foreclosures and led to numerous refinements of the subsequent empirical analysis. Estimates of the own foreclosure discount generally range from near zero to 28 percent, with most estimates greater than 12 percent. Estimates of the discount resulting from spillover effects of nearby foreclosed houses are generally less than 2 percent and diminish rapidly with distance. (JEL R31)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2016, *98*(3), pp. 159-88.
http://dx.doi.org/10.20955/r.2016.159-188

Developments in the residential housing market had major impacts on overall U.S. economic activity in the run-up to the Great Recession and its aftermath.[1] The housing boom was characterized by liberal credit availability, high rates of construction, and rapid price increases that increased the wealth and consumption of many homeowners. However, during the housing bust many homeowners became "underwater"—the amount they owed on their mortgages exceeded the value of the associated property—and they either could not or possibly chose not to stay current on their mortgage payments. As a consequence, sales of so-called distressed properties, often after a foreclosure, became commonplace.

Jeffrey P. Cohen is an associate professor of real estate and finance in the School of Business at the University of Connecticut. Cletus C. Coughlin is a senior vice president and chief of staff at the Federal Reserve Bank of St. Louis. Vincent Yao is an associate professor of real estate in the Robinson College of Business at Georgia State University. Jonas Crews provided research assistance.

Distressed sales can be viewed as an anomaly in most housing markets. They are not typical arm's-length transactions and generally account for a small subset of housing transactions. However, the number and relative share of distressed sales rose substantially during the housing bust. Not surprisingly, this spurred much interest in various aspects of these sales.

Generally speaking, distressed property is sold in one of the following ways: (i) As an alternative to foreclosure, the lender allows a short sale (i.e., the proceeds of the sale are less than the amount owed on the property) by the borrower. (ii) The lender initiates the foreclosure process under a notice of default and the property is sold during the process by the borrower. Or (iii) the lender forecloses on the property, takes title, and then sells the property as real estate owned (REO).[2] As discussed in Clauretie and Daneshvary (2011), these alternatives present lenders with trade-offs involving various costs, such as the price discount and marketing time.

Concerning the third option, there are two methods for foreclosing on a property: judicial and nonjudicial.[3] When no power of sale clause is required in the state, the mortgage holder must file a lawsuit and obtain court approval to foreclose. Once granted, the property can be sold. A nonjudicial foreclosure is allowed when a power of sale clause is required in the state. The property owner is given a period to become current on his or her payment status and another period before the foreclosed property goes on the market. As a result, the time required to implement judicial foreclosures tends to be longer than for nonjudicial foreclosures.[4]

The bursting of the housing bubble led to numerous research articles examining various empirical and theoretical issues relating to the sale of distressed residential property. Given the accumulation of research, now is an appropriate time to take stock of what we have learned.[5] Our review cannot be characterized as exhaustive, as we focus on three topics: the impact of changes in housing prices on foreclosures; the impact of foreclosure on the sales price of the foreclosed house; and the impact of foreclosure on the sales prices of nearby houses. Prior to examining the research on these topics, we provide some background information on the housing market and the basic issues that we review.

AN OVERVIEW OF THE RESIDENTIAL HOUSING MARKET OVER THE PAST 20 YEARS

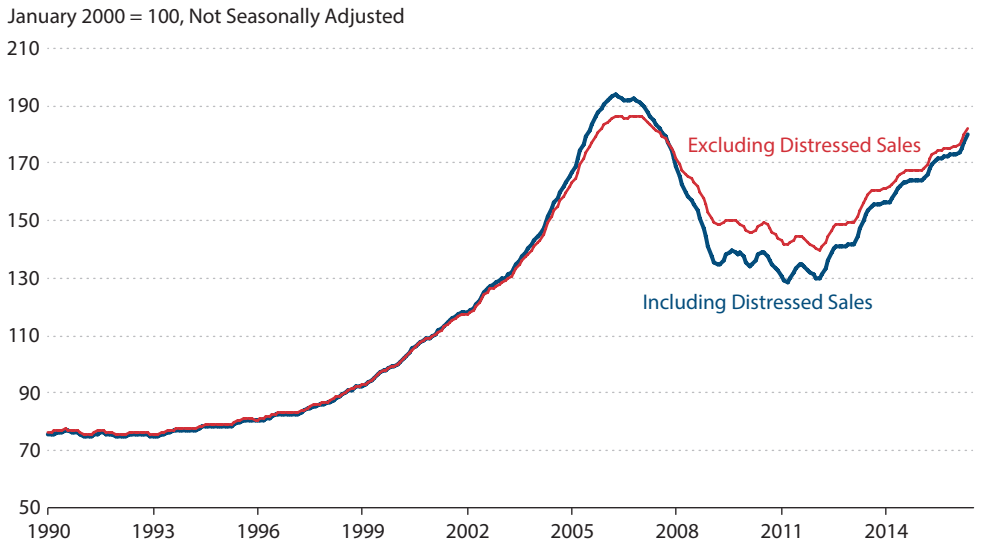
To provide some context for our review, we begin by summarizing some basic information about the housing market over recent years. This information is organized into five categories—prices, foreclosures, homeownership, construction, and sales.

The Rise and Fall in House Prices

Housing prices began to accelerate in the late 1990s. According to the CoreLogic National Home Price Index, housing prices peaked in April 2006.[6] The level of this index, including and excluding distressed sales, is shown in Figure 1, while year-over-year changes are shown in Figure 2. Focusing on the measure of the index including distressed sales (i.e., the CoreLogic National Home Price Index in the figures), Figure 1 highlights the rapid run-up in housing

Figure 1

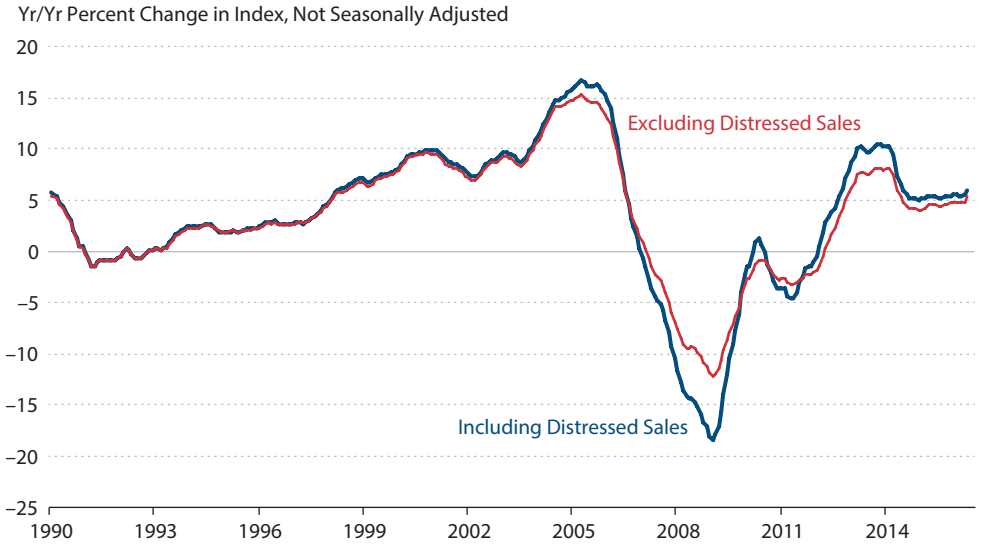
Home Price Indexes



SOURCE: CoreLogic/Haver Analytics.

Figure 2

Home Price Growth



SOURCE: CoreLogic/Haver Analytics.

prices. Figure 2 shows that during this run-up, year-over-year increases consistently exceeded 5 percent from January 1998 and reached as high as 16.6 percent in April 2005, much faster than the rate of increase in consumer prices.[7]

Most agree that housing prices in the mid-2000s reflected a bubble; however, there is much disagreement as to the causes of the bubble. Common explanations include the following: excess credit supply, excessively accommodative monetary policy, a global savings glut, government policies encouraging homeownership, irrationally optimistic beliefs about future housing price appreciation, inelastic housing supply, and an excess of mispriced mortgage finance.[8] Ultimately, the bubble burst. Beginning in April 2006, housing prices declined until reaching a trough in March 2011. During this period housing prices declined 34 percent.

Figure 1 also provides suggestive evidence that the prices of distressed sales have deviated from nondistressed sales, especially around and during the bust. Movements in the two indexes are quite similar until 2004. At that time, which is near the end of the housing boom, the price index including distressed sales rose faster than the corresponding index excluding distressed sales. This puzzling fact has not been addressed in the literature. After the peak in housing prices in April 2006, the index including distressed sales declined more rapidly than the index excluding distressed sales. In other words, sharper declines occurred in the prices of distressed property than in nondistressed property. The magnitude of this differential price behavior has drawn the attention of many researchers.

During recent years the two indexes have risen similarly. Including distressed sales, from March 2011 housing prices have risen steadily (roughly 40 percent) through May 2016. One development that differentiates the recent run-up in housing prices from previous run-ups in the postwar era is that it is not driven by increased demand for owner-occupied housing.[9] The decline in homeownership rates highlighted below suggests that private and institutional investors have found opportunities to take advantage of the current environment for housing purchases.

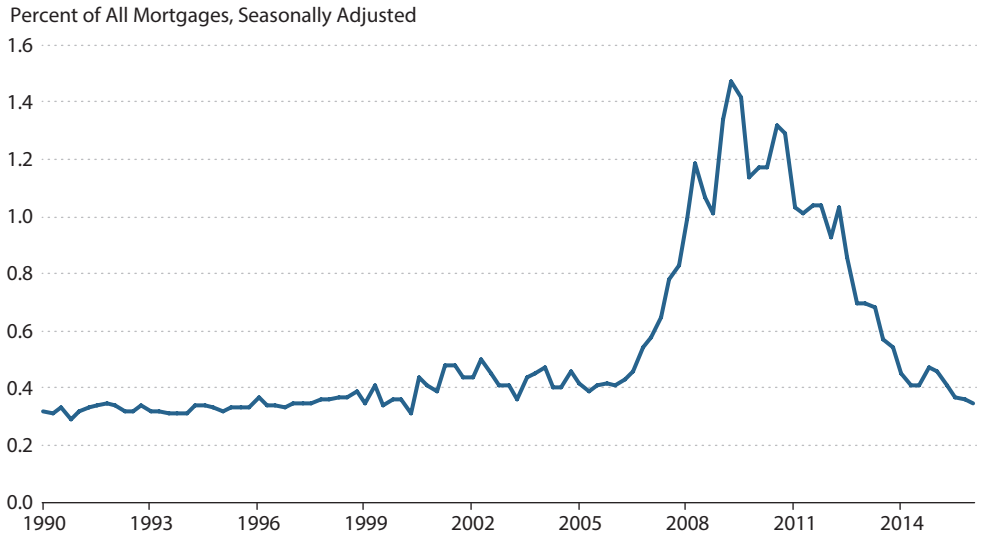
The Rise and Fall in Foreclosures

Declining housing prices have both been caused by foreclosures and contributed to foreclosures. Figure 3 shows the rise and subsequent decline in new foreclosures during the financial crisis. From quarterly levels generally less than 0.50 percent before 2007, new foreclosures rose rapidly during 2007, 2008, and early 2009 as housing prices declined, reaching a peak of 1.5 percent in 2009:Q2.[10] However, because the foreclosure process entails costs for the lender, lenders likely factor in numerous considerations in their foreclosure decisions. Negotiations with the borrower might be a preferred route for the lender.

Legal obligations also affect the results. During the bust, Mian, Sufi, and Trebbi (2015) found that those states without a judicial requirement had twice the foreclosure rates of those states with a judicial requirement. Subsequently, the national rate of new foreclosures has declined to less than 0.5 percent and has remained there since 2014.[11] Coinciding with this normalization of foreclosure rates, Mian, Sufi, and Trebbi (2015) found that the foreclosure rates in judicial and nonjudicial states had converged.

Figure 3

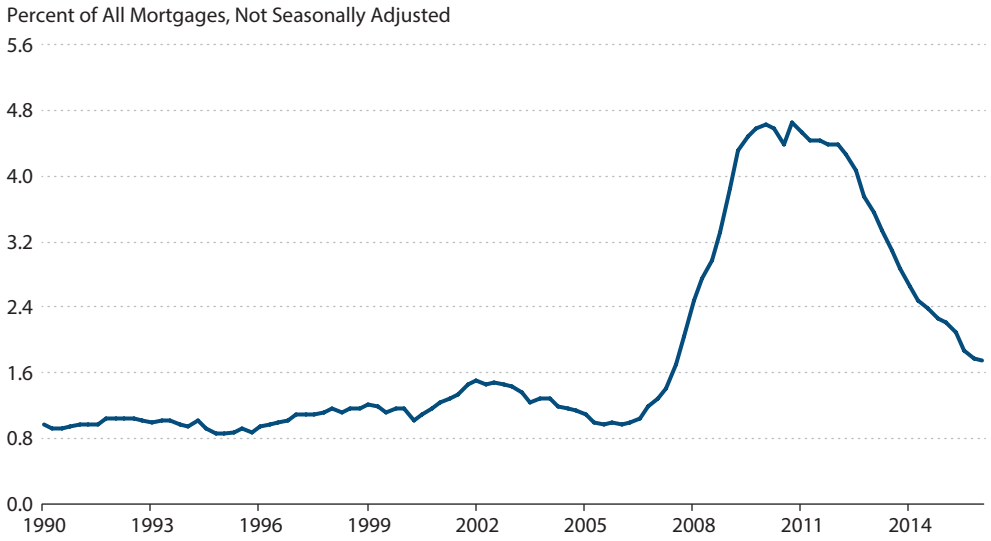
New Foreclosures



SOURCE: Mortgage Bankers Association/Haver Analytics.

Figure 4

Mortgages in Foreclosure



SOURCE: Mortgage Bankers Association/Haver Analytics.

Figure 4 shows another dimension of foreclosures by showing the total number of loans in the legal process of foreclosure as a percentage of the total number of mortgages in a specific quarter. From rates of roughly 1 percent, this rate rose sharply during 2006-10, peaking at 4.6 percent in 2010:Q4. While this inventory declined to 1.7 percent in 2016:Q1, it remains above the levels of the mid-2000s.

Mayer, Pence, and Sherlund (2009) explain the rise in subprime mortgage defaults and suggest that the relaxed underwriting standards—manifested most dramatically by lenders allowing borrowers to forgo down payments entirely—and stagnant to falling house prices in many parts of the country appear to be the most immediate contributors to the rise in mortgage defaults. The reason for the surge in defaults for mortgages with low or no documentation is due mostly to underwriting that had deteriorated along other dimensions. However, subprime defaults are not the entire story. Ferreira and Gyourko (2015) reinterpret the U.S. foreclosure crisis as more of a prime, rather than a subprime, borrower issue. They find that traditional mortgage default factors associated with the economic cycle, such as negative equity, completely account for the foreclosure propensity of prime borrowers relative to all-cash owners and for three-quarters of the analogous subprime gap.

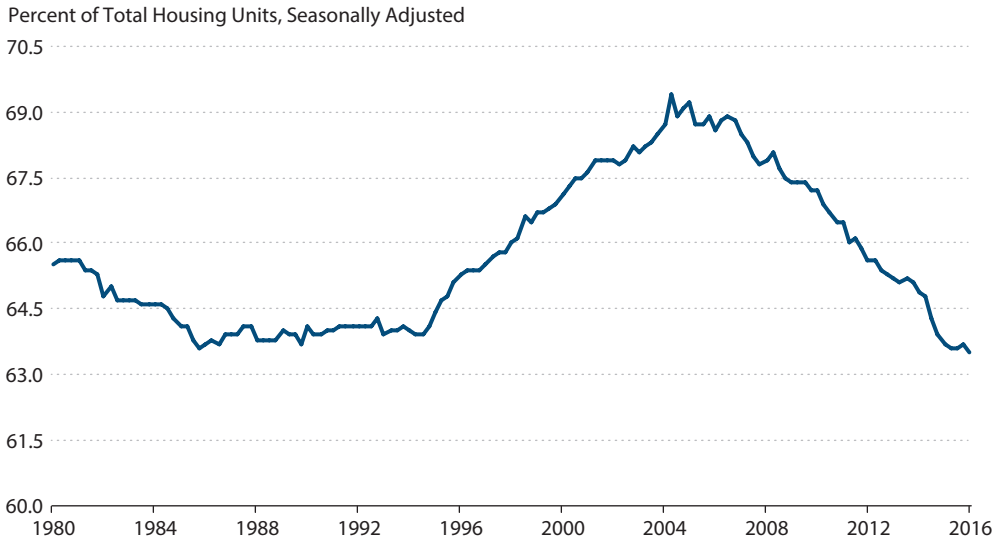
The Rise and Fall in Homeownership

In the mid-1990s national leaders began a broad effort to increase homeownership, which is defined as the percentage of homes owned by their occupants. This bipartisan effort began during the Clinton administration and was later embraced by the Bush administration.[12] As shown in Figure 5, using seasonally adjusted rates, this effort generated increased homeownership, albeit temporarily. After maintaining a rate of roughly 64 percent for nearly 10 years, homeownership began to rise in the mid-1990s, rising in a consistent manner until reaching 69.4 percent in 2004:Q2. The demand associated with this rising homeownership propelled a housing market boom that stimulated rapid growth overall in the United States during this period. With the advent of the housing crisis and recession, however, homeownership began a nearly continuous descent that led to its lowest level in the past 36 years, 63.1 percent in 2016:Q2.

This declining homeownership leads to many questions, none of which we examine thoroughly here. For example, what is a normal homeownership rate? The housing bubble suggests that 69.4 percent in 2004:Q2 is abnormal. Thus, one should expect the rate to be lower than 69.4 percent, but how much lower? Also, are the rates in the mid-1990s (i.e., prior to the housing boom) a reasonable guide? If so, then rates of roughly 64 percent are reasonable. A related question concerning homeownership, especially if one considers the current homeownership rate to be too low, is why more renters aren't becoming homeowners. Fuster, Zafar, and Cocci (2014) explore whether there might be changed attitudes toward housing or whether the answer might be due to a combination of low incomes, weak personal finances, and difficulties in securing mortgages. Their conclusion is that the latter combination of factors is more accurate.

Figure 5

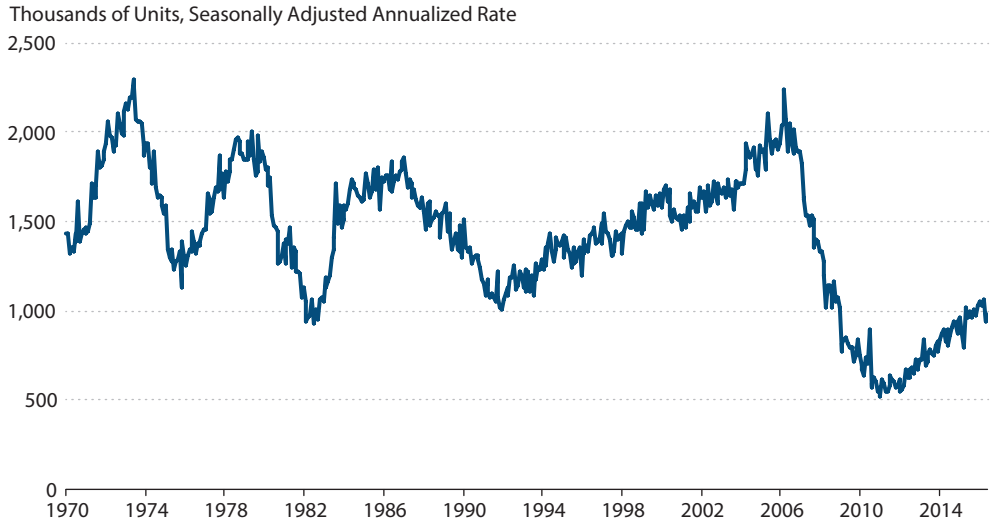
Homeownership Rate



SOURCE: Census Bureau; FRED®, Federal Reserve Bank of St. Louis.

Figure 6

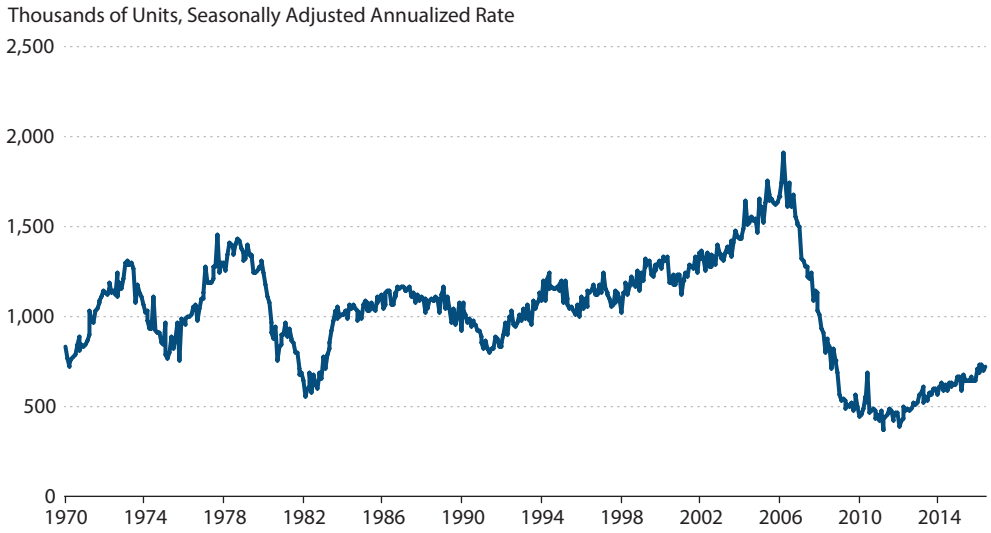
Housing Completions



SOURCE: Census Bureau; FRED®, Federal Reserve Bank of St. Louis.

Figure 7

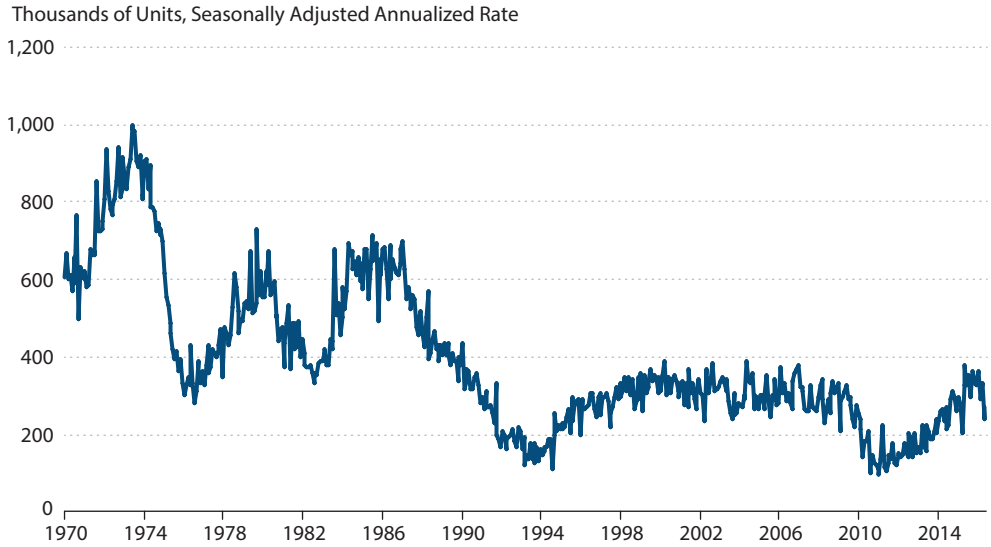
Single-Family Housing Completions



SOURCE: Census Bureau; FRED®, Federal Reserve Bank of St. Louis.

Figure 8

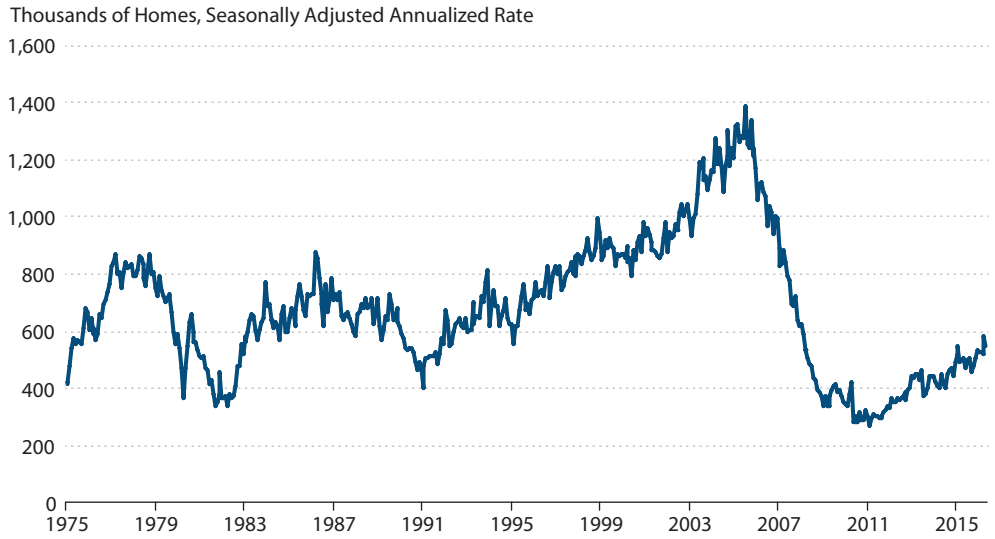
Multifamily Housing Completions



SOURCE: Census Bureau; FRED®, Federal Reserve Bank of St. Louis.

Figure 9

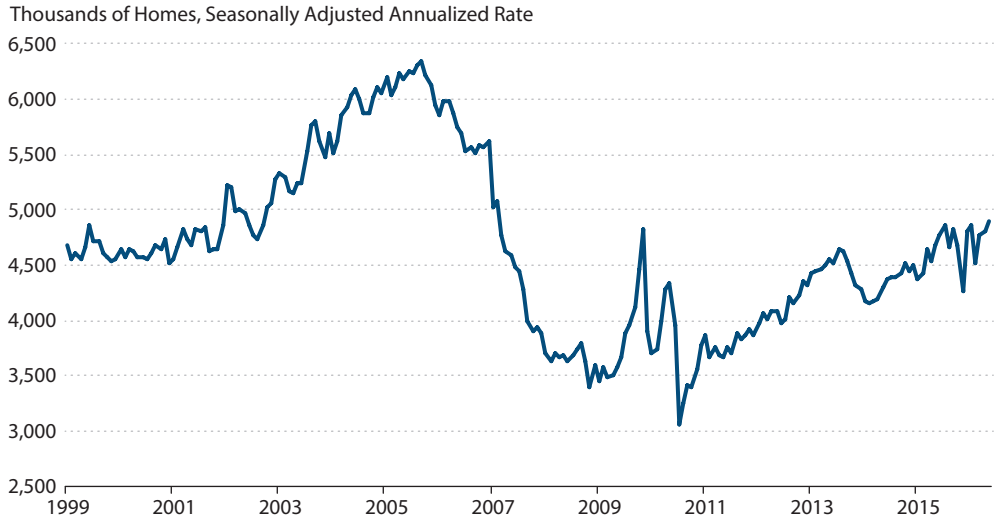
New Single-Family Home Sales



SOURCE: Census Bureau; FRED®, Federal Reserve Bank of St. Louis.

Figure 10

Existing Single-Family Home Sales



SOURCE: National Association of Realtors; FRED®, Federal Reserve Bank of St. Louis.

Changes in Construction of Single-Family Versus Multifamily Housing

Not surprisingly, housing construction activity has mirrored the boom and bust in housing prices. Housing construction trended upward during the boom and dropped precipitously during the bust. Moreover, the recovery of housing construction from the housing/financial crisis has been slow. Figure 6 shows that housing completions have trended upward since early 2011, but that current levels remain far below the levels in the late 1990s/early 2000s.

As the current expansion continues, one might anticipate that the long-run prospects are much more favorable for multifamily housing than single-family housing. As stressed by Rappaport (2013), reduced population growth, which tends to reduce the demand for housing, and the aging of the Baby Boomers, which tends to shift demand toward multifamily from single-family housing, are key demographic factors supporting such an outlook. Figures 7 and 8 indicate that construction has shifted in such a direction. As shown by Figure 7, completion of single-family structures has increased only marginally since 2010, while Figure 8 shows that the completion of multifamily structures has increased substantially. Moreover, the single-family completions remain far below levels prior to the boom, while multifamily completions have returned to levels comparable to those seen during the early and mid-2000s.

Housing Sales

The relatively low levels of completed single-family houses are, not surprisingly, reflected in the sales of new, single-family houses. Despite some recovery in recent years, Figure 9 shows that such sales remain weak compared with the levels in the late 1990s/early 2000s. In terms of sales of existing single-family houses, as shown in Figure 10, the recovery has been far from steady but has reached levels existing in the late 1990s/early 2000s.

EXAMINING FORECLOSURES: SOME BASIC ISSUES

Statistical evidence reveals a negative correlation between foreclosures and house prices. In other words, foreclosures increase (decrease) when house prices decrease (increase). A simple correlation, however, does not answer very basic questions about the direction of causality. Both researchers and policymakers would like answers to the following questions: (i) Do declining house prices cause increased foreclosures? (ii) Do increased foreclosures cause declining house prices? It is easy to provide reasons to suggest that the answer to both questions is yes. If that is the case, then any estimation of the impact of declining house prices on foreclosures or the impact of foreclosures on house prices must address this possibility.[13] This issue is discussed in more depth later when we examine specific studies.

If the answer to either of these questions is yes, then the quantitative relationship is also of interest. For example, if housing prices decline by a given percentage, what is the associated percentage increase in foreclosures? Similarly, what is the impact of a foreclosure on the price of the foreclosed house as well as on the prices of nearby houses? Furthermore, understanding the underlying mechanisms that connect foreclosures and prices can be useful for designing policy remedies. This topic is discussed immediately below. The answer to the first question is likely yes because declines in house prices contribute to a rise in foreclosures by

putting more homeowners underwater. However, because the foreclosure process is costly and various financial conditions and expectations influence both borrower and lender behavior, being underwater is not a sufficient condition for foreclosure.

Concerning the second question, if increased foreclosures cause declining house prices, the following questions come to mind: What mechanisms produce this result? Also, is the decline in price restricted to the foreclosed property or are the values of nearby homes affected as well? With respect to the foreclosed property itself, at issue is whether the discount is due to a "stigma effect" or a "proxy effect." The former reflects a discount for no reason other than the status of the property as foreclosed. Meanwhile, the proxy effect refers to a discount caused by other characteristics that may affect prices negatively, such as deteriorated physical condition of the property and/or neighborhood conditions. Moreover, sellers of foreclosed or soon-to-be-foreclosed properties may be highly motivated (i.e., have a lower reservation price or accept a lower selling price) because of a desire for shorter marketing time, lower direct and indirect carrying costs of the property, or the seller's need for liquidity.

Regarding the impact on nearby property, one possibility is that foreclosed properties are a disamenity in that they can be an eyesore (because of a lack of maintenance) or induce crime and vandalism.[14] A second possibility is through a competitive effect. A foreclosed property adds to the supply of houses available for purchase and this increased supply can lead to lower prices. Would one expect the foreclosure effect to be temporary or permanent? After the foreclosed property is sold, the eyesore/crime issues should be eliminated, but is there a lasting neighborhood effect?

An article relevant to most of the econometric studies that we examine is Coulson and Zabel (2013). Their focus is on the consequences of disequilibrium for hedonic estimations.[15] Given the large number of foreclosures during 2007-11, it is reasonable to argue that during recent years the housing market in many cities was not in equilibrium. These authors suggest several approaches for controlling for potential disequilibria in hedonic housing price models.[16] First, a dummy variable can be added to the hedonic house price regression, which equals 1 if a property is a foreclosure and 0 otherwise. This foreclosure dummy can also be interacted with the other explanatory variables in the hedonic regression to control for the disequilibrium impacts on the housing characteristics and other neighborhood amenities/disamenities. Finally, a variable can be added to represent the distance from other foreclosures, which can control for disequilibria resulting from large numbers of nearby foreclosures. These methods are closely related to our discussion in the following section on foreclosures and house prices.

The Impact of Changing Housing Prices on Foreclosures

Foreclosures are important events that define the ultimate default. The literature on the determinants of mortgage default has been evolving for 30 years. Foster and Van Order (1984, 1985) were among the first to model default as a "put option."[17] When a homeowner has a mortgage and can extinguish his or her obligation by relinquishing the house to the lender, the owner has a put option as well as equity in the house. The put option's value is a function of the drift and volatility of house prices: It is more valuable when house prices are more likely

to fall, and more valuable when house prices are more volatile. That is because, under both circumstances, the probability that the house will be worth less than the mortgage, and therefore in the money, is greater. The Foster and Van Order articles used house price volatility as a covariate for predicting default and found it was a significant and important predictor of default. Many subsequent articles recognize a put option in the money is only a necessary, but not sufficient, condition of default and identify other factors besides home prices that explain the default.[18]

Gerardi, Shapiro, and Willen (2008) found that price declines beginning in the summer of 2005 were the dominant factor in causing a large increase in foreclosures during the downturn (2006 and 2007) for subprime borrowers in Massachusetts. This focus on subprime borrowers is partially due to the fact that homeowners with subprime mortgages are six times more likely to end up in foreclosure than those with prime mortgages. Similarly, Bajari, Chu, and Park (2008) found that changing home prices were a significant determinant of the probability of default of subprime and Alt-A mortgages nationwide.[19]

In another study focused on property in Massachusetts, Fisher and Lambie-Hanson (2012) study a suburb of Boston—Chelsea, Massachusetts—and assess how various factors, including house prices and whether properties are investor owned, affect the probability of foreclosure. They find that local investor-owned properties had a foreclosure rate that was nearly double that of owner-occupied and nonlocal investment properties.[20] They also find a lower foreclosure probability when there is greater house price appreciation (and vice versa).

While declining housing prices contribute to homeowner distress and, therefore, are likely to be associated with mortgage default and foreclosure, the existing literature reveals that mortgage default is a nuanced topic. Theory and empirical evidence suggest that negative home equity is necessary, but not sufficient, for triggering a default. In other words, not all households with negative home equity end up in default. For example, Foote, Gerardi, and Willen (2008), using a dataset of Massachusetts homeowners, found that fewer than 10 percent of borrowers likely to have had negative equity at year-end 1991 actually experienced a foreclosure in the following three years. As such, one needs a model of the default decision to underpin the empirical analysis of the connection between housing prices and distressed sales. Numerous models have been developed; however, a thorough scrutiny of these models is beyond the scope of this review.[21] We restrict our analysis to selected results that provide some insights into the nuances.

One idea that has been explored is the "double-trigger" theory of default. Default is said to be triggered by coinciding events—the borrower experiences both negative equity and an adverse life event, such as a job loss. However, as stressed by Foote, Gerardi, and Willen (2008) and others, the double-trigger theory has been found to be lacking as a sufficient explanation for defaults. In other words, the existence of the two triggers does not guarantee default. Thus, one must provide additional theory concerning the conditions under which the double-trigger model is likely to fail. Given forward-looking agents, the expected changes in a house's price are likely to play a key role in a household's default decision. An expectation of an increasing price is likely to deter a default, while an expectation of a decreasing price is likely to increase the probability of default.

Credit constraints are another factor that can play a role in the default decision. Campbell and Cocco (2015) find that households with high loan-to-value ratios at mortgage origination are more likely to experience negative home equity when house prices decline. The level of negative home equity that produces a default depends on the extent to which households are borrowing constrained. Households with high loan-to-income ratios are subject to tighter borrowing constraints. Across mortgage types, defaults by households with adjustable-rate mortgages increase when nominal interest rates increase and when these households suffer adverse idiosyncratic labor income shocks. Defaults by households with fixed-rate mortgages are higher when interest rates and inflation are low. Finally, interest-only mortgages trade off an increased likelihood of negative home equity against a relaxation of borrowing constraints.

Guiso, Sapienza, and Zingales (2013) use survey data to measure households' propensity to default on underwater mortgages even if they can afford to pay them (strategic default). They find that homeowners' willingness to default increases in both the absolute and the relative size of the home equity shortfall. They also find that this willingness is affected by both pecuniary and nonpecuniary factors, such as views about fairness and morality.

Gerardi, Herkenhoff, Ohanian, and Willen (2015) find that households experiencing a job loss, divorce, or large medical expenses are more likely to default. However, a larger percentage of financially distressed households do not default. For example, 80 percent of unemployed households with negligible savings (i.e., less than one month of mortgage payments) were found to be current on their mortgage payments. Moreover, the role of strategic default appears to be minimal. First, defaulting households generally have relatively low net asset levels, and second, high-wealth households with underwater mortgages generally choose not to default.

The Impact of Foreclosure on the Sales Price of the Foreclosed House

Here we examine the effect of a foreclosure on the sales price of the distressed (foreclosed) property itself. Ideally, the estimate of the foreclosure discount would be the difference in the sales price of the house sold in a normal transaction and the sales price of the same house sold under distress. Obviously, both prices cannot be observed. This leads to the classic "treatment effect" problem.

A normal transaction implies that the transaction would be undertaken at fair market value.[22] Fair market value is the price that would occur in a competitive housing market. A competitive housing market is characterized by many buyers and sellers, relevant information that is equally available to buyers and sellers, and access to financial resources without regulatory or institutional barriers. In addition, related to the previous characteristics, houses must be on the market for a sufficient period to allow for a market determination of the equilibrium price.

A forced sale associated with a foreclosure does not meet the necessary conditions for a competitive housing market. Unlike a voluntary transaction, in a forced sale one of the participants is not entirely a willing participant. Moreover, the buyer possesses less information about the property than in a normal housing market transaction. This lack of information could be due to a time constraint on the date of sale or there could be sale procedures that

preclude on-site inspections of the property. Forced sales are also not as widely advertised as normal sales. Finally, forced sales differ from normal sales in that financing options are more limited for the former than the latter. All these considerations tend to reduce the price of forced sales relative to the price that would occur in a competitive market.[23]

Generally, the mortgagee acquires the residential property through the foreclosure process. One could argue that these REO properties should sell for their market value. A fundamental question is whether the bank or financial institution would behave in the same manner as the seller in a normal transaction. Possibly the bank or financial institution would value the time on the market more highly than a normal seller. One simple reason is that the bank or financial institution is holding a vacant property, while that may not be the case in many normal transactions. If so, this consideration could lead to a lower price. However, such a price difference does not necessarily indicate the existence of potential excess returns.

At issue is whether the discount is due to a stigma effect or a proxy effect. As discussed earlier, the former reflects a discount for no reason other than the status of the property as foreclosed. No characteristics differentiate the foreclosed property from a non-foreclosed property. A large stigma effect suggests the possibility of excess returns for potential buyers who can purchase and then resell quickly, capturing a windfall. While housing markets are not perfectly efficient, it is hard to believe that such returns could be large. Meanwhile, the proxy effect refers to a discount related to other characteristics that may affect price negatively, such as a deteriorated physical condition and/or neighborhood conditions. One reason that foreclosed properties sell at a discount is that they are in worse condition than nearby properties. Moreover, sellers of foreclosed or to-be-foreclosed properties may be highly motivated (i.e., have a lower reservation price or accept a lower selling price) because of their desire for shorter marketing time, lower direct and indirect carrying costs of the property, or the seller's need for liquidity. As a result, the foreclosure status variable is a proxy for other omitted variables. Omitted variables may produce a biased estimate of a pure foreclosure (stigma) effect.[24] For example, vacant houses sell for lower prices. To the extent that foreclosed houses are more likely vacant, omitting this variable can lead to a biased foreclosure effect. Another example involves cash transactions, which often lead to lower sale prices. If foreclosed sales tend to consist of relatively more cash sales than non-foreclosed sales, then omitting this variable can lead to a biased foreclosure estimate.

Zabel (2014) develops a dynamic model of the housing market where he allows for the possibility of vacancies that are part of the stochastic process in the regression's error terms, while controlling for excess demand and excess supply (i.e., disequilibria). When he estimates his model with annual metropolitan statistical area (MSA) data for the United States for the years 1990-2011, he finds that both excess demand and excess supply respond more to changes in the market during 2006-11. This model sheds light on the issue of disequilibria in housing markets and how prices respond to these disequilibria.

An additional complication is that lenders face a choice concerning how to handle defaults by way of short sales, sales during foreclosures, or REO sales. Clauretie and Daneshvary (2011) find that the price discount is highest for REO transactions (13.5 percent) and is smallest for short sales (5.6 percent). However, there are also time-on-the-market costs. The short sales option has the highest costs associated with marketing time.

Table 1

Own Price Decline of a Foreclosure

Authors (date)	Geographic area	Period	Estimation	Estimated decline* (%)
Shilling, Benjamin, & Sirmans (1990)	Baton Rouge, LA	1985	OLS	21.3
Forgey, Rutherford, & VanBuskirk (1994)	Arlington, TX	1991-93	OLS	20.4
Hardin & Wolverton (1996)	Phoenix, AZ	1993-94	OLS	22.2
Springer (1996)	Arlington, TX	1989-93	OLS	3.7
Carroll, Clauretie, & Neill (1997)	Las Vegas, NV	1990-93	OLS	0.17 to 2.6†
Pennington-Cross (2006)	United States	1995-99	Calculations‡	2§
Clauretie & Daneshvary (2009)	Las Vegas, NV	2004-07	GS2SLS	7.5
Campbell, Giglio, & Pathak (2011)	Massachusetts	1987-2008	OLS	27.6
Clauretie & Daneshvary (2011)	Las Vegas, NV	2007-08	3SLS	13.5
Harding, Rosenblatt, & Yao (2012)	13 MSAs	1990-2008	Various	No excess returns
Siebert (2015)	Hollywood, FL	2000-08	WLS	4.7
Siebert (2015)	Fort Lauderdale, FL	2000-08	WLS	12
Siebert (2005)	Lafayette, IN	2000-08	WLS	16.1

NOTE: GS2SLS, generalized spatial two-stage least squares; OLS, ordinary least squares; 3SLS, three-stage least squares; WLS, weighted least squares.

* The estimate presented is not the only estimate contained in these articles. Our goal is to provide comparability across studies in terms of the focus of our review. We present exact estimates—by calculating $100*(e^b - 1)$, where b is the coefficient associated with the foreclosure dummy in log-linear models, in all instances where such a calculation is applicable. Any discrepancies between the estimates provided and those in the corresponding articles can be attributed to this conversion.

† The estimates were statistically insignificant.

‡ The author calculates the percent appreciation in house prices in a foreclosed home's MSA from the origination of the foreclosed home's mortgage through the post-foreclosure sale of the home, and subtracts from that the foreclosed home's price appreciation over that period to obtain a measure of the appreciation discount for the home. The author then calculates the average discount.

§ The author reports an appreciation discount of 22 percent, which results in an overall price discount of 2 percent for the sample average MSA-level appreciation from mortgage origination through post-foreclosure sale (10 percent).

Estimates of the foreclosure discount in the literature we survey range from near zero to 28 percent across studies that cover different geographic areas in the United States and span various years—as early as 1985 up through 2008. These estimates are summarized in Table 1, and we elaborate on the details of some of these studies below.

We briefly examine a number of the early studies on the foreclosure discount.[25] Shilling, Benjamin, and Sirmans (1990) find a 21.3 percent discount on foreclosed condos in Baton Rouge, Louisiana, in 1985. They stress the sellers' desire to sell quickly to avoid carrying costs and the buyers' requirement of a discount to compensate for carrying costs prior to leasing. A similar discount was estimated by Forgey, Rutherford, and VanBuskirk (1994). They found a 20.4 percent discount on foreclosed single-family properties in Arlington, Texas, from 1991 to 1993.

Hardin and Wolverton (1996) also estimated a foreclosure discount in excess of 20 percent. They found a 22 percent discount on foreclosed apartment complexes in Phoenix,

Arizona, in 1993-94, which they attribute to seller motivation. This finding is in contrast to that of Springer (1996), who found a 4 percent foreclosure discount on single-family houses in Arlington, Texas. This study accounted for motivation of the seller, but not endogeneity of time on the market, property condition, or cash sales. Carroll, Clauretie, and Neill (1997) found a discount of 0.17 to 2.6 percent on residential properties in Las Vegas, Nevada, during 1990-93. As a foreshadowing of the conclusions in Clauretie and Daneshvary (2009), they argue that the larger estimates in other articles result from failing to control for neighborhood quality. In a study of foreclosed single-family properties nationwide from 1995-99, Pennington-Cross (2006) found a 2 percent price discount that resulted from a 22 percent discount in the appreciation of foreclosed homes, relative to their respective MSAs, from the origination date of the mortgage through the post-foreclosure sale of the property.

Clauretie and Daneshvary (2009) examined distressed sales in Las Vegas for the period covering November 2004 through November 2007. They show that by accounting for certain variables, such as the physical condition of the property and the relationship between marketing time and price, plus correcting for two types of spatial price interdependence, the previous estimates of the foreclosure discount are biased high.[26-28] In their preferred estimation, which controls for property condition, occupancy status, and payment method, in addition to commonly controlled for characteristics, the foreclosure effect is 7.5 percent. Without these additional controls, the foreclosure effect is 10 percent. Thus, the size of potential excess returns is much smaller than other studies have suggested.

Campbell, Giglio, and Pathak (2011), using a typical hedonic regression, find large foreclosure discounts, about 28 percent on average, with larger discounts for houses in low-quality neighborhoods. In a more recent study in this area, Siebert (2015) found that foreclosed homes in Hollywood, Florida; Fort Lauderdale, Florida; and Lafayette, Indiana, sold for 4.7 percent, 12 percent, and 16.1 percent less, respectively. The vast majority (e.g., 92 percent) of these differences is the result of a proxy effect of lower quality. Therefore, very little relates to a motivation by REO owners for a quick sale to avoid forgone investment opportunities. Siebert (2015) also found much heterogeneity across house size, price, and geographic area. Also, for a topic we examine next, Siebert found that the existence of any nearby foreclosed homes has a negative impact on the values of non-foreclosed properties, with the effect ranging from 0.8 percent to 4.7 percent.

Large discounts—certainly those in excess of 20 percent—suggest the possibility of large returns from buying and then shortly thereafter reselling the property. Using a large sample of repeat sales pairs, Harding, Rosenblatt, and Yao (2012) find that REO purchasers do not earn excess returns—in other words, the real estate market operates efficiently. Thus, there is no evidence that banks are selling houses at fire-sale prices. REO properties and buyers differ from their counterparts in the non-distressed market segment (which can be considered an "endowment effect"), and the attribute prices of REO properties differ from those of non-distressed properties (i.e., a "coefficient effect"). Each of these factors accounts for roughly half of the price difference.

The Impact of Foreclosures on the Sales Prices of Nearby Houses

From a microeconomic perspective, increased homeowner distress (foreclosures) could be causing declining housing prices of nearby properties.[29] However, in the absence of a foreclosure discount, there is no reason to expect a negative price impact in the form of an externality. Given a foreclosure discount, a key empirical challenge is disentangling the supply effect of foreclosures from the potential disamenity effect of foreclosures. Note that the disamenity can be viewed as a reduction in neighborhood quality. However, the supply effect of another house on the market might not be fully portrayed by the characteristics of the house if the house has been allowed to physically deteriorate. This underscores the importance of the assumption that hedonic house price models represent equilibrium prices. The literature reveals much diversity in terms of the geographic (i.e., local) and temporal (i.e., up to 5 years) impacts of foreclosures. Nearby foreclosures do decrease the sales prices of nearby nondistressed properties. A standard finding is that this effect decreases rapidly over both distance and time. The variation in foreclosure discounts and spillover estimates is a result of differences in data, geographies and time periods, and the underlying empirical models.[30] The use of a repeat sales approach is more appropriate than a standard hedonic approach; however, modified hedonic approaches can generate insights when repeat sales data are unavailable.

One of the first studies to estimate the foreclosure externality was that of Immergluck and Smith (2006a) (Table 2). Focusing on Chicago in the late 1990s, they estimated the effects of foreclosures 1 to 2 years after they occurred and found that a foreclosure causes a 0.9 percent decline in house values for all homes within a ⅛-mile radius.[31] The percentage impact is larger for low- and moderate-income areas. A shortcoming of this article, noted by Lee (2008), is the lack of adequate handling of reverse causation.

An important issue is how the price impact of foreclosures might change as the number of foreclosures in a neighborhood increases. Using data on New York City, Been (2008) found that the marginal spillover effects of additional properties with pending foreclosure petitions tend to diminish. This negative effect shrinks with both time and distance. Been's work on how the price effect changes with time and distance has been extended by Lin, Rosenblatt, and Yao (2009). They found that for conforming mortgages, foreclosures have a clear negative impact on prices of local houses. This effect is larger during a downturn than during other times, which suggests one must control for the stage of the housing cycle. Lin, Rosenblatt, and Yao (2009) found their largest effect of foreclosure of 8.7 percent for closely neighboring properties during a bust year: This effect diminishes with distance and time but can last up to five years after the foreclosure.

Further extensions of this literature were made by Leonard and Murdoch (2009), who argue that neighborhood quality can be viewed as a local public good and is an important determinant of housing prices. Changes in nearby foreclosures indicate changes in neighborhood quality. In the Dallas, Texas, area they found a negative effect of housing distress on prices, an effect that decreases as distance from the foreclosed property increases. In models controlling for both spatial dependence in housing prices and in the errors, the authors find that an additional foreclosure within 250 feet of a sale affects the selling price of an average

Table 2

Price Decline of a Nearby Foreclosure*

Authors	Area	Proximity†	Time proximity‡	Period§	Estimation	Interpretation‖	Estimated decline# (%)
Immergluck & Smith (2006a)	Chicago	660 ft	Foreclosure 2 yrs before	Late 1990s	OLS	Per foreclosure	0.9
Been (2008)	New York	500 ft	Foreclosure 2 yrs before	2000-05	OLS	First/second foreclosure (D)	1.8
Schuetz, Been, & Ellen (2008)	New York	250 ft	Foreclosure 18 mos before	2000-05	OLS	Any foreclosure (D)	0.8
Lin, Rosenblatt, & Yao (2009)	Chicago	330 ft	Foreclosure 2 yrs before	2004-06	Heckman	Per foreclosure	8.7
Lin, Rosenblatt, & Yao (2009)	Chicago	330 ft	Foreclosure 2 yrs before	2001-03	Heckman	Per foreclosure	5.0
Leonard & Murdoch (2009)	Dallas	250 ft	Foreclosure in same calendar yr or yr before**	2005-06	GMM	Per foreclosure	0.83
Rogers & Winter (2009)	St. Louis	600 ft	Foreclosure 6 mos before	1998-2007	GMM	First foreclosure	0.66
Harding, Rosenblatt, & Yao (2009)	140 Zip codes	300 ft	Foreclosure 3 mos before	1989-2007	OLS	Per foreclosure	1.1
Campbell, Giglio, & Pathak (2011)	Massachusetts	260 ft	Foreclosure 1 yr before	1987-2009	Weighted OLS	Per foreclosure	0.85††
Daneshvary & Clauretie (2012)	Las Vegas	528 ft	REO sale 3 mos before	2008-09	GS2SLS	Per foreclosure	1.1
Whitaker & Fitzpatrick (2013)	Cuyahoga, OH	500 ft	Public sale 1 yr before	2009-11	GMM	Per foreclosure	4.6‡‡
Hartley (2014)	Chicago	260 ft	Foreclosure 1 yr before	1999-2011	OLS	Per foreclosure	1.3§§
Anenberg & Kung (2014)	4 MSAs	528 ft	REO listed while property listed‖	2007-09	OLS	Per foreclosure	1.6##
Turnbull & van der Vlist (2014)	Orange County, FL	1,320 ft	Post-foreclosure sale 90 days before/after	2007-12	OLS	Per foreclosure	0.8***
Siebert (2015)	Florida	Neighborhood	Same calendar yr	2000-08	OLS	Any foreclosure (D)	0.8
Siebert (2015)	Indiana	Neighborhood	Same calendar yr	2008-08	OLS	Any foreclosure (D)	4.7
Fisher, Lambie-Hanson, & Willen (2015)	Boston	Same address and association	Foreclosure active at sale date†††	1987-2012	OLS	Per foreclosure	2.5‡‡‡
Gerardi, Rosenblatt, Willen, & Yao (2015)	15 MSAs	528 ft	REO at time of sale	2001-10	OLS	Per foreclosure	1§§§

See notes on p. 177.

Table 2, cont'd

Price Decline of a Nearby Foreclosure*

NOTES:

* The estimate presented is not the only estimate contained in these papers. Our goal is to provide comparability across studies in terms of the focus of our review.

† Proximity refers to the maximum distance, not always physical, between the properties being analyzed and nearby foreclosures.

‡ Time proximity generally refers to the maximum amount of time between some key point during the foreclosure process (including, but not limited to, the actual date of foreclosure, date of REO sale, and so on), and the sale date of the home. When the sale date of the analyzed home is not the time marker in the paper, the alternative time marker is provided. Footnotes have been added for situations in which the time proximity requires more description.

§ A study's time period is considered by the authors of this paper to span the beginning to the end of the period created by overlapping the time periods for the sales data and foreclosures data used to produce the estimates.

‖ The interpretation column provides the proper way to interpret the estimates provided. To add additional clarity, a (D) is added when the estimate is derived from the coefficient attached to a dummy variable.

As in Table 1, whenever log-linear models are used, we provide exact estimates of the price effect instead of the rough estimates provided by coefficients.

** For empirical reasons, the authors used only foreclosures that occurred within the same year as a home sale (2006) and foreclosures from the previous year (2005), given that the properties foreclosed on in the previous year were also foreclosed on during the year following the home sale (2007).

†† This estimate results from using both a variable counting the number of foreclosures the year before the sale and a variable counting the number of foreclosures the year after the sale, and subtracting the former from the latter. The authors argue that the coefficient on the after variable captures the effect of economic shocks that result in a noncausal negative relationship between the number of foreclosures before a sale and the sales price. Subtracting the after-sale coefficient from the before-sale coefficient removes such shocks.

‡‡ In later models, the authors demonstrated that the reported estimate is inflated.

§§ This estimate corresponds specifically to the effect of single-family foreclosures on sales of other single-family homes. Assuming the market for housing is segmented into single-family and multifamily units, then a supply effect explains more than 90 percent of the overall price decline.

‖‖ The price effect is measured by estimating the change in a home's list price that occurs when an REO home is listed nearby.

The authors attribute almost all of the price decline to a supply effect. Disamenity effects were found only for neighborhoods with high housing density and low property values, and the effect was roughly 1.5 percent.

*** The estimate consists of a 0.5 percent disamenity effect and a 0.3 percent competitive effect.

††† A foreclosure is active, according to the authors, during the year before the actual foreclosure date and the two years following that date.

‡‡‡ While the author produced estimates for externalities of multiple different types of foreclosed properties onto condos, this estimate specifically refers to the effect of foreclosed condos on condos in the same association and with the same address (same building). The authors argue that, because condos in the same association should be substitutes for one another, the fact that there is a very small and statistically insignificant effect on condo prices if a condo in the same association, but at a different address, is foreclosed gives evidence that this reported effect is largely a disamenity effect.

§§§ Their models explicitly measure the effect on home price growth, and they find that such a property experiences a growth rate that is 1 percent less for each foreclosure or seriously delinquent property in excess of the numbers of such properties present when the home was sold the first time. Because the average appreciation from the first to the second sale was 0 percent, and most houses had neither a foreclosure nor seriously delinquent property nearby at the time of each sale, the 1 percent reduction in price growth can be roughly interpreted as a 1 percent reduction in sales price.

($200,000) house negatively by $1,666 (the direct effect is $1,000, while the total effect is $1,666).

For St. Louis County, Missouri, Rogers and Winter (2009) found that foreclosures have a negative (1 percent or less) impact on prices. Surprisingly, the marginal impact on prices of additional foreclosures *declines* as foreclosures increase. They acknowledge the simultaneity issue but were unable to find an instrument for foreclosures.

In an analysis of foreclosures in New York City from 2000 to 2005, Schuetz, Been, and Ellen (2008) found that proximity to foreclosed properties was associated with reduced sales prices and that the magnitude of the discount increased with the number of foreclosed properties, albeit not in a linear manner. In addition, the authors found evidence of a threshold effect (i.e., being near a small number of foreclosed properties may not have a price impact) and found that housing prices were lower, even before the foreclosures, in neighborhoods in which foreclosures occurred. Thus, failure to control for this latter possibility will produce selection bias.

In a recent study of home prices in Massachusetts, Campbell, Giglio, and Pathak (2011) use a novel identification strategy with hedonic regressions and find that each nearby foreclosure (i.e., within a radius of 260 feet) lowers the selling price of a non-foreclosed house by roughly 1 percent or more. In another closely related article, Harding, Rosenblatt, and Yao (2009) use a repeat sales approach to address the reverse causality and simultaneity issue between local home price trends and foreclosures in the immediate neighborhood. Their estimated discount is roughly 1 percent per nearby foreclosed property. This discount tends to vanish rapidly as the distance from the distressed property increases.

Daneshvary and Clauretie (2012) use single-family detached home transactions from January 2008 through June 2009 in Las Vegas, Nevada, and find foreclosure spillover effects much larger than those found for the same market in previous studies, ranging from 1.1 percent to 2.9 percent per foreclosure. The new results are attributed to controlling for the overall trend in market prices, the neighborhood average price, and unobserved neighborhood characteristics. No additional effect from short sales is found.

Hartley (2014) argued that foreclosure externalities work through two channels: an increase in supply and a disamenity effect if the property is not maintained or is vandalized, possibly while vacant. Both of these channels are likely to have negative effects on sale prices. As a result, a failure to control for the supply effect will likely lead to an overestimate of the disamenity effect. He found that each single-family home foreclosure within 260 feet led to a 1.3 percent price reduction in single-family houses. Also, foreclosures of multifamily units do not exert spillover effects on single-family units. Assuming segmentation, then any spillover effect in an average neighborhood is almost completely the result of the supply effect.

Foreclosures have a causal effect on nearby house prices, according to Anenberg and Kung (2014). The price effects are due to competition (i.e., an additional house for sale) and disamenities. Competition effects are important in all parts of a geographic area, while disamenity effects are found only in high-density, low-price neighborhoods. Also, while REO properties have a negative impact on nearby houses for sale, the effect is only slightly more pronounced than that of non-REO sales.

Fisher, Lambie-Hanson, and Willen (2015) find that a foreclosed condo leads to a 2.5 percent reduction in sales price for a condo in the same association and at the same address, while it has virtually no price effect for a condo in the same association but at a different address. Because condos in the same association can be considered close substitutes, the authors conclude that the foreclosure causes a price decline through a disamenity effect rather than a supply effect.

Turnbull and van der Vlist (2014) use data from Orange County, Florida. They separate the effects of foreclosures and new construction and find that nearby foreclosures reduce property prices. Their disamenity externality estimate is 0.5 percent.

Recent research has begun to take a closer look at the externality issue by attempting to estimate differences across submarkets. For example, Whitaker and Fitzpatrick (2013) estimate the impacts of foreclosures, as well as two other features of the market related to foreclosures—property tax delinquency and house vacancy—on the value of neighboring houses in high- and low-poverty submarkets. Using sales data from low-poverty submarkets in Cuyahoga County, Ohio, the authors find that an additional property within 500 feet that is vacant or delinquent, but not foreclosed, is associated with a reduction in a neighboring house's selling price of 1 percent or 2 percent, respectively. In the same submarkets, the negative impact of a home being vacant and tax delinquent, but not foreclosed, is 4.6 percent.

Mian, Sufi, and Trebbi (2015) use the difference in state foreclosure laws to address endogeneity. They find that a one-standard-deviation increase in the average number of foreclosures per homeowner results in a 5 to 7 percent decline in house prices over two years. They use listing data to show that foreclosures lead to a net increase in housing inventory at the zip code level, and note that this finding complements the theory that foreclosures lower neighboring house prices largely through a supply effect.

Gerardi, Rosenblatt, Willen, and Yao (2015), using data covering 15 large MSAs, provide new evidence on the size and source of the externalities. They find that the temporal impact of the externality extends from the time when the borrower becomes seriously delinquent until well after the bank sells the property.[32] Non-distressed properties within 0.1 miles of a seriously delinquent or foreclosed property sell, on average, for 1 percent less per distressed property.[33] This decline is sensitive to the condition of the foreclosed property, with those in poor condition having a much larger negative effect (2.6 percent) than those in better condition. These spillovers shrink rapidly with distance and disappear completely within one year after the bank sells the property.

Our focus in this section of our review has been on the price effects of foreclosures. A closely related issue is the possibility of foreclosure contagion. Mortgage defaults are contagious if a given default increases the default probability of another mortgage on a nearby property. Harding, Rosenblatt, and Yao (2009) found that the contagion effect grows from the onset of borrower distress through the foreclosure sale, with the effect stabilizing roughly when the lender's sale to the third party occurs. The focus of Towe and Lawley (2013) is on how a foreclosed property affects the probability of foreclosure of a neighboring property. They estimate that the probability of another default increases by 18 percent. Goodstein et al. (2011), after controlling for borrower and loan characteristics, local demographic and eco-

nomic conditions, and changes in property values, find that the mortgage default probability increases by as much as 24 percent given a one-standard-deviation increase in the zip-code-level foreclosure rate. Finally, Rauterkus et al. (2012) address whether there is a tipping point in foreclosure rates. In other words, is there a foreclosure rate above which the foreclosure rate increases at an increasing rate? If so, an area, say a neighborhood, may be at risk of failure. The authors find evidence of contagion using data for Chicago from 2003 to 2008, but it is restricted to a small subset of markets.

An important issue for policymakers is how to mitigate the impacts of any negative externality. To design effective policy instruments, it is important to explore the transmission mechanisms of foreclosure contagions. Gerardi, Rosenblatt, Willen, and Yao (2015) find that the contagion effects are worse for poorly maintained distressed properties. Their results indicate the important role of disinvestment and the value of policies to transition from delinquencies to foreclosures quickly so that normal homeowners can resume the maintenance. The implied strategy is for lenders and government to avoid fire sales or dumping foreclosures to the market all at once. Hartley's (2014) results also support the notion that a supply effect plays a more important role in the channels of contagion, while the disamenity effect is near zero.

A recent article by Cheung, Cunningham, and Meltzer (2014) examines the possible role that a homeowners association might play. A homeowners association, by monitoring foreclosed property and ensuring some minimal levels of maintenance, can reduce the magnitude of the negative externality. Properties in neighborhoods with homeowners associations were found to be less affected by homeowner distress than properties in neighborhoods without homeowner associations. Another relevant article for policymakers is by Gangel, Seiler, and Collins (2013). They found that the size of the foreclosure contagion effect is not as important for market collapse as the time a foreclosed property remains unsold (i.e., stays on the market).

CONCLUDING COMMENTS

The effect of the Great Financial Crisis on housing markets and foreclosures is a key focus of this article. Given the long-lasting effects that foreclosures can have on the health of neighborhoods, major issues of concern are what happened with respect to the spatial aspects of foreclosures and what can we learn from these effects. Increased understanding and accurate empirical relationships provide the foundations for designing policy responses.

An extensive literature exists for the three major topics examined in this article—specifically, (i) how housing price changes affect foreclosures; (ii) how a foreclosure affects the sales price of the foreclosed house; and (iii) how foreclosures in the vicinity of a property affect the sales price of this nearby house. Concerning the first major topic, declining housing prices are associated with increasing foreclosure rates; however, negative equity need not trigger a default. Various other factors, such as a job loss, a major medical issue, financing options, one's views on fairness, or housing price expectations, can also play an important role.

Our literature review has highlighted various theoretical and econometric issues that have raised doubts about the accuracy of estimated price impacts of foreclosures and led to numerous refinements of the subsequent empirical analysis. As is standard in empirical

analyses, issues arise concerning the inclusion and exclusion of specific variables, such as those capturing housing quality and the supply effects of foreclosures. Clearly, there is potential endogeneity/simultaneity between foreclosures and sale prices, but little known effort has focused on this issue (with the exception of some research using spatial econometrics techniques, among a select few other studies).

In addition to the issue of the discount with respect to topics (ii) and (iii) above, we have motivated the issue of foreclosures and housing price studies by elaborating on the importance of considering whether or not property markets are in disequilibrium and how a researcher might control for this possibility, which was highly likely during the Great Recession. Noting that foreclosures and/or vacancies are a form of departure from equilibrium housing market conditions, Coulson and Zabel (2013) provide an excellent review of how researchers should control and have controlled for market disequilibria in the context of valuing willingness to pay for environmental quality. Some of the studies we survey have used the Coulson and Zabel (2013) prescriptions of including a dummy variable for foreclosure properties in a hedonic regression and/or including some function of distance to nearby foreclosures. Even though concern about simultaneity between house prices and foreclosures still remains, the literature has evolved in a manner that has attempted to address these issues. But there is clearly more room in the literature for simultaneity to be handled in a rigorous manner.

We have synthesized many of the estimates of the own foreclosure discount and have found this ranges between almost zero and 28 percent, with the majority of estimates greater than 12 percent. However, much remains to be learned about the fundamental determinants of this discount, especially the specifics of proxy effects. We have also compared estimates of the discount resulting from spillover effects of nearby foreclosed houses, which is much smaller than the own foreclosure discount. Specifically, the nearby foreclosure discount ranges from less than 1 percent to approximately 9 percent, with most estimates below 2 percent. This effect diminishes rapidly with distance. While nearby foreclosures are important determinants of house prices, a much more important determinant of house prices is whether a particular property is a foreclosure.

Not surprisingly, many extensions of the existing literature are possible. We have already highlighted the potential importance of handling the issues of simultaneity and market disequilibria. Two other issues are potentially very important in our view. First, the probability of a foreclosure can vary across locations in a city, and we anticipate extending the Fisher and Lambie-Hanson (2012) analysis to allow for this type of variation. Second, an examination of land value estimates, such as that of Davis and Palumbo (2008) for some major U.S. cities (including Atlanta), indicates a dramatic (and perhaps implausibly large) drop-off in land prices beginning in 2007. One might argue that foreclosures affect land or location values rather than the characteristics and value of the structure of the house. This is a topic deserving further consideration. ■

NOTES

[1] This topic has been examined by many, including Leamer (2007) and Mian, Sufi, and Trebbi (2015).

[2] While our review focuses on sales of distressed residential property, a sale is not the only outcome for a borrower in distress. For a study examining not only sales but also other possibilities, such as a loan modification, see Chan et al. (2014). According to Gerardi and Li (2010), a review of recent foreclosure-prevention programs reveals poor results in reducing foreclosures based on high rates of redefault.

[3] Analyses of the effects of different laws on mortgage outcomes in foreclosure processes have become more frequent recently. A thorough analysis of this topic is beyond the scope of this article. For examples of recent papers, see Desai, Elliehausen, and Steinbuks (2013); Fitzpatrick et al. (2014); Price et al. (2015); and Mian, Sufi, and Trebbi (2015).

[4] See Gerardi, Lambie-Hanson, and Willen (2013) and Cordell et al. (2015) for additional discussion of this issue.

[5] Although slightly dated, some of the topics in our review have been summarized previously. For example, see Lee (2008) and Frame (2010).

[6] Various house price indexes exist and these differing indexes yield slightly different results. For example, the S&P/Case-Shiller Home Price Index peaked in February 2007.

[7] While our overview is focused on the national economy, Cohen, Coughlin, and Lopez (2012) found substantial diversity across metropolitan areas during the boom and bust. During the boom, housing prices tended to rise much faster in metropolitan areas in the East and West Coast regions than in the interior. In addition, metropolitan areas with larger price booms tended to experience larger price busts. For an examination of the overall performance of metropolitan regions beginning in the early 1990s, see Arias, Gascon, and Rapach (2016).

[8] Numerous references for this topic exist. Two analyses providing many references are Foote, Gerardi, and Willen (2012) and Levitin and Wachter (2012). A recent article that studies the relationship between credit supply and house prices is Favara and Imbs (2015).

[9] Garriga (2013) and Molloy and Zarutskie (2013) discuss recent business investor activity in the housing market.

[10] Looking at the annual levels, the foreclosure rate peaked at 5.4 percent in 2009. For comparison, in the late 1990s this rate averaged less than 1.5 percent and over the past four quarters (2015:Q2–2016:Q1) was 1.5 percent.

[11] For an overview of regional variation in subprime delinquencies rates, see Doms, Furlong, and Krainer (2007).

[12] Increased homeownership enjoyed bipartisan support because it was viewed as a valuable way to build wealth and provide upward mobility. However, Bayer, Ferreira, and Ross (2016) found that minority homeowners were quite vulnerable during the bust.

[13] The questions highlighted above have generally been examined from a microeconomic perspective. However, using data aggregated to the state level, Calomiris, Longhofer, and Miles (2013) examine the direction of causality and the magnitudes of the impacts of shocks. They find causality in both directions. In addition, they find that increased foreclosures have a negative effect on housing prices, but that the negative impact of housing prices on foreclosures is much larger. Specifically, the impact of prices on foreclosures is 79 percent larger than the impact of foreclosures on prices. The relatively small impact of foreclosure starts on prices is noteworthy because research has tended to focus on this question as opposed to the impact of prices on foreclosures.

[14] Numerous studies, such as Immergluck and Smith (2006b); Goodstein and Lee (2010); Katz, Wallace, and Hedberg (2013); and Ellen, Lacoe, and Sharygin (2013), have found a connection between foreclosures and various crimes. In a recent article, Cui and Walsh (2015) find that the foreclosure itself has no effect on crime, but rather that foreclosed houses that become vacant are associated with increased violent crime. Once a house is reoccupied, the crime effects dissipate.

[15] Hedonic regressions are used commonly in housing price studies. In a hedonic housing price regression, the house is decomposed into its individual characteristics (including characteristics associated with its location) and then estimates of how each characteristic contributes to the equilibrium price of the house are generated.

[16] A key point of Coulson and Zabel (2013) is that the hedonic housing price approach, as proposed by Rosen (1974), assumes markets are in equilibrium. Therefore, hedonic housing price estimates generated for time periods where there is disequilibrium in housing markets can be biased.

[17] Formally, a put option gives the owner the right, but not the obligation, to sell an asset at a specified price by a predetermined date to a given party.

[18] Kau, Keenan, and Kim (1993) and Kau and Keenan (1999) show that even in the absence of transactions costs, borrowers will not necessarily default immediately when the option is in the money. A further contribution by Deng, Quigley, and Van Order (2000) shows that default models are best modeled in a competing-risk framework, where default competes with prepayment.

[19] An Alt-A mortgage is a type of U.S. mortgage that, for various reasons, is considered riskier than a prime mortgage and less risky than a subprime mortgage.

[20] Foreclosure rates for owner-occupied properties and nonlocal investor-owned properties were not found to be statistically significantly different.

[21] Gerardi, Herkenhoff, Ohanian, and Willen (2015) provide numerous references and some comments about the applicability of different types of models.

[22] See Mitchell, Malpezzi, and Green (2010) for a more thorough discussion of fair market value.

[23] Mitchell, Malpezzi, and Green (2010) raise the possibility that, in addition to the discount arising from a forced sale, a discount related to one's ethnicity or race might also exist, producing a "double discount."

[24] Increased time on the market might also be viewed as a stigma. Time on the market might be a signal that the house is overpriced or has a flaw that has been discovered by other potential buyers. As time passes, sellers lower their reservation price; this produces a lower price and longer time on the market. However, foreclosure status could also reduce the price and the time on the market. Listing price is related to the time a property remains unsold.

[25] When discussing own price foreclosure discounts, the corresponding sales price is generally the price received from the REO sale.

[26] Time on the market is an endogenous variable in the price equation.

[27] One form of spatial dependence is addressed by a spatial autoregressive model, where the prices of neighboring houses affect the price of the house in question. The other form of spatial dependence is spatially correlated disturbance terms, where the source is the endogenous spatially lagged variable. If the error terms associated with houses *i* and *j* are correlated, the price of house *j*, which is the lagged explanatory variable for the price of house *i*, will be correlated with the error term in the equation for the price of house *i*. Spatially correlated disturbance terms can lead to inefficient parameter estimates and, in turn, insignificant *t*-statistics. Meanwhile, if prices of neighboring houses do influence house prices, failing to use a model such as a spatial autoregressive model will result in biased parameter estimates.

[28] The references to prior estimates are by Shilling, Benjamin, and Sirmans (1990); Forgey, Rutherford, and VanBuskirk (1994); Hardin and Wolverton (1996); Springer (1996); Carroll, Clauretie, and Neill (1997); and Pennington-Cross (2006). Clauretie and Daneshvary (2009) point out a fundamental problem associated with ordinary least squares (OLS) estimation of the foreclosure effect. While many characteristics of the real estate are controlled for, the condition of the property is often not. This causes the magnitude of the discount to be overestimated because the foreclosure index is inversely related to the condition of the property. Another potential source of bias is whether the valuation of the characteristics is the same for the buyers of the two types of properties.

[29] The existence of negative externalities provides a theoretical justification for public policies and funding to mitigate these adverse effects. This topic is beyond the scope of our article.

[30] See Frame (2010) for a detailed review of mortgage foreclosure effects on surrounding property values.

[31] The authors use a technique termed "spatial contextual expansion." Simply put, variables of *lat*, *long*, lat^2, $long^2$, and *lat*long* are included in the regression. This allows the impact of the neighborhood and property characteristics to vary across space. If significant, then spatial submarkets within an area appear to exist.

[32] For many, but not all cases, "seriously delinquent" is defined as delinquent for 90 or more days. See Geraldi, Rosenblatt, Willen, and Yao (2015) for further details.

[33] Their models explicitly measure the effect on home price growth, and they find that such a property experiences a growth rate that is 1 percent less for each foreclosure or seriously delinquent property in excess of the numbers of such properties present when the home was sold the first time. Because the average appreciation from the first

to the second sale was 0 percent and most houses had neither a foreclosure nor seriously delinquent property nearby at the time of each sale, the 1 percent reduction in price growth can be roughly interpreted as a 1 percent reduction in the sales price.

REFERENCES

Anenberg, Elliot and Kung, Edward. "Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures." *American Economic Review*, August 2014, *104*(8), pp. 2527-51; http://dx.doi.org/10.1257/aer.104.8.2527.

Arias, Maria A.; Gascon, Charles S. and Rapach, David E. "Metro Business Cycles." Working Paper No. 2014-046C, Federal Reserve Bank of St. Louis, May 2016; https://research.stlouisfed.org/wp/2014/2014-046.pdf.

Bajari, Patrick; Chu, Chenghuan Sean and Park, Minjung. "An Empirical Model of Subprime Mortgage Default from 2000 to 2007." NBER Working Paper No. 14625, National Bureau of Economic Research, December 2008; http://www.nber.org/papers/w14625.pdf.

Bayer, Patrick; Ferreira, Fernando and Ross, Stephen L. "The Vulnerability of Minority Homeowners in the Housing Boom and Bust." *American Economic Journal: Economic Policy*, February 2016, *8*(1), pp. 1-27.

Been, Vicki. "External Effects of Concentrated Mortgage Foreclosures: Evidence from New York City." Testimony before the Subcommittee on Domestic Policy, Committee on Oversight and Government Reform, U.S. House of Representatives, May 21, 2008; http://www.neighborhoodindicators.org/sites/default/files/publications/testimonybeen.pdf.

Calomiris, Charles W.; Longhofer, Stanley D. and Miles, William R. "The Foreclosure-House Price Nexus: A Panel VAR Model for U.S. States, 1981-2009." *Real Estate Economics*, Winter 2013, *41*(4), pp. 709-46; http://dx.doi.org/10.1111/reec.12011.

Campbell, John Y. and Cocco, João F. "A Model of Mortgage Default." *Journal of Finance*, August 2015, *70*(4), pp. 1495-554.

Campbell, John Y.; Giglio, Stefano and Pathak, Parag. "Forced Sales and House Prices." *American Economic Review*, August 2011, *101*(5), pp. 2108-31; http://dx.doi.org/10.1257/aer.101.5.2108.

Carroll, Thomas M.; Clauretie, Terrence M. and Neill, Helen R. "Effect of Foreclosure Status on Residential Selling Price: Comment." *Journal of Real Estate Research*, 1997, *13*(1), pp. 95-102.

Chan, Sewin; Sharygin, Claudia; Been, Vicki and Haughwout, Andrew. "Pathways After Default: What Happens to Distressed Mortgage Borrowers and Their Homes?" *Journal of Real Estate Finance and Economics*, 2014, *48*(2), pp. 342-79; http://dx.doi.org/10.1007/s11146-012-9400-1.

Cheung, Ron; Cunningham, Chris and Meltzer, Rachel. "Do Homeowners Associations Mitigate or Aggravate Negative Spillovers From Neighboring Homeowner Distress?" *Journal of Housing Economics*, 2014, *24*, pp. 75-88; http://dx.doi.org/10.1016/j.jhe.2013.11.007.

Clauretie, Terrence M. and Daneshvary, Nasser. "Estimating the House Foreclosure Discount Corrected for Spatial Price Interdependence and Endogeneity of Marketing Time." *Real Estate Economics*, Spring 2009, *37*(1), pp. 43-67; http://dx.doi.org/10.1111/j.1540-6229.2009.00234.x.

Clauretie, Terrence M. and Daneshvary, Nasser. "The Optimal Choice for Lenders Facing Defaults: Short Sale, Foreclose, or REO." *Journal of Real Estate Finance and Economics*, 2011, *42*(4), pp. 504-21; http://dx.doi.org/10.1007/s11146-009-9201-3.

Cohen, Jeffrey P; Coughlin, Cletus C. and Lopez, David A. "The Boom and Bust of U.S. Housing Prices from Various Geographic Perspectives." Federal Reserve Bank of St. Louis *Review*, September/October 2012, *94*(5), pp. 341-67; https://research.stlouisfed.org/publications/review/12/09/341-368Cohen.pdf.

Cordell, Larry; Geng, Liang; Goodman, Laurie S. and Yang, Lidan. "The Cost of Foreclosure Delay." *Real Estate Economics*, Winter 2015, *43*(4), pp. 916-56; http://dx.doi.org/10.1111/1540-6229.12107.

Coulson, N. Edward and Zabel, Jeffrey E. "What Can We Learn from Hedonic Models When Housing Markets Are Dominated by Foreclosures?" *Annual Review of Resource Economics*, June 2013, *5*(1), pp. 261-79; http://dx.doi.org/10.1146/annurev-resource-091912-151759.

Cui, Lin and Walsh, Randall. "Foreclosure, Vacancy and Crime." *Journal of Urban Economics*, May 2015, *87*, pp. 72-84; http://dx.doi.org/10.1016/j.jue.2015.01.001.

Daneshvary, Nasser and Clauretie, Terrence M. "Toxic Neighbors: Foreclosures and Short-Sales Spillover Effects from the Current Housing-Market Crash." *Economic Inquiry*, January 2012, *50*(1), pp. 217-307; http://dx.doi.org/10.1111/j.1465-7295.2011.00380.x.

Davis, Morris A. and Palumbo, Michael G. "The Price of Residential Land in Large US Cities." *Journal of Urban Economics*, January 2008, *63*(1), pp. 352-84; http://dx.doi.org/10.1016/j.jue.2007.02.003.

Deng, Yongheng; Quigley, John M. and Van Order, Robert. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica*, March 2000, *68*(2), pp. 275-307; http://dx.doi.org/10.1111/1468-0262.00110.

Desai, Chintal A.; Elliehausen, Gregory and Steinbuks, Jevgenijs. "Effects of Bankruptcy Exemptions and Foreclosure Laws on Mortgage Default and Foreclosure Rates." *Journal of Real Estate Finance and Economics*, October 2013, *47*(3), pp. 391-415; http://dx.doi.org/10.1007/s11146-012-9366-z.

Doms, Mark; Furlong, Frederick and Krainer, John. "House Prices and Subprime Mortgage Delinquencies." Federal Reserve Bank of San Francisco *Economic Letter*, June 8, 2007, No. 14; http://www.frbsf.org/economic-research/publications/economic-letter/2007/june/house-prices-subprime-mortgage-delinquency/.

Ellen, Ingrid Gould; Lacoe, Johanna and Sharygin, Claudia Ayanna. "Do Foreclosures Cause Crime?" *Journal of Urban Economics*, March 2013, *74*, pp. 59-70; http://dx.doi.org/10.1016/j.jue.2012.09.003.

Favara, Giovanni and Imbs, Jean. "Credit Supply and the Price of Housing." *American Economic Review*, March 2015, *105*(3), pp. 958-92; http://dx.doi.org/10.1257/aer.20121416.

Ferreira, Fernando and Gyourko, Joseph. "A New Look at the U.S. Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012." NBER Working Paper No. 21261, National Bureau of Economic Research, June 2015; http://www.nber.org/papers/w21261.pdf.

Fisher, Lynn M. and Lambie-Hanson, Lauren. "Are Investors the Bad Guys? Tenure and Neighborhood Stability in Chelsea, Massachusetts." *Real Estate Economics*, Summer 2012, *40*(2), pp. 351-86; http://dx.doi.org/10.1111/j.1540-6229.2011.00317.x.

Fisher, Lynn M.; Lambie-Hanson, Lauren and Willen, Paul S. "The Role of Proximity in Foreclosure Externalities: Evidence from Condominiums." *American Economic Journal: Economic Policy*, February 2015, *7*(1), pp. 119-40; http://dx.doi.org/10.1257/pol.20130102.

Fitzpatrick, Thomas J.; Nelson, Lisa; Richter, Francisca G.-C., and Whitaker, Stephan. "Can Local Housing Ordinances Prevent Neighborhood Destabilization?" Working Paper No. 12-40R, Federal Reserve Bank of Cleveland, October 2014.

Forgey, Fred A; Rutherford, Ronald C. and VanBuskirk, Michael L. "Effect of Foreclosure Status on Residential Selling Price." *Journal of Real Estate Research*, 1994, *9*(3), pp. 313-18.

Foote, Christopher L.; Gerardi, Kristopher S. and Willen, Paul S. "Negative Equity and Foreclosure: Theory and Evidence." *Journal of Urban Economics*, 2008, *64*(2), pp. 234-45; http://dx.doi.org/10.1016/j.jue.2008.07.006.

Foote, Christopher L.; Gerardi, Kristopher S. and Willen, Paul S. "Why Did So Many People Make So Many Ex Post Bad Decisions? The Causes of the Foreclosure Crisis." NBER Working Paper No. 18082, National Bureau of Economic Research, May 2012; http://www.nber.org/papers/w18082.pdf.

Foster, Chester and Van Order, Robert. "An Option-Based Model of Mortgage Default." *Housing Finance Review*, 1984, *3*(4), pp. 351-72.

Foster, Chester and Van Order, Robert. "FHA Terminations: A Prelude to Rational Mortgage Pricing." *Journal of the American Real Estate and Urban Economics Association*, 1985, *13*(3), pp. 273-91; http://dx.doi.org/10.1111/1540-6229.00355.

Frame, Scott W. "Estimating the Effect of Mortgage Foreclosures on Nearby Property Values: A Critical Review of the Literature." Federal Reserve Bank of Atlanta *Economic Review*, 2010, *95*(3); https://www.frbatlanta.org/research/publications/economic-review/2010/vol95no3_effect-of-foreclosures-on-property-values.aspx.

Fuster, Andreas; Zafar, Basit and Cocci, Matthew. "Why Aren't More Renters Becoming Homeowners?" Federal Reserve Bank of New York *Liberty Street Economics* (blog), September 8, 2014; http://libertystreeteconomics.newyorkfed.org/2014/09/why-arent-more-renters-becoming-homeowners.html#.V1mk2k32Z88.

Gangel, Marshall; Seiler, Michael J. and Collins, Andrew. "Exploring the Foreclosure Contagion Effect Using Agent-Based Modeling." *Journal of Real Estate Finance and Economics*, February 2013, *46*(2), pp. 339-54; http://dx.doi.org/10.1007/s11146-011-9324-1.

Garriga, Carlos. "The Recent Boom in House Prices: Why Is this Time Different?" Federal Reserve Bank of St. Louis *Economic Synopses*, 2013, No. 32, November 22, 2013; https://research.stlouisfed.org/publications/es/13/ES_32_2013-11-22.pdf.

Gerardi, Kristopher; Herkenhoff, Kyle F.; Ohanian, Lee E. and Willen, Paul S. "Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default." NBER Working Paper No. 21630, National Bureau of Economic Research, October 2015; http://www.nber.org/papers/w21630.pdf.

Gerardi, Kristopher; Lambie-Hansen, Lauren and Willen, Paul S. "Do Borrower Rights Improve Borrower Outcomes? Evidence from the Foreclosure Process." *Journal of Urban Economics*, January 2013, *73*(1), pp. 1-17; http://dx.doi.org/10.1016/j.jue.2012.06.003.

Gerardi, Kristopher and Li, Wenli. "Mortgage Foreclosure Prevention Efforts." Federal Reserve Bank of Atlanta *Economic Review*, 2010, *95*(2); https://www.frbatlanta.org/research/publications/economic-review/2010/vol95no2_foreclosure-prevention-efforts.aspx.

Gerardi, Kristopher; Rosenblatt, Eric; Willen, Paul S. and Yao, Vincent W. "Foreclosure Externalities: New Evidence." *Journal of Urban Economics*, May 2015, *87*, pp. 42-56; http://dx.doi.org/10.1016/j.jue.2015.02.004.

Gerardi, Kristopher; Shapiro, Adam H. and Willen, Paul S. "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures." Working Paper No. 07-15, Federal Reserve Bank of Boston, May 4, 2008; https://www.bostonfed.org/economic/wp/wp2007/wp0715.pdf.

Goodstein, Ryan M. and Lee, Yan Y. "Do Foreclosures Increase Crime?" Working Paper No. 2010-05, FDIC Center for Financial Research, May 2010.

Goodstein, Ryan M.; Hanouna, Paul; Ramirez, Carlos D. and Stahel, Christof W. "Are Foreclosures Contagious?" Working Paper No. 2011-4, FDIC Center for Financial Research, February 2011; https://www.fdic.gov/bank/analytical/cfr/2011/wp2011/CFR_WP_2011_04.pdf.

Guiso, Luigi; Sapienza, Paola and Zingales, Luigi. "The Determinants of Attitudes toward Strategic Default on Mortgages." *Journal of Finance*, August 2013, *68*(4), pp. 1473-515; http://dx.doi.org/10.1111/jofi.12044.

Hardin, William G. III and Wolverton, Marvin L. "The Relationship between Foreclosure Status and Apartment Price." *Journal of Real Estate Research*, 1996, *12*(1), pp. 101-9.

Harding, John P.; Rosenblatt, Eric and Yao, Vincent W. "The Contagion Effect of Foreclosed Properties." *Journal of Urban Economics*, November 2009, *66*(3), pp. 164-78; http://dx.doi.org/10.1016/j.jue.2009.07.003.

Harding, John P.; Rosenblatt, Eric and Yao, Vincent W. "The Foreclosure Discount: Myth or Reality?" *Journal of Urban Economics*, March 2012, *71*(2), pp. 204-19; http://dx.doi.org/10.1016/j.jue.2011.09.005.

Hartley, Daniel A. "The Effect of Foreclosures on Nearby Housing Prices: Supply or Dis-Amenity?" *Regional Science and Urban Economics*, November 2014, *49*, pp. 108-17; http://dx.doi.org/10.1016/j.regsciurbeco.2014.09.001.

Immergluck, Dan and Smith, Geoff. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values." *Housing Policy Debate*, 2006a, *17*(1), pp. 57-79; http://dx.doi.org/10.1080/10511482.2006.9521561.

Immergluck, Dan and Smith, Geoff. "The Impact of Single-Family Mortgage Foreclosures on Neighborhood Crime." *Housing Studies*, November 2006b, *21*(6), pp. 851-66; http://dx.doi.org/10.1080/02673030600917743.

Katz, Charles M.; Wallace, Danielle and Hedberg, E.C. "A Longitudinal Assessment of the Impact of Foreclosure on Neighborhood Crime." *Journal of Research in Crime and Delinquency*, August 2013, *50*(3), pp. 359-89; http://dx.doi.org/10.1177/0022427811431155.

Kau, James B. and Keenan, Donald C. "Patterns of Rational Default." *Regional Science and Urban Economics*, November 1999, *29*(6), pp. 765-85; http://dx.doi.org/10.1016/S0166-0462(99)00024-1.

Kau, James B.; Keenan, Donald C. and Kim, Taewon. "Transaction Costs, Suboptimal Termination and Default Probabilities." *Journal of the American Real Estate and Urban Economics Association*, 1993, *21*(3), pp. 247-63; http://dx.doi.org/10.1111/1540-6229.00610.

Leamer, Edward E. "Housing *Is* the Business Cycle." Proceedings of the 2007 Economic Policy Symposium, "Housing, Housing Finance, and Monetary Policy," Jackson Hole, Wyoming, 2007. Kansas City, MO: Federal Reserve Bank of Kansas City, 2007, pp. 149-233; https://www.kansascityfed.org/publicat/sympos/2007/PDF/Leamer_0415.pdf.

Lee, Kai-yan. "Foreclosure's Price-Depressing Spillover Effects on Local Properties: A Literature Review." Community Affairs Discussion Paper No. 2008-01, Federal Reserve Bank of Boston, September 2008; https://www.bostonfed.org/commdev/pcadp/2008/pcadp0801.pdf.

Leonard, Tammy and Murdoch, James C. "The Neighborhood Effects of Foreclosure." *Journal of Geographical Systems*, December 2009, *11*(4), pp. 317-32; http://dx.doi.org/10.1007/s10109-009-0088-6.

Levitin, Adam J. and Wachter, Susan M. "Explaining the Housing Bubble." *Georgetown Law Journal*, 2012, *100*(4), pp. 1177-258; http://dx.doi.org/10.2139/ssrn.1669401.

Lin, Zhenguo; Rosenblatt, Eric and Yao, Vincent W. "Spillover Effects of Foreclosures on Neighborhood Property Values." *Journal of Real Estate Finance and Economics*, May 2009, *38*(4), pp. 387-407; http://dx.doi.org/10.1007/s11146-007-9093-z.

Mayer, Christopher; Pence, Karen and Sherlund, Shane M. "The Rise in Mortgage Defaults." *Journal of Economic Perspectives*, Winter 2009, *23*(1), pp. 27-50; http://dx.doi.org/10.1257/jep.23.1.27.

Mian, Atif; Amir, Sufa and Trebbi, Francesco. "Foreclosures, House Prices, and the Real Economy." *Journal of Finance*, December 2015, *70*(6), pp. 2587-633; http://dx.doi.org/10.1111/jofi.12310.

Mitchell, Thomas W.; Malpezzi, Stephen and Green, Richard K. "Forced Sale Risk: Class, Race, and the 'Double Discount.'" *Florida State University Law Review*, Spring 2010, *37*(3), pp. 589-658.

Molloy, Raven and Zarutskie, Rebecca. "Business Investor Activity in the Single-Family-Housing Market." *FEDS Notes*, December 5, 2013; https://www.federalreserve.gov/econresdata/notes/feds-notes/2013/business-investor-activity-in-the-single-family-housing-market-20131205.html.

Pennington-Cross, Anthony. "The Value of Foreclosed Property." *Journal of Real Estate Research*, April-June 2006, *28*(2), pp. 193-214.

Price, Russell; Lindsey-Taliefero, Debby; Kelly, Lynne and Brent, William. "Do Pre-Sale and Post-Sale Foreclosure Statutes Impact Foreclosure Starts?" *Journal of Financial Risk Management*, June 2015, *4*(2), pp. 82-91; http://dx.doi.org/10.4236/jfrm.2015.42008.

Rappaport, Jordan. "The Demographic Shift from Single-Family to Multifamily Housing." Federal Reserve Bank of Kansas City *Economic Review*, Fourth Quarter 2013, pp. 29-58; https://www.kansascityfed.org/publicat/econrev/pdf/13q4Rappaport.pdf.

Rauterkus, Stephanie Y.; Miller, Norman G.; Thrall, Grant I. and Sklarz, Michael A. "Foreclosure Contagion and REO versus Non-REO Sales." *International Real Estate Review*, Winter 2012, *15*(3), pp. 307-24.

Rogers, William H. and Winter, William. "The Impact of Foreclosures on Neighboring Housing Sales." *Journal of Real Estate Research*, 2009, *31*(4), pp. 455-79.

Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, January-February 1974, *82*(1), pp. 34-55; http://dx.doi.org/10.1086/260169.

Schuetz, Jenny; Been, Vicki and Ellen, Ingrid Gould. "Neighborhood Effects of Concentrated Mortgage Foreclosures." *Journal of Housing Economics*, December 2008, *17*(4), pp. 306-19; http://dx.doi.org/10.1016/j.jhe.2008.09.004.

Shilling, James D.; Benjamin, John D. and Sirmans, C.F. "Estimating Net Realizable Value for Distressed Real Estate." *Journal of Real Estate Research*, 1990, *5*(1), pp. 129-40.

Siebert, Ralph. "The Impact of Foreclosure on Housing Prices." CESifo Working Paper No. 5196, January 2015; https://www.econstor.eu/bitstream/10419/107373/1/cesifo_wp5196.pdf.

Springer, Thomas M. "Single-Family Housing Transactions: Seller Motivations, Price, and Marketing Time." *Journal of Real Estate Finance and Economics*, November 1996, *13*(3), pp. 237-54; http://dx.doi.org/10.1007/BF00217393.

Towe, Charles and Lawley, Chad. "The Contagion Effect of Neighboring Foreclosures." *American Economic Journal: Economic Policy*, May 2013, *5*(2), pp. 313-35; http://dx.doi.org/10.1257/pol.5.2.313.

Turnbull, Geoffrey K. and van der Vlist, Arno J. "Foreclosures, New Construction, and House Prices: Neighborhood Externalities or Supply Effects?" Working Paper No. 1304, University of Central Florida School of Real Estate, September 29, 2014.

Whitaker, Stephan and Fitzpatrick, Thomas J. IV. "Deconstructing Distressed-Property Spillovers: The Effects of Vacant, Tax-Delinquent, and Foreclosed Properties in Housing Submarkets." *Journal of Housing Economics*, June 2013, *22*(2), pp. 79-91; http://dx.doi.org/10.1016/j.jhe.2013.04.001.

Zabel, J. "A Dynamic Model of the Housing Market: The Role of Vacancies." *Journal of Real Estate Finance and Economics*, May 2014, online article; http://dx.doi.org/10.1007/s11146-014-9466-z.

The Visible Hand: The Role of Government in China's Long-Awaited Industrial Revolution

Yi Wen and George E. Fortier

China is undergoing its long-awaited industrial revolution. There is no shortage of commentary and opinion on this dramatic period, but few have attempted to provide a coherent, in-depth, political-economic framework that explains the fundamental mechanisms behind China's rapid industrialization. This article reviews the New Stage Theory of economic development put forth by Wen (2016a). It illuminates the critical sequence of developmental stages since the reforms enacted by Deng Xiaoping in 1978: namely, small-scale commercialized agricultural production, proto-industrialization in the countryside, a formal industrial revolution based on mass production of labor-intensive light consumer goods, a sustainable "industrial trinity" boom in energy/motive power/infrastructure, and a second industrial revolution involving the mass production of heavy industrial goods. This developmental sequence follows essentially the same pattern as Great Britain's Industrial Revolution, despite sharp differences in political and institutional conditions. One of the key conclusions exemplified by China's economic rise is that the extent of industrialization is limited by the extent of the market. One of the key strategies behind the creation and nurturing of a continually growing market in China is based on this premise: The free market is a public good that is very costly for nations to create and support. Market creation requires a powerful "mercantilist" state and the correct sequence of developmental stages; China has been successfully accomplishing its industrialization through these stages, backed by measured, targeted reforms and direct participation from its central and local governments. (JEL B00, H10, H40, H70, K00, L10, N00, O10, O20, O30, O40, O50, P00)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2016, *98*(3), pp. 189-226.
http://dx.doi.org/10.20955/r.2016.189-226

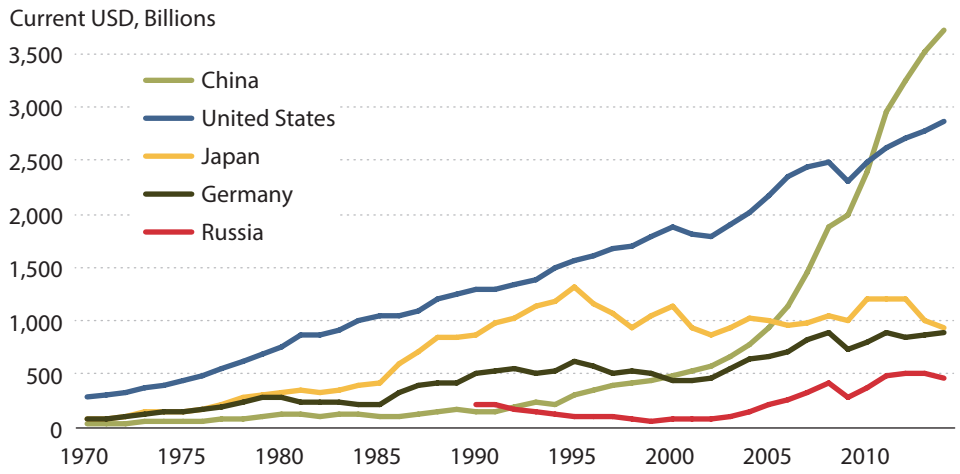
INTRODUCTION

China's economic transformation has astonished the world. Even as recently as 20 years ago, few would have predicted China's dominance as a *regional* industrial power, let alone a *global* superpower. In merely one generation's time, China has created more productive forces than have the past 5,000 years of its previous dynasties and transformed from an impoverished

Yi Wen is an assistant vice president and economist in the Research Division of the Federal Reserve Bank of St. Louis. This article is based on his working paper from the Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/wp/more/2015-006) and book (http://www.worldscientific.com/worldscibooks/10.1142/9885): *The Making of an Economic Superpower—Unlocking China's Secret of Rapid Industrialization*. George E. Fortier is the managing editor of the Federal Reserve Bank of St. Louis *Review*. The authors thank Alex Monge-Naranjo and David Wheelock for insightful and helpful comments.

Figure 1

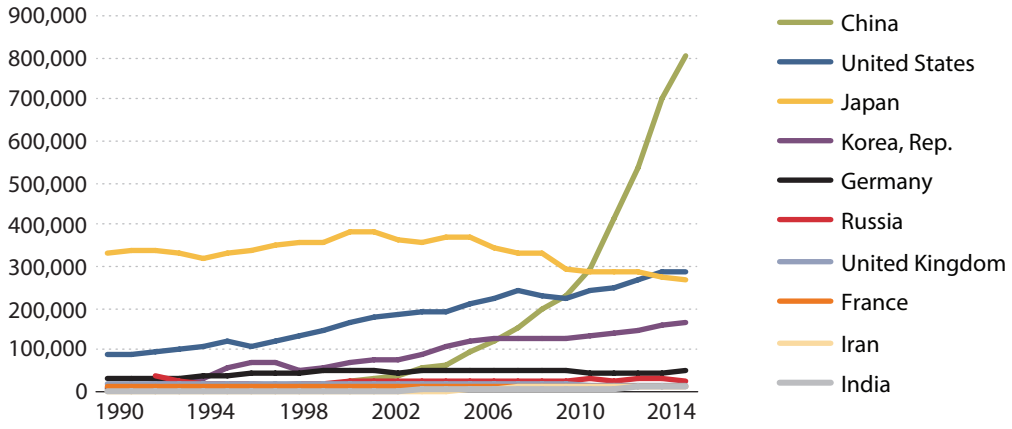
Manufacturing Output (1970-2014), Top 5 Countries in 2014



SOURCE: United Nations.

Figure 2

Patent Applications (1985-2014), Top 10 Countries in 2014



SOURCE: World Intellectual Property Organization (WIPO); see also Agence France-Presse (2014).

agrarian nation into the world's largest and most vigorous manufacturing powerhouse. (See Figure 1.)

In one year, China can produce 50 billion T-shirts (more than seven times the world's population), 10 billion pairs of shoes, 800 million metric tons of crude steel (50 percent of global supply and 800 percent of the U.S. level of production), 2.4 gigatons of cement (nearly

60 percent of world production), and close to 4 trillion metric tons of coal (burning almost as much coal as the rest of the world combined). China is the world's largest producer of passenger cars, high-speed trains, ships, tunnels, bridges, highways, machine tools, cell phones, computers, robots, air conditioners, refrigerators, washing machines, furniture, fertilizer, agricultural crops, pork, fish, eggs, cotton, copper, aluminum, books, magazines, television shows, as well as college students (see Wen, 2016a).

Moreover, China is now the world's number one industrial patent applicant. For example, China's industrial patent applications were more than the sum of those in the United States and Japan in 2014. (See Figure 2.)

How did China achieve all this in a mere 35 years, when many observers were actually betting on its collapse? Critics called attention to the Tiananmen Square incident, the collapse of the Soviet Union and Eastern European communism, the Asian financial crisis, and the 2008 global recession, which cut China's total exports persistently by more than 40 percent below trend. Yet, China persisted through its industrial revolution and has achieved an astonishing 30-fold expansion of real GDP since 1978. This transformation was unexpected not merely because of China's pervasive backwardness after centuries of turmoil and economic regress, but also because of its enduring "extractive" and authoritarian political institutions. According to the new institutional theories of economic development, the existence of these obstacles predicted nothing but dismal failure for China. For example, the celebrated book, *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, articulates this perspective (Acemoglu and Robinson, 2012).

This skepticism does have some historical support: Since 1860, all of China's previous attempts at industrialization had failed. China in the 1950s, under the government of Mao Zedong, was on the threshold of true economic growth. Its state-owned enterprises at the time were motivated by the goal of rapidly catching up with the Western industrial powers, such as Great Britain and the United States. But China attempted to achieve this goal by full-fledged industrialization through central planning based on (i) "leapfrog" developmental strategies biased toward heavy industry and (ii) industrial policies of self-reliance and self-sufficiency. Yet this newly created industrial base produced goods to meet only very thin, limited domestic demand in China. Thus, these enterprises were highly unproductive and inefficient. *Sustainable* industrialization was once again out of reach.

Two decades later, in the late 1970s, China detonated a true industrial revolution. Today, China's government and firms are guided by well-known economic principles. But these are not contemporary principles; these are the well-known yet often-ignored principles of Adam Smith (*Wealth of Nations*, 1776). Smith explained the wealth of nations by the division of labor based on the size of the market, using examples from 18th century pin factories. More specifically, it has been China's approach to creating markets that has laid the foundation for its success. Instead of taking as given the neoclassical assumption that the free market automatically exists (and would be automatically created by free individuals on the supply side in the absence of any government effort or intervention), the Chinese government has taken the initiative and expended enormous effort to create both domestic and international markets for Chinese firms. This approach is analogous to what the European monarchies and powerful

merchants (such as the English East India Company) in the 16th to 18th centuries had done since the Age of Discovery, including the colonization of the Americas.

Unlike the European nations of the 16th to 18th centuries, however, China in the late 1970s did not have a class of wealthy, savvy, entrepreneurial merchants to create markets by organizing the means of production, commerce, and transportation. The Chinese government relied, instead, on government officials with desirable leadership characteristics: These were capable, business-minded administrators who helped create local, national, and international markets for local business by supporting village firms with low taxes and cheap land, attracting outside investment, advertising local products, negotiating business deals, and building distribution networks. Such a structure could have resulted in bureaucratic stagnation. But under Deng Xiaoping's system of merit-based selection and competition, any officials who were ineffective in finding ways to bring material wealth to local populations would lose their positions under fierce intra-national competition for economic success in the villages, townships, counties, cities, and provinces. This pragmatism effectively turned all levels of Chinese government officials, through the administrative networks initially established by Mao Zedong during his 30 years of communist central planning experiments, into a highly motivated "public merchant" class. These public merchants were China's market creators.[1]

With an enormously expanded and deepened market, China eventually set off its long-awaited industrial revolution. Indeed, China's modern firms, regardless of their ownership type, operate according to the Smithian market-size principle to compete and meet the demand of well-developed and well-enriched domestic and international markets. Many of China's modern firms, while state-owned, have been highly productive, competitive, and profitable because they have the mass market to support their large-scale mass operations; comparable Chinese firms in the 1960s were highly unprofitable because they had no such markets or market mechanisms.

The objective of this article is to provide a brief summary and road map for Wen's (2016a) New Stage Theory (NST) of market creation and economic development, drawn from China's growth experiences and the economic history of the West.[2] We will describe the stages (and the sequence of those stages) of market creation and identify the sometimes easily overlooked steps that China's government took to successfully generate a full-fledged industrial revolution after 1978. This last point is critical: The Chinese central and local governments relied on China's state banking system and public land ownership to help create one of the largest unified manufactured goods markets in world economic history; this market nurtured, supported, and incentivized firm entry and industrial upgrading through the demand-side-driven adoption and market-oriented invention of modern manufacturing technologies and industrial organizational changes. Only toward the end of its second industrial revolution (which featured mass production of heavy industrial goods) did China begin to seriously engage in creating a financial market, pushing for the internationalization of its currency, and establishing market-based financial regulatory institutions to manage financial capital flows. China's prudent sequence of market creation explains the absence of any recurrent and destructive financial crises that have dominated the developmental history of the West and Latin America. China's sequential and "engineered" market-creation process thus offers a new model of

economic development for developing countries. Behind the core economic validity of this market-creation strategy is an equally compelling reality: Political stability and social trust are the most fundamental pillars of the "free" market; forces that undermine stability and trust (such as premature and radical top-down political-economic reforms) can undermine the market itself. And it is worth noting here the distinction between the concept of an absolutely free market and the actual "free" markets in modern China, which are vibrant but directed by the government. This direction, or intervention, as noted throughout this article, is based on balancing the creative powers and the destructive powers of market forces.

CHINA'S FAILED ATTEMPTS AT INDUSTRIALIZATION

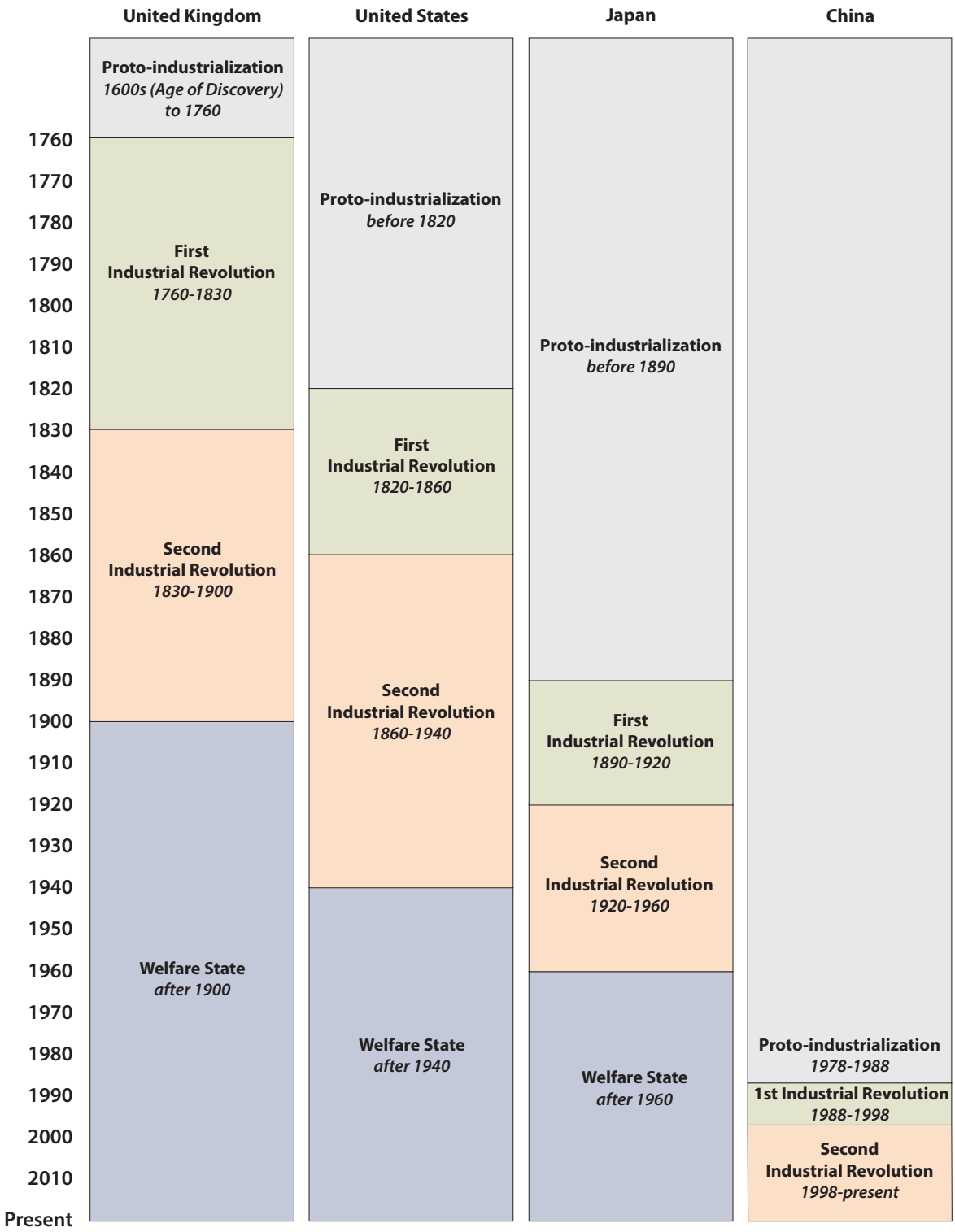
The thousands of years of China's history include technological innovations, cultural advances, and global voyages that have preempted or surpassed those of many nations. Yet, in the middle of the 20th century, China remained one of the poorest nations on earth, with one of the lowest standards of living and life expectancy and a per capita income just one-third of the average sub-Saharan African country.[3]

Of course, China did try to instigate economic growth—as well as increase military power to protect its national interests and encourage national pride, among other efforts. The economic reform in 1978 was certainly not China's first state-led attempt to industrialize. In fact, it was the *fourth* attempt since the Second Opium War.

After China was defeated by the British in the Second Opium War in 1860, the late Qing monarchy attempted to modernize its agrarian economy with the establishment of, among other things, a modern navy and an industrial infrastructure. The effort was a gigantic failure. The event that crystalized that failure was China's defeat in 1895 at the hands of the Japanese in the First Sino-Japanese War. As with earlier conflicts against the British, the war was a lopsided defeat. Despite China's hopes for true industrialization, even semi-industrialized Japan severely outmatched an underdeveloped China. Half a century passed, and by 1910 the nation was in turmoil, the Chinese government was deep in debt, and the hoped-for industrial base was nowhere in sight.

The Qing government's repeated failure to defend China against foreign aggression triggered demand for political reform. Social unrest ultimately led to the 1911 Xinhai Revolution that overthrew the Qing monarchy and established the Republic of China, the first "inclusive" government in Chinese history. This new republican government, based on a Western-style constitution, also tried to industrialize China by mimicking U.S. political institutions such as democracy and the separation of powers (that is, the checks and balances of the legislative, executive, and judicial branches of government). The Chinese people at that time adopted the slogans "Of the people, by the people, and for the people" and "Only science and democracy can save China." The educated elite revolutionaries believed that the Qing monarchy's failure to industrialize and China's overall backwardness was due to its lack of democracy, political inclusiveness, and pluralism—exactly as modern institutional theory has argued (again, see Acemoglu and Robinson, 2012). The political leaders of the republic established an inclusive form of government based on several premises: open access to political power (by including

A Macro View of Industrial Revolution



SOURCE: Estévez-Abe (2008) and Wen (2016a,b) and the references therein.

Proto-industrialization

This stage involves the rural production of the most basic goods.

In the United Kingdom, production was coordinated and financed by a class of wealthy merchants under the "putting out" system (see endnote 13). Japan gained from the political stability of the Edo period (1603-1868) and continued its proto-industrialization in the early Meiji period (1868-1890). In China, this and other stages made use of collectively (not privately) owned enterprises.

First Industrial Revolution

This stage involves the mass production of textiles, through the use of rudimentary systems such as wood-framed and water-powered machinery, as well as imported technologies (notably in Japan).

In the United States, mass production of textiles was driven by water power, especially along New England's fast-moving rivers, such as the first cotton-spinning mill (Blackstone River, Pawtucket, RI) founded by Samuel Slater. China made progress in this stage in much the same way as the previous three nations did, becoming the largest producer and exporter of, among others, textiles, cotton, furniture, and toys.

Second Industrial Revolution

This stage is divided into two components: (i) a boom in the industrial trinity of energy, transportation, and locomotive power and (ii) the mass production of the means of mass production.

In the United Kingdom (as well as in nations to follow), coal was a major source of power used to produce, e.g., iron, steel, and chemicals; advances in transportation included "macadam" roads, railroads, and canals and the steam engine. The United States had a railroad surge from the 1820s through the 1870s and relied on steam power after the Civil War; the production of the means of mass production included automobiles, telecommunications, and mechanized agricultural systems.

Japan's second industrial revolution progressed into the Second World War and continued under Allied occupation (1945-1952). In China, key developments have included a major surge in the extraction and consumption of coal and production of steel, cement, machine tools, and infrastructure (e.g., 70,000 miles of highway, almost 50 percent more than the U.S. total).

Welfare State

This stage involves an expansion of political rights and social services.

Examples in the United Kingdom include the National Insurance Act in 1911 and universal suffrage in 1928. In the United States, examples include the civil rights movement in the 1960s and the Violence Against Women Act in 1994. Japan enacted a minimum wage in 1959 and their modern, current national health insurance system in 1961. China has not yet completed its second industrial revolution and has not yet entered the welfare state.

even the Communist Party in the government),[4] modern corporations and financial institutions, new private property laws, and new versions of public universities. These reforms encouraged free trade, welcomed foreign capital, and fully embraced the bourgeois ethic throughout China, especially in large commercial cities such as Shanghai. But 40 years passed and, in 1949, China remained one of the poorest nations on earth in terms of average living standard and life expectancy and literacy.

This second failed attempt at industrialization in China was illuminated, again, by Japan's almost effortless yet ruthless invasion and occupation of China in the late 1930s and early 1940s, including the Nanking Massacre. The republic government's ineffectiveness in solving the problem of China's poverty also made it vulnerable to revolt; the communist peasant army defeated the regime in 1949 with the support of millions of impoverished peasants. Mao declared that "the Chinese people have finally stood up!" and initiated a third ambitious attempt to industrialize China—this time by mimicking the Soviet Union's social planning model instead of the West's capitalism and democracy. Thirty years passed and this third attempt at industrialization failed again: In 1978, China remained essentially in the same "Malthusian" poverty trap with per capital income no different from what it was around the Second Opium War.

To be fair, each of these failed attempts made some progress, but not enough to set off an industrial revolution. For example, Mao's regime managed to establish a basic (though highly unprofitable) industrial base and national defense system, which relied heavily on government subsidies through heavy taxation on agriculture. Agricultural productivity did improve, with the exception of the "Great Leap Forward" period. Life expectancy increased from about 35 years in 1952 to 68 years in 1982, and infant mortality fell from about 300 deaths for every 1,000 live births in 1952 to 31 deaths in 1999; rates of infection and disease, such as malaria, as well as deaths from floods and drought, also fell precipitously (Cook and Dummer, 2004; Blumenthal and Hsiao, 2005). In addition, China's literacy rate reached 66 percent in the 1960s. However, these improvements immediately translated into an expanded population—from 600 million in 1950 to 1 billion in the late 1970s, leaving income per capita barely changed from 1949, when the communist regime assumed power. Despite their potential for success, these changes did not provide food security or an escape from the Malthusian poverty trap. But they did lead directly to Deng Xiaoping's successful economic reforms in 1978.[5]

DENG XIAOPING'S PRAGMATIC APPROACH TO ECONOMIC SUCCESS

> *"It does not matter if the cat is black or white as long as it catches the rat."*
> **—Deng Xiaoping**

Deng Xiaoping worked alongside Mao Zedong through the political, social, and economic strife from the 1950s through the 1970s, including the Great Leap Forward. Deng developed a reputation for inner strength and philosophical flexibility, and Mao once described him as "a needle inside a cotton ball" (Vogel, 2013, p. 26). Although Deng was a high-ranking official in the Communist Party of China, he did not always agree with the Party's rigid, ideological

approaches to economic reform (and in fact was purged twice from the Party as a result of those disagreements).[6] When Deng assumed power in the late 1970s, he had much cautionary evidence to consider: three major failures at industrializing China over a period of 120 years, spanning three different political regimes.

> Some officials were bold enough to suggest that the real cause of the problems China was facing was Mao Zedong himself, but Deng believed that a single person should not be held responsible for the failures of the previous two decades… [I]n Deng's view the larger problem was the faulty system that had given rise to those mistakes. The effort to gain control of the political system down to the household had overreached, creating fear and lack of initiative. The effort to gain control of the economic system had also overreached, causing rigidities that stymied dynamism. How could China's leaders loosen things up while keeping the country stable? (Vogel, 2013, pp. 21-22)

Ultimately, Deng Xiaoping's pragmatic and patient reforms led to a sustainable industrial revolution in China and paved the way for continued economic growth.

After observing major political upheavals and miscalculations based on ideology, Deng established fundamental principles behind his approach to economic reform: No socialist economy can achieve sustainable growth without market elements; but no market economy can flourish and continue to prosper without state-led industrial policy, social order, and political stability, which, in China's case, would be established by a strong state government. Deng's reforms were enacted, modified, and supported by high-ranking central government officials and implemented by local-level officials in the counties and countryside. China would "cross the river by touching the stones" and "seek the truth from its own practice": That is, China would not adhere to a strictly dogmatic approach, but rather would embrace persistent pragmatism, step by step.

Ultimately, China's development since those reforms has been very much outside any typical or traditional strategy enacted to promote economic growth, such as those suggested by the "Washington consensus" and "shock therapy."[7] Rather, industrialization in China emerged from a more pragmatic process of trial and error according to the sociopolitical conditions in China at the time. Chinese government officials might have embraced existing economic theories and conventional advice to guide them (as did governments in Africa, Latin America, Russia, and Eastern Europe in the 1980s and 1990s), but Deng Xiaoping's government would forgo that advice and forge its own path. To be sure, the path to development after 1978 was a bumpy one and the Chinese government made many mistakes; fortunately, none of them has been ruinous, although some did inflict unnecessary pain on the Chinese people. But in its process of trial and error, the Chinese government under Deng also made many correct decisions that turned out to be critical for setting off China's truly long-awaited industrial revolution.

Key Steps

China's industrialization can be characterized by the following key steps:

(i) Solve the food security problem. The Chinese government established basic food security through a primitive agricultural revolution based on small-scale farming

China's Village Firms

In China, among other countries, industrialization began with the village firm. In their most basic form, these rural small-scale operations allowed farmers to manufacture simple goods outside of the growing season, to supplement income. Village firms existed during China's Great Leap Forward (1958-1962), when millions of farmers were relocated to unproductive centrally planned firms that tried to meet only local demand. Failure and famine followed. After 1978, the collectively owned village firms, although managed by government officials, were free to choose what to produce based on market demand. They were also guided toward long-distance trade and international exports with supportive commercial and credit policies. These were Deng Xiaoping's innovations. And, because the market expanded, the number and size of village firms also expanded and commerce flourished.

The output of village firms grew, on average, 28 percent per year from 1978 to 2000, doubling every 3 years. Adjusted for inflation, growth was still 21 percent per year (twice as fast as China's real GDP growth), doubling every 3.7 years.

Characteristics
Rural
Led by local officials (often democratically elected) serving as entrepreneurs/merchants/CEOs
Collective ownership but also private decisionmaking responsive to market demand
Residual claims and profit sharing
Institutional constraints but also market incentives, including competition and profit opportunities
Governmental support in securing credit and commercial information, conducting negotiations, coordinating supply chains, smoothing inventories, etc.

Progress, 1978-88

	1978	1988	Increase
Industrial gross output	51.5 billion yuan (14% of GDP)	702 billion yuan (46% of GDP)	13.5-fold
Number of village firms	1.5 million	18.9 million	12.5-fold
Workers' aggregate wage income	8.7 billion yuan	96.3 billion yuan	12-fold
Total capital stock	23 billion yuan	210 billion yuan	9-fold
Number of workers	28 million	95 million	3-fold
Workers as fraction of total rural labor force	9 percent	23 percent	2.5-fold

SOURCE: Wen (2016a) and the references therein. For the original data, see Zhang and Zhang (2001 [in Chinese], Appendix Table 1).

and collective ownership of land. Government officials encouraged commercialized farming and supplementary (sideline) work to generate additional income for farmers.

(ii) Start a primitive rural industrialization based on township-village enterprises. This stage was critical because it channeled local surplus labor in rural areas into simple industries; this process would ferment the mass market needed to support mass production that would emerge from China's forthcoming industrial revolution.

(iii) Initiate a true industrial revolution of mass production of light consumer and industrial goods based on obsolete or imported technologies, with a well-fermented domestic market from the previous stage of rural industrialization as well as an international market for these goods.

(iv) Engineer a boom in the "industrial trinity" of energy, motive power, and infrastructure (especially transportation) based on the savings accumulated from the rural industrialization and the first industrial revolution. This boom in the industrial trinity naturally initiates a second industrial revolution, featuring the mass production of the *means* of mass production and mass distribution: These means (goods or tools) include steel, cement, and other intermediate goods used in buildings, highways, and railroads and the machinery used in light industries. The industrial trinity is the flagship industry during the initial phase of a second industrial revolution and a linchpin between the first industrial revolution and the second industrial revolution. This initial phase of the second industrial revolution becomes feasible, affordable, and profitable because of the thick market (enormous market demand) for energy, motive power, and infrastructure created through the earlier development stages, especially the first industrial revolution stage. Also, a later phase of the second industrial revolution (featuring mainly the mass production of various types of machine tools) naturally follows because an industrial trinity boom broadens and deepens the market for heavy industrial goods and machinery created through the earlier stages and especially through the industrial trinity boom itself.

China is currently engaged in its second industrial revolution. Once China finishes that second industrial revolution, the entire system of industrial production will be complete, forming a positive feedback loop such that all essential goods and commodities can be mass produced, including the means of mass production and mass distribution itself. This system is also flexible enough to respond to changes in consumer demand, unlike the rigid centrally planned system set up by the Soviet Union during the Cold War era. In the economic development of the West, large financial markets were created in stage iv (as previously described) mainly to facilitate this heavy industrial revolution (e.g., during the age of steel and railways), which required large sums of both public and private debt and credit. China has only recently begun to seriously engage in creating a large financial market, now that it is in the middle of finishing its second industrial revolution. Once this process is complete and China has a modern financial market, it will be ready to enter the next stage, a welfare state, which Western economies have enjoyed since the middle of the 20th century—or at least the end of World

War II—after finishing their own second industrial revolutions. This welfare state includes economic welfare, such as a social safety net, and political welfare, such as universal suffrage.

Therefore, the key to successful economic development and avoidance of the so-called "middle income trap" is to eventually finish a second industrial revolution, to successfully forge the industrial feedback loop to make mass production and technology creation self-sustainable. However, the correct steps to achieve this do not lead through heavy industrial buildup and financial liberalization in the early stages of development, but instead follow a continuous, specifically sequential creation of markets to nurture and stimulate industrial upgrading *over time*. Attempting to leapfrog by skipping the earlier, primitive developmental stages and entering the welfare stage prematurely can lead to development disorders, debt crises, and political chaos.

APPLYING THE NEW STAGE THEORY OF DEVELOPMENT TO EACH STEP OF CHINA'S INDUSTRIALIZATION

China's path to industrialization, as it turned out, mimics the sequence of the original Industrial Revolution, which occurred in Great Britain from the 18th to the late 19th centuries. In fact, Wen (2016a) argues that almost all successfully industrialized nations—such as northwestern European nations, the United States, and Japan—have followed a similar *bottom-up* and sequential approach to industrialization despite dramatic differences in their political systems. Many nations (including China in its earlier three failed attempts) have failed to kick-start their industrial revolution or have gotten stuck in the middle of their industrialization process because they have taken a *top-down* approach by skipping important earlier stages in the sequence of the original industrial revolution. With this top-down approach, the governments in developing nations build up advanced industries and systems during the very early stages of industrialization: capital-intensive industries such as those for chemicals, steel, and automobiles; modern financial systems such as a floating exchange rate, free international capital flows, and fully fledged privatization of state-owned properties and natural resources; and modern political institutions such as democracy and universal suffrage.

This top-down approach fails for a simple reason. A mass market is required to make mass production profitable and sustainable. Yet, it is extremely costly to create a mass market, especially one for heavy industrial goods, because the mass market requires not only political stability and enormous social trust but also a system of mass distribution. Developing countries simply do not have such a mass-distribution system and the purchasing power to support the mass production of heavy industrial goods. Again, a direct consequence of such top-down approaches is political instability and unbearable financial burdens.

Hence, an industrial revolution requires the correct procedure and the correct sequence of steps to create mass markets to support mass production. Traditional development strategies such as import substitution industrialization (ISI), the Gerschenkronian (1962) heavy-industry-biased "big push," "shock therapy," and the structural adjustment program based on the Washington consensus have failed precisely because they have all ignored certain key ideas: (i) the Smithian principle that the division of labor is limited by the extent of the market,

(ii) the enormous social costs of market creation (which requires powerful state capacity and a strong-willed "mercantilist" government), and (iii) the correct sequence of creating the mass market under correct industrial policies, as illustrated below.[8]

Step 1: China's Primitive Agricultural Revolution

Throughout history, many nations have experienced pervasive market failures in agriculture, despite highly secured private land ownership and institutions that protect alienable land contracts. Even the governments of the Qing Dynasty and the Republic of China protected private property and contracts. But, as development economist Joe Studwell (2013) notes, in such a market-failure equilibrium, although land is privately owned with alienable contracts, powerful Darwinian forces eventually concentrate land in the hands of a few landlords and the majority of the population become tenants.[9] The population grows, land becomes increasingly scarce over time, and landlords can then easily lease out plots at higher and higher rents. Landlords also act as money lenders and are able to impose high interest rates (usury). Within such an equilibrium, tenants have no incentives to make the investments to improve land productivity (e.g., through fertilizers or irrigation systems) because they have little security in maintaining access to that land and must face stiff rents and carry expensive debts. Landlords also have no incentives to invest in fertilizer and irrigation systems because they profit easily from merely collecting rent and lending at high rates. When tenant debts are not paid, landlords reclaim the plots of land along with collateral and then lease them out to other tenants. Nations in such an equilibrium of low agricultural productivity cannot withstand even minor natural shocks, such as drought or flood, and thus are constantly in a state of chronic famine. Évariste Régis Huc (1813-60), a French missionary Catholic priest who traveled through China from 1839 to 1851, bears witness to such conditions in his book *A Journey through the Chinese Empire*:

> [U]nquestionably there can be found in no other country such a depth of disastrous poverty as in the Celestial Empire. Not a year passes in which a terrific number of persons do not p[e]rish of famine in some part or other of China; and the multitude of those who live merely from day to day is incalculable. Let a drought, a flood, or any accident whatever occur to injure the harvest in a single province, and two thirds of the population are immediately reduced to starvation. You see them forming up into numerous bands—perfect armies of beggars—and proceeding together, men, women, and children, to seek some little nourishment in the towns and villages…Many faint by the wayside and die before they can reach the place where they had hoped to find help… (Quoted by Landes, *The Wealth and Poverty of Nations*, 1999, p. 346)

The 1911 Xinhai Revolution, as profound as it may be in modern Chinese history in moving toward a democratic political system, did not change China's miserable rural landscape. The revolution introduced pluralist political structures and inherited private land ownership from the Qing Dynasty. R.H. Tawney, the British economic historian who visited China in the late 1920s (10 years after the Xinhai Revolution and more than 70 years after Évariste Régis Huc), wrote about the devastating situation of Chinese peasant farmers: "There are districts in which the position of the rural population is that of a man standing permanently up to the neck in water, so that even a ripple is sufficient to drown him… in Shanxi province

at the beginning of 1931, three million persons had died of hunger in the last few years, and the misery had been such that 400,000 women and children had changed hands by sale" (in Studwell, 2013, p. 17).

The American sociologist and writer William Hinton, who conducted research in China's Shanxi province in the 1940s, also wrote about "the mundane realities of death by starvation during the annual 'spring hunger' when food reserves ran out, and of the slavery (mostly of girls), landlord violence, domestic violence, usury, endemic mafia-style secret societies and other assorted brutalities that characterized everyday life" (in Studwell, 2013, p. 18).

As dramatic as this may sound, many pre-industrial agrarian societies face these hardships. In China, they served as the socioeconomic foundation for the rise of communism and radical land reform led by Mao Zedong's communist party. Ironically, after the nationalist government was defeated by the communist army and fled to Taiwan, they conducted essentially the same type of land reform as the communists did in the mainland by taking the available land from landlords and dividing it up and distributing it equally among the farming population. Such a land reform triggered Taiwan's economic takeoff and industrial revolution.

However, Mao Zedong's plan to boost agricultural productivity after land reform by reorganizing individual farming units into large collectives was a dramatic failure. Agricultural production (with both traditional and modern techniques) requires special attention and is not easily or quickly converted to a system that might function well for other industries.

Historically, individual family farms have been fairly self-contained and have required few contributions from individuals outside the family. But during the Great Leap Forward, each farming collective assembled hundreds or even thousands of farmers within a militaristic organizational structure. In agriculture, the rate of return gained from this sort of division of tasks, specialization, and coordination among a large labor force is low and extremely limited—unlike the pin factory visited by Adam Smith, or the labor-intensive mass-production textile factories in late 19th century England, or the Ford automobile assembly lines in early 20th century United States. Growing crops is governed entirely by the natural biological cycle of plants, cannot be arbitrarily divided into many intermediate stages or intermediate goods, and is land intensive and nature sensitive. Hence, it is subject to rapidly diminishing returns from an increased supply of labor or a large-scale organization of labor. Moreover, because of the natural lack of complementarity among individual farmers' efforts in agricultural production, free-rider moral-hazard problems can easily arise in large organized farms that are based on teamwork.

Even in the development history of Western industrial countries, agriculture has always been the last sector to be industrialized or to achieve the economies of scale associated with the use of heavy machinery. For example, fully fledged mechanized farming did not take place in the United States until the 1940s, compared with the mechanization of the textile industry, which took place in the mid-1800s.

This premise holds true across different forms of ownership and property rights associated with land and the farm. Although a free market system would have likely avoided the Great Leap Forward and the malfunctioning of the communes within Mao's centrally planned collective farming, by no means does that imply that a free market would have automatically

solved China's food security problems and set off China's agricultural revolution and then its industrial revolution. The Qing Dynasty's free markets did not bring about these events. And neither did the Republic era's free markets under political democracy. So why would they do so in the 1950s or the 1980s? Ultimately, agriculture would flourish in China. But China would first need to take a few critical steps for that to happen.

Deng's 1978 reform started with tearing down the large farming collectives and reverting to the traditional family-based farming units. This turned out to be the correct step to raise agricultural productivity. But this change meant returning to the simpler modes of production that were in use before communism, as in the Qing Dynasty (before 1911) and the Republic era (1912-1949). But this system of agricultural production was not sufficient for and did not lead to agricultural self-sufficiency and food security. As noted, private property rights were not the easy answer: The Qing Dynasty and the Republic era had those rights but did not make any significant progress in agricultural stability and food security.

For Deng's policies to be successful, they had to address other fundamental obstacles: (i) Peasant farmers lacked residual claim rights in the so-called "market determined" contracts that existed in Chinese history between land owners and those farmers. In other words, the farmers would not share in any profits and thus had no incentive to innovate or increase production beyond the bare minimum expected of them. (ii) There was no network of village-level irrigation systems or public roads connecting the villages and townships. Isolated farmers risked starvation—in the face of droughts, floods, or other natural disasters—if they moved from subsistence farming to some kind of agricultural specialization or commercialized farming aimed at long-distance trade, where they would need to sell their surplus of specific goods to buy other, necessary goods. (iii) The market was not large or stable enough to support commercialized farming and agricultural product specialization. And (iv) no rural industrialization had taken place to absorb the surplus labor in the countryside and dramatically raise farmers' productivity through manufacturing.

Mao's government had actually solved two of these problems during the Great Leap Forward and afterward (mostly during the government-engineered "rural corporative movement"): It had built up the infrastructure of public irrigation systems and local roads to accommodate the collective farms. This infrastructure provided the necessary improvements in technology and transportation during Deng's agricultural reform era to help increase the productivity of family-based agricultural efforts and the profitability in agricultural trade. Another easily overlooked but critical accomplishment is the establishment of rural factories (eventually known as township-village enterprises after being called communal factories during Mao's era) based on the cooperative spirit that Mao helped to create among Chinese peasant farmers in the vast rural areas. This "social capital"—a pillar of the "free" market—turned out to be a crucial factor for detonating China's rural industrialization after Deng's 1978 economic reform.[10]

It is worth considering Adam Smith's perspective on Scotland's economic environment in the 18th century:

> In the lone houses and very small villages which are scattered about in so desert a country as the highlands of Scotland, every farmer must be butcher, baker, and brewer, for his own

> family. In such situations we can scarce expect to find even a smith, a carpenter, or a mason, within less than twenty miles of another of the same trade. The scattered families that live at eight or ten miles distance from the nearest of them, must learn to perform themselves a great number of little pieces of work, for which, in more populous countries, they would call in the assistance of those workmen…. There could be little or no commerce of any kind between the distant parts of the world. What goods could bear the expense of land-carriage between London and Calcutta? Or if there were any so precious as to be able to support this expense, with what safety could they be transported through the territories of so many barbarous nations? (*The Wealth of Nations*, Chapter III)

Mao's public land ownership and rural corporative movement transformed similarly isolated Chinese farms and villages, challenged by the same limited options as those described by Adam Smith, into rural corporations. These corporations were not profitable because of the limitations of the market. It was Deng's series of reforms that extended a unified national market for rural corporations leveraging those rural infrastructures built by Mao's government and the social trust among the farmers nurtured by the movement and ethic (including the Great Leap Forward) of the commune.

Deng's reforms also provided solutions to other obstacles that Mao's communist regime could not solve: Farmers were given the incentives to work harder than before because their compensation was now linked to their individual efforts—again, despite public ownership of the land. Farmers were given a 15- or 30-year lease to work the land and the freedom to decide what crops to grow (based on market demand and other factors) and when and how long to work. The productivity of land varies greatly, depending on the type of soil and the crops planted. Deng's system of public ownership with private decisionmaking allowed farmers the independence to choose how to maximize their output, for example, by diversifying their crops, targeting crops most suitable for the soil, and being able to respond to market demand.

More importantly, under Deng's new incentive mechanism, farmers became the residual claimers on the output they produced, after meeting government quotas. Hence, farmers worked harder and longer hours and could fully use evenings and seasonal leisure time as they desired. All these factors combined to form an environment that produced an unprecedented boom in agricultural productivity in China soon after 1978.

As a result of this primitive agricultural revolution, China's aggregate agricultural output increased significantly and steadily. For example, crop output rose permanently by more than 20 percent in 1980 alone (see Wen, 2016a). Now, this permanent increase in agricultural output could have been used as surplus, as has been the case throughout history, to support additional children in the family, which in China would have amounted to millions of new babies. But China's Malthusian trap did not continue and the multitude of additional children did not come. The main reason is China's infamous one-child policy implemented in 1979 by the central government. Another reason is that a different sort of revolution—a rural industrialization—was taking place. This early, limited industrialization began to offer an increasing variety of goods to consume. And suddenly the populace had a choice: work harder and use the surplus to consume more goods or use it to raise children. Essentially, they chose the former.

Step 2: China's Proto-industrialization

A well-known and well-documented phenomenon soon after the 1978 reform was the emergence of the so-called township-village enterprises across China's vast countryside. These collectively owned enterprises flourished because (i) farmers were eager to find new sources of income beyond what their subsistence-level farming offered and (ii) local village and township governments also wanted (and in fact were required by the central government) to rapidly develop their local economies to improve the conditions of the local population. But although the existence of these township-village enterprises is well studied, their relation to Western economic history and the Industrial Revolution is not. These enterprises were, in fact, the key to triggering China's industrial revolution. And, although these enterprises may seem highly specific to China alone, this form of rural industrialization is in fact reminiscent of the primitive (proto) industrialization that flourished in 17th and 18th century Great Britain over the two hundred years or so before the British Industrial Revolution.[11]

Affluence in industrial societies often means the ability to mass-produce manufactured goods, such as processed food and textiles and shelter and means of transportation. In an agrarian society, before machinery and other forms of capital are invented or used in production and can be mass-produced, labor is the most important and perhaps the only means of producing manufactured goods. (Of course, human labor produces these goods with the help of primitive tools.) But the majority of the members of the labor force do not engage in manufacturing; instead, they reside in the countryside and devote themselves to agricultural production to maintain food security.

Between the 17th century and the middle of the 18th century in Great Britain, more and more peasant families were engaging in small-scale manufacturing and choosing to specialize in textiles and other products as the market deepened. More and more rural households were transformed into commerce-based proto-industries involving specialization and long-distance trade. Over a century and a half, the market fermented and organizational developments took hold; these part-time manufacturing workers and village firms eventually transformed into full-time workers and large-scale factories. Mass production became the critical means for merchants and other capitalists to compete for domestic and international market shares.

Such a proto-industrialization was necessary for detonating the British Industrial Revolution because mass-production-based industrialization requires a large and deep market to make the division of labor and large-scale cooperative efforts profitable. Industrialization also relies on sufficiently high incomes (wages) and the purchasing power of the grassroots population, which in turn requires that a large pool of the autarkic peasants move away from subsistence farming and engage in cooperative manufacturing within a framework of industrial organization. And all this must be accomplished without jeopardizing food security. In addition, factories are erected on land, and both land and labor are cheaper and more readily available in rural areas than in urban areas. For one thing, providing board and housing for peasant workers in the cities would be extremely costly.[12] Hence, using local land and local surplus labor in the rural areas to produce manufactured goods for long-distance trade is the most economical way of starting a proto-industrialization, regardless of who owns the property, as long as (i) the right to make decisions resides in the firms and (ii) business-failure

risks are borne by these rural industries. It is precisely this proto-industrialization that is needed to create the mass market and mass distribution networks to support a genuine first industrial revolution based on labor-intensive mass production.

There is one fundamental difference between the origination of China's proto-industrialization and the origination of Great Britain's proto-industrialization: In Great Britain (and other parts of Europe), it was mainly the merchants that took the initiative to finance and organize the village industries: They engaged and recruited the peasants to work cooperatively; they coordinated the production systems and cartage in the manufacturing of light consumer goods (mostly textiles); they took charge of long-distance trade and sales and provided the needed trade credit and raw materials for continuous production (e.g., from the emergence of the rudimentary "putting-out" system of local production all the way to the emergence of large factories in rural areas).[13] So, in many parts of Europe, the catalysts (or "economic enzymes," if you will) of market creation and rural industrial organization were the merchants, not the owners of the production factors (labor, land, and tools).[14] What motivated and financed the merchant class's proto-industrial activities through the putting-out system was the enormous world market and profit opportunities created by the mercantilist governments of Europe in general and Great Britain specifically and their trade policies based on colonialism, imperialism, and the slave trade.[15]

In China, however, that entrepreneurial role of domestic and global market creation and rural production-organization was played essentially by the central government acting in concert with the local village-level and township-level governments. These government officials were key agents in facilitating the organization of the factors and methods of production.

As Adam Smith observed, in a primitive agrarian society, the family is the basic unit of production and exchange. The family members produce everything they need. They *might* have the incentive to specialize and produce more than what is needed, through the division of labor, and sell that surplus to other families. But, because of the lack of an organized market, it is risky to specialize in producing one type of household good and to depend on other sources for other necessary goods. Clearly, food security is the highest priority, and the lack of any "insurance" for failed sales in the market is daunting. Yet the division of labor and separation of demand and supply through social specialization is the key to improving labor productivity. Even the most primitive form of rural factories requires peasants from different families to be organized into teams (essentially, a corporation) to engage in coordinated production and to share the profits and business risks. Such an organization requires initial capital (more than a hundred or thousand times a farmer's annual family income) as well as fundamental trust among the workers and the organizers. Moreover, success depends critically on efficient long-distance distribution channels to ensure sales and supply of raw materials (which may not be available locally, such as cotton and wool).

As noted earlier, a new and powerful merchant class emerged in the 16th to 18th centuries in Europe; they were backed by mercantilist and militarized state support and motivated by enormous monopoly profits from global trade (based on colonialism, imperialism, and the slave trade).[16] These merchants created markets and served as the organizers of proto-industries, which paved the way for the British Industrial Revolution—which hinged on

Banking and Finance

Throughout China's industrial revolution, all of its banks have been state owned. Although those banks continue to be state owned today, the Chinese government restructured and reformed its banking industry in the late 1970s and early 1980s.

The People's Bank of China, which was China's sole bank during Mao's regime, engaged in both commercial and central banking; in 1984, it became China's central bank exclusively. The four major state-owned banks that provide banking services are listed here, along with the year they were restructured:

- Bank of China (1979)
- Agricultural Bank of China (1979)
- China Construction Bank (1981)
- Industrial & Commercial Bank of China (1984)

These major state-owned banks have been responsible mainly for financing China's large state-owned enterprises. These major banks did not provide loans to China's small village firms in the early stages of industrialization, despite the strong profit motives of those village firms. Initially, risk and distance prevented these banking relationships. How, then, did village firms raise the funds to purchase materials, secure equipment, and pay salaries?

Initially, village firms were "self-financed" in the 1980s and 1990s by two complementary methods: pooling farmers' somewhat meager savings and direct loans from local, collectively owned credit unions. As Lu (2006) states, "smaller commercial banks and many nonbank deposit institutions… organized on a shareholding basis…serve[d] local needs." Most of these firms also relied on trade credits to finance working capital.

Market-oriented reform of the banking industry emerged from the October 1992 national congress, when the Communist Party leadership agreed to establish the "socialist market economic system."

> This decision immediately accelerated reforms on all fronts, including the banking and financial sector. The promulgation of a central bank law and a commercial bank law in 1995 marked a watershed between a centrally planned "monobank" system and a post-reform modern central banking system based on fractional reserves. (Lu, 2006, p. 6)

Since then, during China's second industrial revolution, many private firms have invested in new technologies that have been partially financed by China's major state banks and partially self-financed through these firms' retained cash flows.

China's entry into the World Trade Organization (WTO) in 2001 imposed requirements for additional reform of its financial sector, including allowing entry to foreign financial firms. That process, which has included measures taken to prepare its domestic banks to compete, is ongoing. By October 2005, for example, 138 foreign banks were approved to conduct yuan-based banking services, with assets amounting to $84.5 billion, equivalent to 2 percent of total assets in China's domestic banking sector at the time.

SOURCE: Lu (2006) and Wen (2016a) and the references therein. Also see Lu for a description of the "inherited links" between China's large state-owned enterprises and large state-owned banks.

Britain's monopoly power and hegemony in the global textile market and cotton trade. Such a powerful merchant class was obviously lacking in 1978 China, and thus China was stymied by a missing-market-creators problem. So, how would China ignite its proto-industrial revolution almost as soon as the reform started in 1978?

Deng's government imposed a national ideology: economic development through all possible means conditioned on political stability and social order. Government officials were expected to find ways to bring material wealth to local people. Fierce competition for economic success in both rural and urban areas effectively turned all levels of Chinese government officials into a highly motivated "public merchant" class. Through merit-based selection and competition with neighboring areas, there emerged a new generation of very capable business-minded administrators who helped create local, national, and international markets for local business by supporting village firms with low taxes and cheap land, attracting outside investment, advertising local products, negotiating business deals, and building distribution networks.

These market creators did not bear the stigma of traditional merchants; they were not seen by the Chinese populace as profiteers, traders, arbitragers, and opportunistic salesmen. They reinvented and revolutionized the historical European merchant-based putting-out system within China, except on a much larger scale and with an overtly nationalistic mission: They provided critical entrepreneurial and managerial services to village firms by acting as CEOs and members of "boards of directors" (à la Jean Oi, 1992), providing credit through China's state banking system, enforcing payments, supplying commercial information, organizing industrial parks and trade exhibition forums, and negotiating with out-of-region entities for the supply of raw materials and intermediate goods needed for production. These officials also sometimes even coordinated the absorption of inventories and the smoothing of supply and demand shocks to firms. They also helped organize farmers in their spare time to build roads, improve irrigation systems, or obtain loans from provincial or national banks to build local infrastructure. According to Oi (1992), "The impressive growth of collective rural industrial output between 1978 and 1988 is in large measure a result of local government entrepreneurship. Fiscal reform has assigned local governments property rights over increased income and has created strong incentives for local officials to pursue local economic development. In the process, local governments have taken on many characteristics of a business corporation, with officials acting as the equivalent of a board of directors."[17]

Hence, this "Chinese style" rural industrialization occurred through the emergence of a large number of collectively owned village firms. This process immediately ended China's shortage economy caused by the central planning of Mao's era. In less than 5 years after the 1978 reform, China had successfully ended all rationing imposed on food (including meat), textiles, and other light consumer products.

Another critical distinction in China's path to industrialization was its unprecedented pace and scale, compared with more than 200 years of proto-industrialization that occurred in 16th to late 18th century Europe. In merely a 10-year period after Deng's reforms began, between 1978 and 1988, the number of village firms, their industrial gross output, and farmers' aggregate wage income all increased more than 10-fold. Employment tripled. China's explosive growth continued throughout the 1990s and 2000s in a type of chain reaction in which expansion led to more expansion and growth led to more growth. By 2000, the number of workers in village firms had reached more than 128 million (not including the migrant workers in the cities), accounting for a remarkable 30 percent of China's entire rural labor force. Village industrial gross output reached 11.6 trillion yuan, a 16.5-fold increase compared with its 1988 value or a 225-fold increase compared with its 1978 value; its average rate of growth was 28 percent per year between 1978 and 2000, doubling every 3 years, and the total increase in real gross output of village industries was at least 66-fold over the 1978-2000 period.[18] This scale and speed of long-lasting economic growth is unique in economic history.

Again, this stage of China's industrial revolution replicates a comparable stage during the British Industrial Revolution, which also started in the countryside. In China's case, tens of millions of village enterprises arose in the vastly impoverished rural areas in the late 1970s and early 1980s. These village firms were organized and managed by the uneducated peasants who were not much different from their Qing dynasty ancestors in the 17th and 18th centuries.

Some economic historians and the human-capital school of development attribute China's failure to attain industrialization at that time to the lack of education among these peasant farmers. But in fact it was equally uneducated peasant farmers who brought China's industrial revolution to fruition in the late 20th century.[19]

So, the puzzle is no longer why a proto-industrialization was suddenly kick-started in China after 1978, but rather why it did not happen earlier in Chinese history, despite private property rights, such as those during its first and second attempts of industrialization in the 19th and early 20th centuries.

The answer to this puzzle is now much clearer: In the Qing Dynasty and the Republic era, China did not have a well-fermented unified domestic or global market to support the division of labor and it did not have a large number of market creators and rural-firm organizers. During these early attempts at industrialization, the absence of markets and market-creators could have been remedied only by some kind of intervention. As noted, previous industrializations were led by a powerful class of merchants supported by a strong-willed and militarily powerful pro-commerce and pro-manufacturing mercantilist government and motivated by monopoly profits in the world market through, for example, armed trade and colonialism.[20]

Step 3: China's First Industrial Revolution

> *"It is not worth my while to manufacture your engine for three counties only, but I find it very well worth my while to make it for all the world."*
> **—English manufacturer Matthew Boulton (1728-1809), to his business partner James Watt (1736-1819), cited by Eric Roll, [1930], 1968, p. 14)**

What good would it do to adopt the division of labor in Adam Smith's pin factory if the market demand were only one pin per day instead of 40,000 per day? Governments in developing countries are often eager to modernize their economies by adopting the latest, most-efficient mass-production technologies: Why bother to use slower, outdated technologies when faster, more-productive technologies are available? But without finishing a proto-industrialization and engaging in a first industrial revolution, such a process is doomed from the start. The relationship between mass production and the size of the market is key.

After a decade of rapidly building up proto-industries and commercial networks, unifying its domestic market, and expanding that market through international trade, China reached the tipping point of its first industrial revolution in the late 1980s. The flagship industry of China's first industrial revolution was textiles and clothing.

China's primitive agricultural revolution allowed for some economic gains among the rural population, and the commerce associated with the proto-industrialization rapidly improved the living standard and purchasing power of that population. Hence, *local* demand for textiles and apparel continued to rise throughout the 1980s, owing to the income elasticity of these goods. Fueled by this rising demand as well as intense competition among small firms, mass production of textiles and garments became profitable. As a result, China's total production of yarn and cotton fabrics increased from 330,000 tons and 1.9 billion meters in 1985 to

8.5 million tons and 32.2 billion meters in 2002, with a 23-fold increase for yarn and a 15-fold increase for cotton fabrics over 17 years (the implied annual growth rate is 20 percent and 17 percent, respectively). Total garment output increased from 1.3 billion pieces in 1985 to 9.5 billion pieces in 1996, with an average growth rate of 22 percent per year. Total chemical fabric production increased from 94.8 thousand tons in 1986 to 991.2 thousand tons in 2002, growing by 16 percent per year on average. As early as 1990, there were already tens of millions of spindles in the east and south of China with well-formed industrial production chains and textile manufacturing clusters. By 1994-95 (more than 6 years before joining the World Trade Organization [WTO]), the number of spindles reached 40 million, one for every 25 people in China.

This growth was driven initially by the large state-owned textile enterprises (which gradually became profitable during the proto-industrialization) because of their scale of operations and easy access to finance, but then was dominated primarily by privately owned enterprises as soon as they caught up with the mass-production technologies through self-financed investment. The profits of these privately owned enterprises grew by 23.5 percent per year between 1990 and 1997.

As a result, the textile and clothing industry became the largest manufacturing industry and major source of foreign exchange in China during its first industrial revolution, between 1988 and 1998. This industry included about 24,000 enterprises and employed about 8 million workers even in the early 1990s; its exports accounted for more than 20 percent of China's total exports. China surpassed the United States and became the world's largest producer and exporter of textiles and clothing in 1995, six to seven years before joining the WTO, and has retained this dominant position ever since (see Wen, 2016a).[21]

Again, the Chinese government played a pivotal role in each of the stages of China's industrialization and certainly did so in this textile-led first industrial revolution. China's government in 1979 chose to implement additional economic reforms and policies to target the nation's textile and clothing industry, which early on was one of the primary industries it promoted. Previous attempts at industrialization, as under Mao's regime, focused on promoting heavy industries such as steel. The Chinese government under Deng promoted the textile and clothing industry for three key reasons: (i) This industry was consistent with China's comparative advantage in its abundance in labor, (ii) it did not require advanced technologies and had relatively low entry costs, and (iii) the domestic and international markets for these products were huge.

To promote the textile industry, the government launched a policy called "Six Priorities," which favored the textile industry in six areas: supply of raw materials, energy and power, bank loans, foreign exchange, imported advanced technology, and transportation (see, e.g., Qiu, 2005). The Chinese government was directly involved in the import and storage of cotton nationwide to smooth domestic cotton prices and demand. Moreover, it established sophisticated organizations to nurture this industry. The following government agencies were created (long before China joined the WTO) to supervise, regulate, and assist the textile and clothing industry in coping with international textile market rules and competition, each with specific functions and areas of focus.

- China Chamber of Commerce for Import and Export of Textiles
- China National Textile Industry Council
- China Petroleum and Chemical Industry Association
- Ministry of Agriculture
- Ministry of Commerce
- National Development and Reform Commission
- State Administration for Quality, Supervision, Inspection, and Quarantine
- State Environmental Protection Administration
- State-owned Assets Supervision and Administration Commission
- Textile Industry Standardization Institute

Raw Material Supply. The Ministry of Agriculture is responsible for key raw material industries including cotton, silk, and wool. The National Development and Reform Commission (NDRC) is responsible for the importation of raw materials, for which import quotas still apply.

Production and Processing. China National Textile Industry Council (CNTIC) guides the production and processing in the textile industry. CNTIC is the legacy agency of the now-defunct Ministry of Textile Industry. Its broad responsibilities include the implementation of industrial development guidelines for the sector.

Export Quota License. The NDRC's Department of Industry supervises the national textile industry. The Bureau of Economic Operation is responsible for formulating policies and controlling the export quota licensing system in the textile industry. However, the Ministry of Commerce is responsible for actually issuing export quota licenses.

Standards-Setting. The State Administration for Quality, Supervision, Inspection, and Quarantine (AQSIQ) is the government agency responsible for setting technical, safety, and environmental protection standards for textile products in China. In the textile sector, AQSIQ functions as a standards-setting coordinator. When setting standards, it seeks technical support from the Textile Industry Standardization Institute (TISI) and consults with the CNTIC. AQSIQ is also the agency responsible for enforcing standards and providing certification of products and enterprises. AQSIQ is also involved in drafting laws and regulations governing industrial standardization in the textile sector.

The textile industry was instrumental in China's first industrial revolution and led the way to China's second industrial revolution in the late 1990s; this progression strongly resembles the pattern of the British Industrial Revolution and hence sheds considerable light on the long-standing puzzle and internal logic of an industrial revolution in general.

The textile industry was also the flagship industry during the first industrial revolution in Great Britain. From the 1760s to the 1830s, a series of inventions of simple yet powerful wood-framed tools and machines rapidly sped up spinning and weaving. However, the British Industrial Revolution was not driven merely by these technological inventions, per se, as the conventional wisdom often assumes. Rather, it was driven mainly by the colossal textile *market* created by British merchants and the government and was the outcome of fierce competition among the European proto-industrial textile firms for market share.

Textile production is much easier to mechanize with simple low-cost tools than growing crops and building shelters; it is also much easier to divide this type of production into many intermediate stages in an environment of division of labor. Textile production is simple enough that even young or otherwise unskilled workers can easily accomplish it. It can involve longer working hours and thus can potentially absorb a huge amount of surplus labor from rural areas in which only agricultural work had been done.

The textile market is potentially the largest and most income-elastic, compared with other light consumer goods such as jewelry, pottery, or furniture; hence, the textile market has the potential for rapid growth as incomes rise and mass-production technology progresses. Moreover, the competition inherent in this market naturally stimulates innovation.[22]

Before the Industrial Revolution, Great Britain had nurtured its textile market for hundreds of years, at least since Elizabeth I (1558-1603) and possibly even earlier. These interventions created Europe's largest textile market and distribution network by the early 18th century, and Great Britain eventually possessed the largest number of early textile firms. However, by the early and middle of the 18th century, after centuries of proto-industrialization and the boom in textile production across Europe, the woolen and linen textile markets for British textile products (based on artisan workshops) appeared virtually saturated. Yet the demand for cotton textiles was growing rapidly while the supply (imports from India) was restricted by the British mercantilist laws to protect its domestic woolen textile market. This environment was immensely competitive. This competition was critical for stimulating technological innovation and discovery of new varieties of consumer products: To survive market competition, firms needed to adapt and gain new market shares. To nurture its textile industry, Great Britain would reshape the market, as exemplified by the government-promoted shift from the traditional woolen textiles to cotton textiles in the 1730s (e.g., as reflected in the Manchester Act in 1736), the shift from workshops to cotton mills in the 1740s, and the subsequent Industrial Revolution.

Hence, it is not surprising that the Industrial Revolution started first in Great Britain and first in this particular industry—because only a massive market with mature distribution networks and highly income-elastic demand could stimulate and sustain profitable mass production through mechanization. Interestingly, this economic logic has not changed since the British Industrial Revolution. Virtually all recently developed nations followed the same path paved by the British to successfully kick-start their own first industrial revolution.[23]

Step 4: China's Industrial Trinity Boom and Second Industrial Revolution

The industrial trinity is defined as three key industries: energy, motive power, and infrastructure. Infrastructure includes but is not limited to transportation and communication. The industrial trinity represents the flagship industries during the initial phase of a second industrial revolution.

China kick-started its massive buildup in energy and infrastructure only around the middle of the 1990s, after finishing or nearly finishing its first industrial revolution, because only then did such capital-intensive industrial projects become affordable and profitable.

The boom in the industrial trinity was triggered and supported by the market demand created by the first industrial revolution. Moreover, the boom itself generates colossal demand

Three Gorges Dam

The Three Gorges Dam is located across the Yangtze River in the town of Sandouping, in the Yiling District of Yichang Prefecture in the province of Hubei. The dam is the world's largest power-production facility: As of 2014, it generated 98 terawatt-hours of electricity. The dam was built also to increase the shipping capacity of the river and reduce flooding downstream.

Damming the Yangtze River was long imagined and supported by Chinese leaders, including Sun Yat-sen, the founder of the Republic of China, and Mao Zedong, after the communist takeover. The National People's Congress, in 1992, approved the project and finally secured enough support and funding.

The dam was fully functional as of July 4, 2012. A ship lift was completed in December 2015. The estimate of full cost recovery is once 1,000 terawatt-hours of electricity is generated, which translates to a yield of 250 billion yuan.

Preparation	Moved 102.6 million cubic meters (134.2 million cubic yards) of earth and more than a million residents
Construction	27.2 million cubic meters (35.6 million cubic yards) of concrete; 463,000 tons of steel
Dam Size	2,335 meters (7,661 feet) long; 185 meters (607 feet) above sea level
Reservoir Size	175 meters (574 feet) above sea level (at its highest level); 110 meters (361 feet) higher than the river level downstream 660 kilometers (410 miles) long; 1.12 kilometers (3,700 feet) wide 39.3 cubic kilometers (31,900,000 acre·feet) of water 1,045 square kilometers (403 square miles) of total surface area 632 square kilometers (244 square miles) of land flooded on completion
Total Cost	180 billion yuan (US$22.5 billion) initially estimated 148.365 billion yuan spent (US$18.5 billion): 64.613 billion on construction, 68.557 billion on relocating residents, and 15.195 billion on financing as of 2008

SOURCE: Mufson (1997), Jones and Freeman (2001), and Chinese National Committee on Large Dams (2010).

for heavy industrial goods and materials, which in turn provides economic forces and markets to support the second industrial revolution, which features mass production of the means of mass production and mass distribution (such as heavy intermediate goods, machinery, and infrastructure).

The second industrial revolution enables a society to provide a large and steady supply of machinery and various intermediate goods as well as means of mass distribution to sustain the continuation of the first industrial revolution. In other words, this environment calls for the mass production and provision of heavy industrial goods such as chemicals, cement, iron, steel, communication equipment, automotive products, ships, cars, trucks, airplanes, and a large organized credit system. Any new discoveries or inventions that facilitate the efficient supply of these goods will necessarily be adopted into the production process, as long as their benefits outweigh their costs. These innovations include new forms of energy, motive power, transportation, and communication and new (man-made) materials. Also, innovations in

High-Speed Rail

"High-speed" rail refers to commercial railway train service that can achieve speeds of 200 km/h (124 mph) or higher, the international standard. Commercial train service in China in 1993 averaged only 48 km/h (30 mph) and was inadequate to satisfy increasing demand for transportation of passengers and cargo. The Chinese government attempted to modernize the railway system by, first, increasing the speed and capacity of existing lines through double-tracking, electrification, grade improvements, reduced turn curvature, and use of continuously welded rail. China's "Speed-Up" campaigns in April 1997, October 1998, October 2000, November 2001, and April 2004 upgraded passenger service on 7,700 km (4,800 miles) of existing track to just below the threshold of "high-speed": 160 km/h (100 mph). Currently, China has the world's longest high-speed rail network: as of January 2016, over 19,000 km (12,000 miles) of track, which is more than the rest of the world's high-speed rail tracks combined. Plans are in place for a network of 30,000 km (19,000 miles) by 2020.

SOURCE: Wen (2016a) and the references therein.

financial services and credit management and a stable and well-managed national banking system are needed to facilitate the large volume of trade.

Some examples: The construction of the world's largest hydroelectric power station, the Three Gorges Dam, began on December 14, 1994. Except for a ship lift that was completed in 2015, the dam project was completed and fully functional on July 4, 2012.

And the buildup of China's high-speed rail network started only in the late 1990s. Since the operation of China's first high-speed railroad in 2008, 28 Chinese provinces (out of 30) are now covered by the world largest and longest high-speed rail network (more than 19,000 kilometers in length and 50 percent greater than current world capacity outside China).

Vast improvements have been made during the past two decades in irrigation systems, sewage systems, street and highway networks, air and rail transportation, electrical grids, gas and oil pipelines, and so on. For example, the total length of public roads reached 4,230,000 kilometers (about 2,643,700 miles), including 111,950 kilometers (about 70,000 miles) of highways at the end of 2014, surpassing the U.S. system as the world's longest highway system. More than 95 percent of China's villages are now connected by asphalt roads. As a result, China now enjoys an exceptionally high ranking in the World Bank Logistics Performance Index (LPI). China is one of the few developing countries to achieve an LPI score comparable to that of high-income nations in international shipments, infrastructure, custom services, logistics competence, tracking and tracing, and timeliness, with an overall LPI score of 3.53 in 2014, ranked 28th in the world, next to Portugal but above richer countries such as Turkey, Poland, and Hungary. (See World Bank, 2014.) Moreover, China's infrastructure-construction boom is still continuing at an unprecedented speed both domestically and internationally. Such remarkable catching-up in infrastructure has no doubt fed-back and made a significant contribution to China's rapid market formation and prepared China well for the next decade of growth in industrialization.

OVERVIEW OF CHINA'S GOVERNMENT INVOLVEMENT AND PRIVATIZATION POLICIES

The Chinese government has had a role in each of these stages of industrialization. Beyond the general expectations for providing social order and political stability, China's government worked to overcome the problems of missing markets and market-coordination failures during each of the stages described here. In addition, the government plays another critical role: Industries generate enormous positive externalities for the national economic system that only the state can fully internalize—especially in the areas of energy, motive power, financial services, and infrastructure, which are pivotal for overall development and national security.[24] China's second industrial revolution, starting in the mid-1990s, benefited tremendously from the large-scale state-owned heavy industries and scientific research institutions established during Mao's era.[25] These heavy industries and research institutions had been highly inefficient and unprofitable and were large financial burdens for China. But they did not remain so. Once China finished its proto-industrialization and its first industrial revolution, it adopted a competitive, profit-driven approach to managing heavy industries and a merit-based reward system for research and innovation.[26]

The Chinese government (wisely, as it turns out) chose to retain its "inefficient" state-owned heavy industries in the 1980s and early 1990s instead of dismantling them through marketization and privatization, which is what Russia did during its initial "shock therapy" reforms in early 1990s. China maintained many important state-owned enterprises and postponed their reform until the late 1990s, after China finished both its proto-industrialization and its first industrial revolution.[27] By the late 1990s, China had become the world's largest market for modern infrastructures and heavy industrial goods. Only a market as large as this would be able to profitably sustain large state-owned heavy industries. Hence, it was much easier for China to engage in market-oriented reform and restructure its state-owned heavy industries at this time than, say, in the late 1970s and 1980s or even early 1990s.[28] Again, in contrast, Russia's state-owned heavy industries were mostly abandoned or destroyed by their "shock therapy" approach to reform and the ensuing so-called "market forces" in the 1990s. China, however, took a more patient approach and leveraged its large domestic market to successfully build its colossal light industrial base and expand its purchasing power to sustainably finance its large-scale heavy industries.

China's national saving rate is nearly 50 percent and its aggregate investment rate is 45 percent; the inflows of foreign direct investment in manufacturing from industrial economies have been extensive since the mid 1990s, as has China's rapid advancement in heavy industrial technologies such as steel, ship-building, high-speed rail systems, and space programs—most of which are state-owned.

An important lesson learned from China's privatization experience (in comparison with Russia's) is that a nation should be extremely cautious in privatizing its state-owned enterprises. It is dangerous to blindly or indiscriminately privatize all industries before market conditions are ready. The market conditions for privatizing a particular industry are ready if and only if (i) the market is large enough to support similar-type private firms; (ii) private

firms in this industry are well developed and sufficiently competitive domestically or internationally in finance, management, and technological innovations; and (iii) privatization does not put national security at risk and key industries (such as natural resources) may be only merged or engaged in joint ventures as opposed to fully privatized.[29]

It is extremely costly to create a market and even more costly to create the regulatory institutions to regulate market activities. Without forceful and appropriate regulations, markets will malfunction and market forces can destroy social trust—which is the very foundation of the market itself. Yet the Washington consensus and the institutional theory have offered no instructions to developing nations on how to build market-specific regulatory institutions to prevent or mitigate the destructive power of market forces and corporate freedom, especially with regard to deregulation, liberalization, marketization, privatization, and democratization.

CONCLUSION

Poverty is always and everywhere a social coordination-failure problem. The problem arises because creating markets and the corresponding economic organizations (based on the principle of the division of labor) are extremely costly and require broad and intense coordination efforts and trust from all market participants. Thus, Wen (2016a) states that the "free" market is the most fundamental public good in a market economy, and its most fundamental pillar is social trust. The benefits of the market are largely social, while its costs (of creation and participation) are largely private.

Hence, development's first and foremost challenge is in overcoming both missing markets and missing market creators. Historically, a natural process of mass-market formation/fermentation has been a lengthy evolutionary process initially accomplished mainly by a large and powerful merchant class that has acted collectively under a nationalistic mercantilist spirit backed fiercely by their government.

China's development experience has shown the world that the centuries-long Western-style "natural" and lengthy market-fermentation process can be dramatically accelerated and re-engineered by the government when it supplies the market creators—in place of the missing merchant class—yet without repeating the Western powers' old development path of primitive accumulations based on colonialism, imperialism, and the slave trade.

China's development experience suggests a new model of economic development: the New Stage Theory (NST) or "Embryonic" Development Theory (EDT). The NST is closely related to the old stage theory of List (1841), Marx (1867), and Rostow (1960) and the other schools of development theory, such as Structuralism and New Structuralism (Justin Yifu Lin) and the ISI and the "Big Push" theory of development (as advocated by Paul Rosenstein-Rodin in 1943 and Kevin M. Murphy, Andrei Schleifer, and Robert W. Vishny in 1989).

NST emphasizes that, for late-developing countries with a tremendous lag in reaching the technological frontier and despite the advantage of backwardness, repeating some of the earlier development stages of now-developed nations is necessary. Consider the study of mathematics: After thousands of years of development, the human race discovered math knowledge sequentially, from numbers to arithmetic to algebra to calculus etc. Although cal-

culus is in today's first-year college textbooks, every generation of children must still repeat humanity's evolutionary process to learn math. They do not jump to calculus at age 6, but instead start with learning numbers (with the help of their fingers, just as our ancestors did) and gradually move up the ladder to more advanced forms. Yet modern development theories focus almost exclusively on adoption of frontier technology or financial liberalization as the key to industrialization for agrarian societies, without realizing that supply does not create its own demand. The mode of mass production is not profitable when a mass market and mass distribution do not exist. So, industrialization is first and foremost a task of market creation. The creation of a mass market must always proceed through several major and distinctive stages—sequentially—with each stage facing its own specific challenges related to market failures and missing market creators. Hence, poverty and the development problem cannot be solved by political democracy, as so many expected or hoped for during the "Arab Spring." It also cannot be solved by an intense effort through "shock therapy" or a one-time colossal national investment boom facilitated by foreign aid or a top-down heavy-industry-based approach, as advocated by the old stage theories and the Washington consensus. Instead, successful economic development requires many rounds of sequential effort led by a powerful mercantilist government under political stability, coordinated between local and central governments, but that begins at the grassroots level.

The new institutional theory (e.g., Acemoglu and Robinson, 2012) suggests that the Industrial Revolution started in Great Britain, instead of 18th century China or India, because it was Great Britain that first developed inclusive political institutions (through the Glorious Revolution) and the rule of law. This view is misleading and inconsistent with historical facts. As economic historian Sven Beckert aptly put it, "The first industrial nation, Great Britain, was hardly a liberal, lean state with dependable but impartial institutions as it is often portrayed. Instead it was an imperial nation characterized by enormous military expenditures, a nearly constant state of war, a powerful and interventionist bureaucracy, high taxes, skyrocketing government debt, and protectionist tariffs—and it was certainly not democratic" (Beckert, 2014, p. xv).

Furthermore, the institutional theory cannot explain (i) why there are many democracies with pervasive economic stagnation and continuous political turmoil, such as Afghanistan, Egypt, Iraq, Libya, Pakistan, Thailand, Tunisia, and Ukraine; (ii) why there are ample extractive institutions that have been economically strong, such as Germany (1850 to World War II) and Soviet Russia (1860 to World War II); and (iii) modern-day Russia's dismal failure in economic reform under democracy and shock therapy, Japan's rapid industrialization during the Meiji Restoration, South Korea's economic takeoff in the 1960s-1980s under dictatorship, and Singapore's post-independence economic miracle; and (iv) the fact that, under identical political institutions, property rights, and the rule of law, there exist pockets of both extreme poverty and extreme wealth and both violent crime and obedience to the law in many cities and regions of the same country, including the United States.

The degree of industrialization is limited by the extent of the market a nation provides for its firms. Therefore, the fundamental reason Great Britain, instead of the politically more liberal Netherlands, kick-started the (first) Industrial Revolution in the late 18th century was

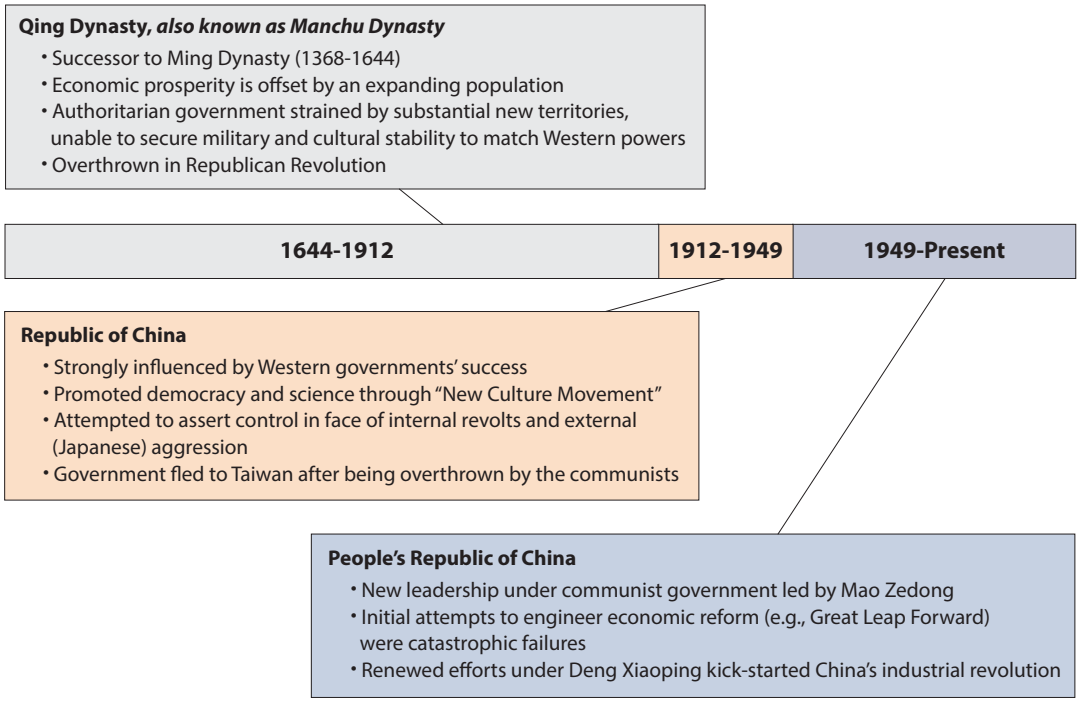
because of its successful creation of the world's largest textile market and cotton supply chains, which made the nationwide adoption of the spinning jenny and factory system profitable and inevitable. Likewise, the fundamental reason the United States, instead of France or Germany, overtook Great Britain to become the next economic superpower was that the United States, with the help of its government, had created an even larger manufactured goods market in the 19th century, which nurtured the world's greatest inventors such as Thomas Edison and industrial giants such as Andrew Carnegie, Henry Ford, J.P. Morgan, John D. Rockefeller, and Cornelius Vanderbilt.[30]

What the NST suggests in general is that industrial policies and development strategies matter and are responsible for the failures and successes of nations when they attempt to escape poverty. Political institutions, which are endogenous to economic development, are not responsible, except in providing political stability to support commerce. In fact, many different political systems can provide such commerce-friendly political stability, such as monarchies, republics, autocracies, or meritocracies. Given that most nations and their governments do want to develop their economies and have tried very hard repeatedly in the past to industrialize, it is difficult to argue that their failures are due to the government's lack of incentives to develop because of vested interests under a non-democratic system (Acemoglu and Robinson, 2012). In fact, it is in the interest of the poor nation's government to develop if they want to stay in power.[31] Unfortunately, many poor nations and their government leaders get stuck in the process of industrialization because of implementing the *wrong* industrial policies and development strategies, just as China did in its previous three failed attempts at industrialization. To be sure, although Western style democracy is not likely the precondition that will sustain China's growth, market mechanisms and good governance are.

Institutions are endogenous and built to facilitate the execution and implementation of developmental policies and strategies and to protect the fruits of industrialization. Therefore, it is reasonable to expect that China's "backward" financial and legal institutions will be history if China continues to develop based on its gradualist development strategy: move up the industrial ladder from light to heavy industries, from labor to capital-intensive production, from manufacturing to financial capitalism, and from a high-saving state to a consumeristic welfare state. China has had only 35 years of genuine industrialization, which can only be described as short when compared with 300 years of volatile capitalism in the Western world. China may require at least another 30 years or so to clearly demonstrate whether it can build a modern financial system to facilitate its enormous economy and a modern legal system to protect the fruits of its industrial revolution. Only then can China be judged more fairly and on a more equal footing with Western nations.[32] ■

APPENDIX

Recent Regimes and Events in China



1842	China defeated by British in First Opium War.
1860	China defeated by British in Second Opium War; Qing monarchy attempts to establish modern navy and industrial infrastructure.
1895	China defeated by Japanese in Sino-Japanese War.
1900	Boxer Rebellion.
1911	Qing monarchy is deep in debt; social unrest leads to the Xinhai Revolution, which overthrows the monarchy and establishes Republic of China (the nation's first "inclusive" government).
1912	Republic of China commences under Sun Yat-sen, who is soon replaced by Yuan Shikai.
1919	May Fourth Movement: Students and workers protest China's acceptance of Treaty of Versailles, which relinquished land (formerly under German control) to Japan. Surge in Chinese nationalism.
1921	Chinese Communist Party organizes in Shanghai.
1937	Japan invades China.
1945	Mao outlines his "New Democracy." War with Japan ends.
1949	Communists / People's Liberation Army occupy Beijing and Shanghai. Mao Zedong proclaims founding of People's Republic of China. Chiang Kai-shek's government flees to Taiwan.
1950-1952	Land reforms implemented.

1953	Five-Year Plan of economic growth and development begins. Mutual Aid Teams organized in Chinese countryside. Chinese Communist Party Central Committee authorizes Agricultural Producers' Cooperatives.
1955-1956	Mao Zedong intervenes to speed up formation of rural Agricultural Producers' Cooperatives, greatly increasing their numbers, but severely disrupts agricultural production.
1958-1961	Great Leap Forward. Food crisis intensifies.
1966	Cultural Revolution begins. Little Red Book is published.
1967	Deng Xiaoping and Liu Shaoqi accused of crimes against Chinese Communist Party.
1969	U.S. partially lifts trade embargo against China.
1972	U.S. President Nixon visits China.
1973-1974	Deng Xiaoping "rehabilitated" as vice premier, addresses United Nations General Assembly, reappointed as vice chairman of Chinese Communist Party.
1976	Deng Xiaoping purged from Party; Mao Zedong dies.
1977	Deng Xiaoping, once restored to Party, begins push for major reforms.
1984	Deng Xiaoping tours Special Economic Zones and advocates for continued economic reform; U.S. President Reagan visits China.
1986	China becomes member of Asian Development Bank.
1989	Tiananmen Square incident.
1992	Deng Xiaoping visits Special Economic Zones in Shenzhen, leading to further economic reforms. China's military spending increases 13 percent. National Party Congress approves plans for Three Gorges Dam project. One million workers laid off from inefficient state-owned enterprises. At 14th National Party Congress, principle of "socialist market economic system" is promoted.
1993	Jiang Zemin appointed president of People's Republic of China.
1994	U.S. extends most-favored nation status to China, separating human rights and trade issues.
1995	China adopts its first banking laws, the Law of the People's Bank of China. China Construction Bank and Morgan Stanley launch China International Capital Corporation, first joint venture investment bank in China. China formally requests to join World Trade Organization.
1996	Kelon becomes first Chinese township enterprise listed on the Hong Kong stock exchange.
1997	Deng Xiaoping dies.
2001	China enters World Trade Organization.
2003	China surpasses U.S. as world's largest recipient of foreign direct investment. Four million families in China own automobiles.
2004	China becomes sixth-largest economy in the world. China publicizes several of its billionaires. Provisions to protect human rights and private property incorporated into Chinese constitution. China reaches 87 million internet users.
2005	China replaces U.S. as Japan's largest trade partner, with foreign exchange reserves second only to Japan. End to global textile quotas leads to surge in exports from China to U.S.
2010	China becomes second-largest economy in the world.
2012	The Three Gorges Dam, the world's largest power station, becomes fully functional after 18 years of construction at a cost of 148 billion yuan ($22.5 billion).

SOURCE: Sullivan (2007), Bruton, Lan, and Lu (2000), and Wen (2016a) and the references therein.

NOTES

[1] For a profile of a typical Chinese local government official and his role as a "public merchant," see Chapter 6 in Wen (2016a).

[2] Also see Wen's (2016b) short article in the Federal Reserve Bank of St. Louis *The Regional Economist*.

[3] In 1949, China's peasant population as a share of its total national population remained at more than 90 percent, not much changed since the Second Opium War around 1860. Average life expectancy remained as low as 30 to 35 years and the literacy rate was only 15 to 25 percent.

[4] For example, the young Mao Zedong was a high-ranking official member of the republican government in the early 1920s.

[5] One extremely important legacy of Mao's era, seldom mentioned and appreciated in the existing literature, is the critical level of "social capital" (including social trust in general and farmers' abilities to organize themselves) and a unified national market with a potential size of four times that of the U.S. market. As mentioned earlier and discussed later, social trust and national unity are fundamental pillars of a unified "free" market on which the division of labor is based.

[6] The first time was in 1966 at the start of the Cultural Revolution, and the second was around 1976 after the Cultural Revolution but before Mao's death. See Vogel (2013).

[7] See Marangos (2007) and Williamson (2004).

[8] To conserve space, this article does not provide a systematic analysis on why such developmental strategies fail. Interested readers are referred to Wen (2016a) for such an analysis and critical evaluation of these development policies.

[9] According to Studwell (2013, p. 17), "In the 1920s, when 85 percent of Chinese people lived in the countryside, life expectancy at birth for rural dwellers was 20-25 years. Three quarters of farming families had plots of less than one hectare, while perhaps one-tenth of the population owned seven-tenths of the cultivable land."

[10] During the Great Leap Forward, there were 6 million village factories created. But most of them were forced to shut down after 1962 due to the great famine. But a fraction of them survived because of the protection of local villagers. The initial 1.52 million village firms in 1978 were the legacy of the Great Leap Forward and served as the catalyst of China's long-awaited rural industrialization. Of course, the benefits achieved by Mao's communist regime do not deny or excuse the hardships and crimes perpetrated against the Chinese people—including violating human rights during the Cultural Revolution. Actually, the Cultural Revolution might have destroyed some of the social capital (mostly social trust) built by Mao in the 1950s and early 1960s. However, this loss of trust did not occur in the rural areas. And it was that sustained social capital in the rural areas that was the most critical for setting off China's proto-industrialization.

[11] See Mendels's (1972, 1981) analysis of the phenomenon of proto-industrialization in European economic history and the literature stimulated by this. What is most puzzling about China's proto-industrialization is that it did not occur in Qing Dynasty China or during the Republic era, despite a market system and well-protected private property rights in the rural areas.

[12] Housing has been one of the major areas of growth during China's industrialization. We refer interested readers to "The Great Housing Boom of China," a working paper by Chen and Wen (2015, forthcoming in *American Economic Journal: Macroeconomics*).

[13] "The putting-out system exploited the benefits of the division of labor to the full" (from the *Oxford Encyclopedia of Economic History*, 2003, p. 101; see also *International Encyclopedia of the Social Sciences*, 2008). The putting-out system was family-based domestic manufacturing prevalent in the rural areas of Western Europe during the 17th and 18th centuries. It appeared even earlier in 16th century Italy. Domestic workers typically owned their own primitive tools (such as looms and spinning wheels) but depended on merchant capitalists to provide them with the raw materials to fashion products that were deemed the property of the merchants. Semi-finished products would be passed on by the merchant to another workplace for further processing, while finished products would be taken directly to market by the merchants. Even independent domestic craftsmen working on their own also relied on merchants to introduce their products in distant markets.

[14] Under the putting-out system, tools of production have often been owned by the peasant households, but sometimes by the merchants who have rented them to the peasant workers.

[15] Mercantilism is economic nationalism for the purpose of building a wealthy and powerful state based on commerce and manufacturing. It has sought to enrich the country by restraining imports of *manufactured* goods and encouraging exports of *manufactured* goods. In short, it emphasizes and promotes manufacturing over agriculture and commercialism over physiocracy. However, most of the literature on mercantilism views it simply as a form of protectionism or pure pursuit of trade surplus or gold reserves and overlooks the key point of commerce and, again, *manufacturing*. An economy relying solely on agriculture has nothing to benefit from mercantilism. But a nation intending to build on manufacturing can benefit greatly from mercantilism because manufacturing stimulates the division of labor and generates the economies of scale. The historical importance of mercantilism in the 16th to 18th centuries in Europe as the prototype of capitalism and the key step leading to the British Industrial Revolution can never be emphasized enough. Indeed, the promotion of manufacturing inherent in mercantilism has seldom been appreciated by classical economists, including Adam Smith and David Ricardo, unlike Friedrich List (1841). One example of the impact of mercantilism on economic development is the 19th century American industrial revolution based on the "American System," which was an economic development strategy envisioned by Alexander Hamilton (1755-1804) in 1791 and vigorously implemented throughout the 19th century to win global competition with Great Britain. (Hamilton's idea was not immediately adopted in the 1790s and the initial decade of the 1800s.) It consisted of several mutually reinforcing parts: high tariffs to protect and promote the American infant manufacturing sector in the North; a national bank to foster commerce, stabilize the currency, and rein in risk-taking private banks; maintenance of high public land prices to generate federal revenue; and large-scale federal subsidies for roads, canals, and other infrastructures to develop a unified national market—financed through tariffs and land sales. Also see Ha-Joon Chang's (2003) *Kicking Away the Ladder: Development Strategies in Historical Perspective* for many examples of mercantilism and the historical role it played in Western economic development. However, many Latin American countries in the middle 20th century also adopted various forms of mercantilism (e.g., import substitution industrialization) but failed miserably. The reasons behind such successes and failures are precisely what Wen's (2016a) book is about.

[16] Some historians believe that slavery and trans-Atlantic trade helped finance the British Industrial Revolution. Plantation owners, shipbuilders, and merchants who were connected with the slave trade accumulated vast fortunes that established banks and large manufactures in Europe and expanded the reach of capitalism worldwide. For scholarly articles on the critical contributions of slavery and trans-Atlantic trade to the Industrial Revolution, see, e.g., Williams (1994). Kriedte, Medick, and Schlumbohm (1981, p. 131) provides a related perspective: "For England, which was politically and militarily the most successful country, the 'virtual monopoly among European powers of overseas colonies,' established during the phase of proto-industrialization, was one of the central preconditions which carried proto-industrialization beyond itself into the Industrial Revolution." Economic historians Pomeranz and Topik (2013, p. 104) argue that opium trade "not only helped create Britain's direct [trade] surplus with China, but made possible even the larger surplus with India. Without those surpluses, Britain could not have remained the West's chief consumer and financer, and the Atlantic economy as a whole would have grown much more slowly."

[17] With China's institutional arrangement of public land ownership and the administrative power of local governments (a legacy of Mao's communism), farmers and peasants were able and willing to pool their savings to form the initial capital (cash and land assets) necessary for an initial investment in an establishment that by design was collectively owned, with profits and work opportunities equally shared among village farmers. Although land had been leased to individual families since 1978 under the family-responsibility system, the nature of the public ownership of land had not changed; acquiring land for industrial purposes, then, was not a great hurdle for the village farmers and the local governments. The managers of such collectively owned establishments were often the village officials, who were often democratically elected and viewed as natural leaders (China's earliest CEOs). Although Deng disbanded the communes that had been created under Mao's regime, the legacy of the Great Leap Forward and its communization movement made it easy to reintroduce collectively owned organizations. The high degree of trust among these village families and the leadership of the local governments enabled Chinese farmers to overcome the prohibitive transaction costs of contracting in an agrarian society where the legal system and law enforcement were lacking. In essence, they trusted fair income distribution and risk-sharing and credit payments.

[18] See the boxed insert on village firms for more details. The source is Wen (2016a) and the references therein. For the original data, see Zhang and Zhang (2001 [in Chinese], Appendix Table 1).

[19] But with one critical difference: Chinese farmers in the 1980s were experienced with self-organizational skills and endowed with social capital gained through Mao's Great Leap Forward and rural corporative movements.

[20] In the 16th to 18th centuries in Europe, the lack of social trust and the associated transaction costs in forming corporations in rural areas were mitigated and overcome by the entrepreneurial, risk-taking, profit-seeking merchants, who were less financially constrained and more experienced in long distance trade. But, again, it took centuries for Europe in general and England specifically to form such a powerful merchant class through commercialism, colonialism, imperialism, mercantilism, and the trans-Atlantic slave trade. This process of forming markets in Europe, England, and elsewhere around the globe under colonialism can be thought of as "natural market fermentation," where the key agents are the powerful merchants. This global market creation process is also extremely costly and requires that trade be secured and enforced by military pressure. European overseas explorations and trade were extremely capital-intensive because of the colossal costs and risks involved. Most long-distance trade carried out by European merchants included armed trade and was endorsed and supported militarily by their governments. This context is captured in the words of the famous Dutch merchant and warrior Jan Pieterszoon Coen to the Dutch monarch: "Your Honours should know by experience that trade in Asia must be driven and maintained under the protection and favour of Your Honours' own weapons, and that the weapons must be paid for by the profits from the trade; so that we cannot carry on trade without war, nor war without trade" (see Bown, 2010, p. 7).

[21] As important as it was in further stimulating China's labor-intensive industry, many economists wrongly attribute China's rapid industrialization to its successful entrance into the WTO. India and Indonesia, for example, both became members of the WTO in early 1995, six to seven years before China did in late 2001. Yet WTO membership did not trigger an industrial revolution in these two countries. The key difference between China and India or Indonesia was that China had already begun its industrial revolution at the time of its WTO entry, whereas India had not. Hence, WTO membership meant very different things for these countries: It meant a larger export market for mass-produced Chinese goods, but simply more inflow of foreign-produced goods for India and Indonesia.

[22] Compared with wool and other types of natural fibers, cotton is also more easily manipulable for the production of clothing.

[23] More details can be found in Wen (2016a). To illustrate, the United States became the world's textile superpower (superseding Great Britain) around the middle of the 19th century before it became the global manufacturing superpower in the late 19th century; Japan became a textile superpower in the early 20th century before it became a manufacturing superpower around the middle of the 20th century; China became the world's textile superpower in 1995 before it launched its second industrial revolution in heavy industries. These same development steps were taken by France, Germany, South Korea, Taiwan, Hong Kong, and many other economies, all with dramatically different geographic locations, population sizes, and cultural and institutional environments.

[24] This is essentially the view of Friedrich List (1841) as established in *The Natural System of Political Economy*. Even institutions in developed nations (such as the U.S. Department of Energy) maintain tight connections with foreign policy and international politics.

[25] China waited until 1997-98 to start substantially reforming its state-owned enterprises (SOEs); by then, China had essentially already finished its first industrial revolution. Because China's SOEs were located mostly in urban areas and large cities, such a measured development strategy enabled the SOEs to perform at least two important functions in facilitating China's economic transition and industrialization: (i) to maintain and stabilize urban employment during the rural-based proto-industrialization and first industrial revolution; and (ii) to play a leadership role in promoting and transferring more advanced production technologies to rural industries. (China's rural industries received most of their technology and engineers from SOEs in nearby cities.) But once rural industries caught up with SOEs in technology and China broadly finished its first industrial revolution in mass-producing labor-intensive light consumer goods, the historical role of China's small to medium-sized SOEs (which were based on mass production technology to begin with) was finished and naturally yielded to newly formed but more-productive and better-managed private or collective enterprises. During the first 2 years of SOE reform between 1998 and 2000, about 21.4 million SOE workers were laid off, mostly in the textile, mining, military defense, and machinery sectors. However, because of prohibitive costs in finance and technological barriers to form large-scale private heavy industries, China privatized only the small to medium-sized SOEs, which could be easily absorbed or substituted by the private sector. But it kept its large heavy-industrial SOEs under the so-called "grasping the large and letting go of the small" nationwide SOE reform. This by no means implied lack of reform for the remaining large SOEs. The government forced the remaining large heavy-industrial SOEs to reform management structure, upgrade technologies, and confront domestic and international competition. The success of China's high-speed rail companies is a good example of such a measured and targeted SOE reform strategy.

[26] The private patent system has not been as important in the advancement of science and technology as institutional economists have claimed—not even during the British Industrial Revolution (see, e.g., Boldrin and Levine, 2008,

and Mokyr, 2008). In fact, Boldrin and Levin use historical evidence (e.g., the inventor James Watt and his steam engine) to argue that intellectual property rights have hindered innovation rather than stimulated it throughout history.

[27] See, e.g., Lau, Qian, and Roland (2000). Also see the literature's discussions on China's "grasping the large, letting go of the small" reform strategy implemented since 1997 for its heavy industries. For a definition, see http://en.wikipedia.org/wiki/Grasping_the_large,_letting_go_of_the_small.

[28] For example, some of China's military defense companies shifted from manufacturing weaponry and tanks to manufacturing durable consumer goods such as motorcycles and automobiles in the early 1990s. The world's largest producer of high-speed trains used to be a money-losing firm that produced steam engines back in the 1960s under Mao.

[29] Judged by such criteria, China's privatization of small to medium-sized firms such as labor-intensive textile firms was extremely successful, but its market-based reforms in the education and healthcare sectors were disastrous. In retrospect, China should have waited until private hospitals and clinics (or private schools) were well developed and sufficiently competitive with their public counterparts before introducing profit-motivated reforms into these public sectors. Such a waiting period could also allow the government to develop sophisticated regulations in such important welfare-sensitive areas. Hence, as China is currently undergoing its second industrial revolution, it must be extremely careful in taking a measured, dual-tracked, and gradualist approach to financial-sector reforms and privatization of its heavy industries. The danger and risk of a Russian-style collapse under "shock therapy" still exists.

[30] During its first industrial revolution period in 1815-1860, the United States spent $188 million to build canals, 73 percent of which was financed by state and local governments (see Chandler, 1977). In the same period, the territory of the United States expanded enormously, after taking Texas and California from Mexico. Then, after preventing the secession of the cotton-rich Southern states through the Civil War and a long period of government-led railroad expansion, the United States successfully created the largest unified domestic market in the world.

[31] For example, Indian leader Jawaharlal Nehru (in 1946) said, "No country can be politically and economically independent, even within the framework of international interdependence, unless it is highly industrialized and has developed its power resources to the utmost." Chinese leader Mao Zedong (in 1943) said similarly that "Without the establishment of heavy industries in China, there can be no solid national defense, no well-being for the people, no prosperity and strength for the nation." (See Lin, 2009, p. 20.)

[32] In other words, universal suffrage is not the same thing as the rule of law, the rule of law is not the same thing as the market mechanism, and the market mechanism is not the same thing as private property rights. For example, research scientists working for Pfizer (one of the largest U.S. pharmaceutical companies) do not own their intellectual property at all, but they still work very hard to develop new drugs. The United States finished its second industrial revolution during 1880-1940 without universal suffrage.

REFERENCES

Acemoglu, Daron and Robinson, James. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business, 2012.

Agence France-Presse. "China Drives Growth in Patent Applications Worldwide." *Industry Week*, December 16, 2014. http://www.industryweek.com/global-economy/china-drives-growth-patent-applications-worldwide.

Beckert, Sven. *Empire of Cotton: A Global History*. Knopf, 2014.

Blumenthal, David and Hsiao, William. "Privatization and Its Discontents—The Evolving Chinese Health Care System." *New England Journal of Medicine*, 2005, pp. 1165-70; http://dx.doi.org/10.1056/NEJMhpr051133.

Boldrin, Michele and Levine, David K. *Against Intellectual Monopoly*. Cambridge: Cambridge University Press, 2008.

Bown, Stephen R. *Merchant Kings: When Companies Ruled the World, 1600-1900*. Macmillan, 2010.

Bruton, Garry D.; Lan, Hailin and Lu, Yuan. "China's Township and Village Enterprises: Kelon's Competitive Edge." *Academy of Management Executives*, 2000, *14*(1), pp. 19-29; http://dx.doi.org/10.5465/ame.2000.2909834.

Chandler, Alfred D. Jr. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, MA: Harvard University Press, 1977.

Chang, Ha-Joon. *Kicking Away the Ladder: Development Strategies in Historical Perspective*. London: Anthem Press, 2003.

Chen, Kaiji and Wen, Yi. "The Great Housing Boom of China." Federal Reserve Bank of St. Louis Working Paper 2014-022B, 2014 (updated September 2015); https://research.stlouisfed.org/wp/more/2014-022.

Chinese National Committee on Large Dams. "Three Gorges Project." 2010; http://www.chincold.org.cn/dams/rootfiles/2010/07/20/1279253974143251-1279253974145520.pdf.

Cook, Ian G. and Dummer, Trevor J.B. "Changing Health in China: Re-evaluating the Epidemiological Transition Model." *Health Policy*, 2004, *3*, pp. 329-43; http://dx.doi.org/10.1016/j.healthpol.2003.07.005.

Estévez-Abe, Margarita. *Welfare and Capitalism in Postwar Japan: Party, Bureaucracy, and Business*. Cambridge: Cambridge University Press, 2008; http://dx.doi.org/10.1017/CBO9780511510069.

Gerschenkron, Alexander. *Economic Backwardness in Historical Perspective*. Cambridge, MA: Belknap Press of Harvard University Press, 1962.

International Encyclopedia of the Social Sciences. "Putting-Out System," 2008; retrieved from http://www.encyclopedia.com/doc/1G2-3045302138.html.

Jones, William C. and Freeman, Marsha. "Three Gorges Dam: The TVA on the Yangtze River." Schiller Institute: Great Infrastructure Projects, 2001; http://www.schillerinstitute.org/economy/phys_econ/phys_econ_3_gorges.html.

Kriedte, P.; Medick, H. and Schlumbohm, J. *Industrialization before Industrialization: Rural Industry in the Genesis of Capitalism*. Cambridge: Cambridge University Press, 1981, p. 131.

Landes, David S. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: Norton, 1999.

Lau, L.; Qian, Y. and Roland, G. "Reform Without Losers: An Interpretation of China's Dual-Track Approach to Transition." *Journal of Political Economy*, 2000, *108*, pp. 120-43; http://dx.doi.org/10.1086/262113.

Lin, Justin Y. *Economic Development and Transition: Thought, Strategy, and Viability*. Cambridge: Cambridge University Press, 2009; http://dx.doi.org/10.1086/262113.

List, Friedrich. *The Natural System of Political Economy*. London: Longmans, Green, and Co, [1841] 1909.

Lu, Ding. "China's Banking Sector Meeting the WTO Agenda." Discussion Paper No. 5, China Policy Institute, University of Nottingham, April 2006.

Marangos, John. "The Shock Therapy Model of Transition." *International Journal of Economic Policy in Emerging Economies*, 2007, *1*(1), pp. 88-123; http://dx.doi.org/10.1504/IJEPEE.2007.015583.

Marx, Karl. *Capital*. Volume 1. 1867.

Mendels, Franklin F. "Proto-industrialization: The First Phase of the Industrialization Process." *Journal of Economic History*, March 1972, *32*(1), pp. 241-61; http://dx.doi.org/10.1017/S0022050700075495.

Mendels, Franklin F. *Industrialization and Population Pressure in Eighteenth-Century Flanders*. New York: Arno Press, 1981.

Mokyr, Joel. "The Institutional Origins of the Industrial Revolution," in Elhanan Helpman, ed., *Institutions and Economic Performance*. Cambridge, MA: Harvard University Press, 2008, pp. 64-119.

Mufson, Steven. "The Three Gorges Dam." *Washington Post*, November 9, 1997; http://www.washingtonpost.com/wp-srv/inatl/longterm/yangtze/yangtze.htm.

Murphy, Kevin M.; Shleifer, Andrei and Vishny, Robert W. "Industrialization and the Big Push." *Journal of Political Economy*, October 1989, *97*(5), pp. 1003-26; http://dx.doi.org/10.1086/261641.

Oi, Jean C. "Fiscal Reform and the Economic Foundations of Local State Corporatism in China." *World Politics*, 1992, *45*(1), pp. 99-126; http://dx.doi.org/10.2307/2010520.

Oxford Encyclopedia of Economic History. Oxford: Oxford University Press, 2003.

Pomeranz, Kenneth and Topik, Steven. *The World That Trade Created*. Third edition. ME Sharpe, 2013.

Qiu, Larry D. "China's Textile and Clothing Industry." Unpublished manuscript, Hong Kong University of Science and Technology, 2005.

Roll, Eric. *An Early Experiment in Industrial Organization: Being a History of the Firm of Boulton & Watt, 1775-1805*. Frank Cass and Co, 1968.

Rosenstein-Rodan, Paul. "Problems of Industrialization of Eastern and South-Eastern Europe." *Economic Journal*, 1943, *53*(210/211), pp. 202–11; http://dx.doi.org/10.2307/2226317.

Rostow, Walt W. *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge: Cambridge University Press, 1960.

Smith, Adam. *The Wealth of Nations*. London: JM Dent & Sons, 1937 [1776].

Studwell, Joe. *How Asia Works: Success and Failure in the World's Most Dynamic Region*. Grove Press, 2013.

Sullivan, Lawrence R. *Historical Dictionary of the People's Republic of China*. Lanham, MD: Scarecrow Press, 2007.

Vogel, Ezra F. *Deng Xiaoping and the Transformation of China*. Cambridge, MA: Belknap Press / Harvard University Press, 2013.

Wen, Yi. *The Making of an Economic Superpower: Unlocking China's Secret of Rapid Industrialization*. Singapore: World Scientific Publishing, 2016a; http://dx.doi.org/10.1142/9885. [In Chinese: 《伟大的中国工业革命——"发展政治经济学"一般原理批判纲要》，清华大学出版社，2016；https://world.tmall.com/item/535482239543.htm?spm=0.0.0.0.sDNoZi&id=535482239543].

Wen, Yi. "China's Rapid Rise." Federal Reserve Bank of St. Louis *The Regional Economist*, April 2016b, *24*(2), pp. 8-14.

Williams, Eric. *Capitalism and Slavery*. Chapel Hill: University of North Carolina Press, 1994.

Williamson, John. "The Washington Consensus as Policy Prescription for Development." Lecture in the series "Practitioners of Development," delivered at the World Bank, January 13, 2004; https://piie.com/publications/papers/williamson0204.pdf.

World Bank. "Connecting to Compete 2014: Trade Logistics in the Global Economy—The Logistics Performance Index and Its Indicators." Washington, DC: World Bank, 2014; https://openknowledge.worldbank.org/bitstream/handle/10986/20399/904190WP0LPI0R00Box385316B00PUBLIC0.pdf?sequence=1&isAllowed=y.

Zhang, Yi and Zhang, Song-song. "A Short History of China's Village Enterprises." [In Chinese: 张毅、张颂颂编著，《中国乡镇企业简史》，中国农业出版社] 2001.

A Taylor Rule for Public Debt

Costas Azariadis

Public debt is an important source of liquidity in economies facing shortages of private credit. It is also a bubble whose current price depends on expectations of what it will buy at future dates. In this article, the author studies how the government must balance the provision of sufficient liquidity against the risk of adverse expectations regarding future debt prices when private liquidity has dried up. The socially optimal balance is captured in a Taylor-like rule that sets a target for real public debt and manages expectations by overreacting to deviations from the target value. Overreaction takes the form of manipulating budget surpluses to absorb excess debt or reverse liquidity shortages. A budget surplus (deficit) is equivalent to an income tax (subsidy) on investors that restrains (raises) their demand for liquid assets. (JEL H60, E52)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2016, *98*(3), pp. 227-38.
http://dx.doi.org/10.20955/r.2016.227-238

ISSUES AND IDEAS

Classical economists such as Thornton (1802) and Bagehot (1873) understood that one important function of public sector liabilities is the provision of liquidity in times of financial distress.[1] Early research stressed the role of central banks in averting financial panics or catastrophic contractions in private credit driven by pessimistic expectations. Public debt can also help in (i) low-collateral economies with a built-in shortage of liquid assets such as those analyzed in Kiyotaki and Moore (1997) and Bernanke and Gertler (1989) and (ii) environments with weak enforcement such as those in Bulow and Rogoff (1989) and Hellwig and Lorenzoni (2009). The downside of public debt is its fragile or bubble nature, which makes it sensitive to speculative attacks and financial panics. As shown in an extensive literature starting with Tirole (1985) and ending with Kocherlakota (2009), bubbles are assets that add nothing to national income and are socially useful only when investors believe they will maintain their value in the future. Under rational expectations or perfect foresight, bubbles are unstable

Costas Azariadis is the Edward Mallinckrodt Distinguished Professor in Arts and Sciences in the Department of Economics at Washington University in St. Louis and a research fellow at the Federal Reserve Bank of St. Louis. The author thanks David Andolfatto for detailed comments and Minhyeon Jeong for valuable research assistance.

equilibria that can quickly unravel if expectations turn pessimistic in the slightest degree. Keeping bubbles alive requires expectations to be almost permanently optimistic. Thoughtful fiscal policy must then strike a balance between the provision of sufficient liquidity and the feeding of adverse forecasts about the future value of national debt.

How does the government achieve and maintain this balance? This article addresses this question. The setting is a hypothetical economy with no collateral income and no private credit; nothing of substance will change if we look instead at an environment with modest amounts of collateral in which good outcomes are impossible without some form of government intervention. Suppose now that B^* is the socially optimal amount of liquidity that, under ideal conditions, would be provided by private sources. If private debt is substantially or completely illiquid, most or all of the amount B^* must come from the sale of public debt to the private sector, with the government imposing a small tax τ^* to pay interest on B^*.

Investors' expectations of the future value of public debt depend critically on fiscal policy. To reassure investors, the government may commit to raise taxes above τ^* whenever public debt exceeds B^* and to lower them below τ^* in the reverse situation. This commitment helps maintain the value of debt near B^* by promising investors that an oversupply (undersupply) of debt will be aggressively countered by a primary budget surplus (deficit).

Ideally, fiscal policy should seek to maintain the optimum level of liquidity by absorbing excess liquidity immediately and making up liquidity shortages with equal dispatch. A Taylor rule for public debt will achieve these goals if policy overreacts to deviations of public debt from its socially optimal value. Taxes impose an income effect on investors that works like an automatic stabilizer in this framework. When public debt lurches above its target value, investor incomes are hit with higher taxes, which rein in the demand for liquidity. In the reverse case, taxes shrink, disposable incomes rise, and the demand for public debt expands.

To understand the need for fiscal policy to overreact, suppose that investors demand liquidity to raise future consumption at the expense of current consumption. How should fiscal policy manipulate taxes to raise or lower disposable incomes? Any change in income will typically affect both current and future consumption. Reducing future consumption, or the demand for public debt, by one unit typically requires current taxes to rise by *more* than one unit and disposable income to fall by a corresponding amount. It is in that sense that policy should be aggressive. Exactly how much overreaction is socially optimal is the main concern of this article.

THE OPTIMUM AMOUNT OF LIQUIDITY

Private Liquidity with Perfect Financial Markets

To fix notation, let us look at the possibilities for consumption smoothing in an economy with deterministic individual incomes, populated by two groups of agents indexed by $i = 1,2$, each with unit mass. Time is discrete, extends to infinity, and is denoted by $t = 0,1,2,\ldots$. Each agent i has preferences given by

$$\sum_{t=0}^{\infty} \beta^t u\left(c_t^i\right) \tag{1}$$

with $0 < \beta < 1$. The aggregate endowment is constant at two units, but its distribution over agents changes deterministically over time. In particular, individual endowments are periodic[2]—that is,

$$(2) \qquad \left(\omega_t^2, \omega_t^1\right) = \begin{cases} (1+\alpha, 1-\alpha), & \text{if } t = 0,2,\ldots \\ (1-\alpha, 1+\alpha), & \text{if } t = 1,3,\ldots \end{cases}$$

with $\alpha \in (0,1)$.

In a standard dynamic general equilibrium model with abundant collateral and perfect enforcement of loan contracts, an equilibrium is an infinite sequence (c_t^i, b_t^i, q_t) that describes, for each period t, the consumption c_t^i for each agent, their security holdings b_t^i, and the price of loans q_t in terms of the consumption good. This sequence satisfies a consumption Euler equation for each person, two budget constraints, and market clearing. These equations are

$$(3) \qquad q_t u'\left(c_t^i\right) = \beta u'\left(c_{t+1}^i\right)$$

$$(4) \qquad c_t^i + q_t b_{t+1}^i = \omega_t^i + b_t^i$$

$$(5) \qquad b_t^1 + b_t^2 = 0; \ \left(b_0^1, b_0^2\right) \text{ given}$$

for all (i,t), where q_t is the price of a one-period private security paying off one unit of consumption at $t+1$, and b_{t+1}^i is the number of those securities purchased by household i at time t.

This setting of perfect financial markets with full commitment to repay loans has a unique competitive equilibrium with perfect consumption smoothing for every $t = 0,1,\ldots$:

$$(6a) \qquad \left(c_t^1, c_t^2\right) = \left(c_1^*, c_2^*\right)$$

supported by a bond price

$$(6b) \qquad q_t = \beta \quad \forall t,$$

which corresponds to equality between the rate of return on private debt and the rate of time preference. Consumption flows (c_1^*, c_2^*) satisfy household budget constraints, equating the present value of consumption with household wealth. Specifically,

$$(7a) \qquad \frac{c_1^*}{1-\beta} = \frac{1+\alpha+\beta(1-\alpha)}{1-\beta^2} + b_0^1$$

$$(7b) \qquad \frac{c_2^*}{1-\beta} = \frac{1-\alpha+\beta(1+\alpha)}{1-\beta^2} + b_0^2.$$

As one might expect, this equilibrium is Pareto optimal and aggregate consumption equals aggregate income—that is,

(8) $$c_1^* + c_2^* = 2$$

because $b_0^1 + b_0^2 = 0$. The allocation of consumption across households depends on the initial distribution of (b_0^1, b_0^2) of financial wealth. A particularly interesting initial distribution is

(9) $$\left(b_0^1, b_0^2\right) = \left(\frac{\alpha}{1+\beta}, -\frac{\alpha}{1+\beta}\right),$$

which gives households with low initial income, $1 - \alpha$, a claim against those with high initial income, $1 + \alpha$. In this case, the unique equilibrium is symmetric with

(10a) $$c_1^* = c_2^* = 1$$

and private debt holdings that alternate between $\alpha/(1 + \beta)$ and $-\alpha/(1 + \beta)$. Specifically,

(10b) $$\left(b_t^1, b_t^2\right) = \begin{cases} \left(\frac{\alpha}{1+\beta}, -\frac{\alpha}{1+\beta}\right) & \text{if } t = 0, 2, \ldots \\ \left(-\frac{\alpha}{1+\beta}, \frac{\alpha}{1+\beta}\right) & \text{if } t = 1, 3, \ldots \end{cases}.$$

Commitment to repay debts is essential in achieving this allocation of resources. If borrowers have little or no collateralizable income, financial markets will work poorly and households may not be able to receive the amount of liquidity needed to support perfect consumption smoothing. In the following text, I briefly review two types of financial market imperfections and then concentrate on the extreme case of financial distress that occurs when private credit dries up completely.

Private Liquidity with Imperfect Financial Markets

Before considering public debt, I briefly discuss what could go wrong with private debt in an environment with financial frictions. A particularly useful friction to consider is limited enforceability of loan contracts, which restricts the amount each household can borrow either by the pledgeable collateral owned by that household or by how much the borrower values a good reputation that permits unfettered access to future credit. Kiyotaki and Moore (1997), for example, analyze credit market imperfections by adding to the household problem of the previous subsection a collateral constraint of the form

(11) $$b_t^i + \lambda_t \omega_t^i \geq 0,$$

where $\lambda_t \in [0,1]$ is an exogenous leverage ratio describing the fraction of a borrower's income that lenders can claim in the event of default.

Kehoe and Levine (1993) propose to make the leverage ratio λ_t endogenous by connecting access to future credit with the reputation of each borrower. In this setting, enforcement

works through reputation. Borrowers who default are shut out of credit markets and forced into permanent autarky—they cannot borrow or lend ever again. Solvency is maintained by credit lines that motivate borrowers to repay and continue trading in the credit market instead of default and autarky.

Self-enforcing loans of that type satisfy an incentive constraint of the form

$$\sum_{s=0}^{\infty} \beta^s \left[u\left(c_{t+s}^i\right) - u\left(\omega_{t+s}^i\right) \right] \geq 0 \tag{12}$$

for each i and t. This constraint keeps the value of solvency at or above the value of default for everyone at all times. It holds with equality for rationed borrowers. Alvarez and Jermann (2000) show how to find leverage ratios that connect equation (12) to equation (11).

It is fairly easy to guess what happens in an economy with debt constraints of the type described in equation (11). If the leverage ratio λ_t is always sufficiently large, the symmetric equilibrium of the previous subsection still goes through. This happens when

$$\lambda_t (1+\alpha) \geq \alpha/(1+\beta) \tag{13}$$

for all $t \geq 0$ in equation (11), and when

$$\frac{u(1)}{1-\beta} \geq \frac{u(1+\alpha)+\beta u(1-\alpha)}{1-\beta^2} \tag{14}$$

in equation (12). Adverse collateral shocks will violate equation (13) and reduce the amount of consumption smoothing achieved in equilibrium. As the leverage ratio shrinks in equation (11), equilibrium allocations will approach autarky—that is,

$$c_t^i \to \omega_t^i \text{ as } \lambda_t \to 0.$$

The same problem crops up in the Kehoe-Levine economy when condition (14) is violated. When that happens, endogenous leverage ratios λ_t tend to depend on expectations of future leverage ratios ($\lambda_{t+1}, \lambda_{t+2}, \ldots$). If future values of λ are small, then the credit market does not help smooth future consumption, which reduces the value of a good reputation to the borrower and makes lenders unwilling to lend now. Pushing this syllogism to the end, autarky turns out to be a socially undesirable but stable equilibrium in which pessimistic expectations of credit panics are self-fulfilling.

More sanguine forecasts of future credit conditions lead to better equilibria, but consumption smoothing will typically be limited by borrowing constraints. One way to avoid catastrophic reductions in private credit is for the central bank to act as a lender of last resort.[3] Significant reductions in non-collateral credit did actually happen in the last quarter of 2008 and the first quarter of 2009 when the stock of commercial paper in circulation dropped by more than 50 percent. How should the government react when that occurs?

The Role of Public Debt

One reason public debt is more tradable than private debt is that the fiscal authority has certain powers of enforcing claims on households that the private sector does not possess. Suppose, in particular, that the government can extract a modest tax

$$\tau_t \leq \overline{\tau}$$

from every high-income individual and use the proceeds to finance transfers to low-income individuals or to repay those it borrowed from in the past. The government budget constraint is then

$$\tau_t = \begin{cases} B_t - q_t B_{t+1} & \text{if } q_t > 0 \\ 0 & \text{if } q_t = 0 \end{cases}, \tag{15}$$

where B_{t+1} is the real value of public debt issued at t and maturing at $t+1$. Taxes amount to a primary budget surplus that pays interest on existing public debt.

The rest of the economy functions in the way described by equations (3), (4), (5), and (15), except for the budget constraint (4) and the market-clearing condition (5), which now reflect the payment of taxes and the existence of a financial asset (public debt) in positive net supply. These equations become

$$c_t^i + q_t B_{t+1}^i = \omega_t^i + B_t^i - \tau_t^i \text{ and} \tag{4'}$$

$$B_t^1 + B_t^2 = B_t. \tag{5'}$$

Here $B_t^i \geq 0$ denotes claims on the government and

$$\tau_t^i = \begin{cases} \tau_t & \text{if } \omega_t^i = 1+\alpha \\ 0 & \text{otherwise} \end{cases}.$$

It is a simple matter to verify from the budget constraint (15), from equation (4′), and from the market-clearing condition (5′) that the policy choice

$$\tau_t = \tau^* := \alpha(1-\beta) \quad \forall t \tag{16}$$

will support a steady-state equilibrium with constant public debt B^*, bond price q^*, and constant consumption c^* for each i, where the high-income person buys the entire stock of public debt from the low-income person, and

$$\left(B^*, q^*, c^*\right) = (\alpha, \beta, 1).$$

In other words, the fiscal authority can stand in for the malfunctioning private credit market and provide the optimum amount of liquidity if high-income households expect that

the price of public debt will remain forever constant at its maximal value β. To understand the circumstances that justify this level of optimism, we need to delve into the dynamics of public debt and of asset bubbles.

MANAGING BUBBLES

Can the government raise the odds that markets will trust the price of public debt to remain at the value β that is needed to support an optimum allocation of resources in the private sector? As a start, we endow the fiscal authority with a policy rule that connects taxes (that is, the primary budget surplus) with the stock of maturing debt. The general form of this rule is

(17) $$\tau_t = T(B_t),$$

where the function T maps the public debt interval $[0,1+\alpha]$ into the feasible tax interval $[0,\bar{\tau}]$. Assuming that high-income households purchase the entire stock of public debt, equilibria again satisfy two budget constraints (one for the high-income household and another for the low-income household):

(18a) $$c_t^H = 1+\alpha-\tau_t - q_t B_{t+1},$$

(18b) $$c_{t+1}^L = 1-\alpha+B_{t+1};$$

one first-order condition for the high-income household (the other household chooses a corner solution):

(19) $$q_t u'(c_t^H) = \beta u'(c_{t+1}^L);$$

and the government budget constraint:

$$\tau_t + q_t B_{t+1} = B_t.$$

Combining equations (17) through (19) gives a simple first-order law of motion for the debt sequence (B_t), which is

(20) $$\left[B_t - T(B_t)\right]u'(1+\alpha-B_t) = \beta B_{t+1} u'(1-\alpha+B_{t+1}).$$

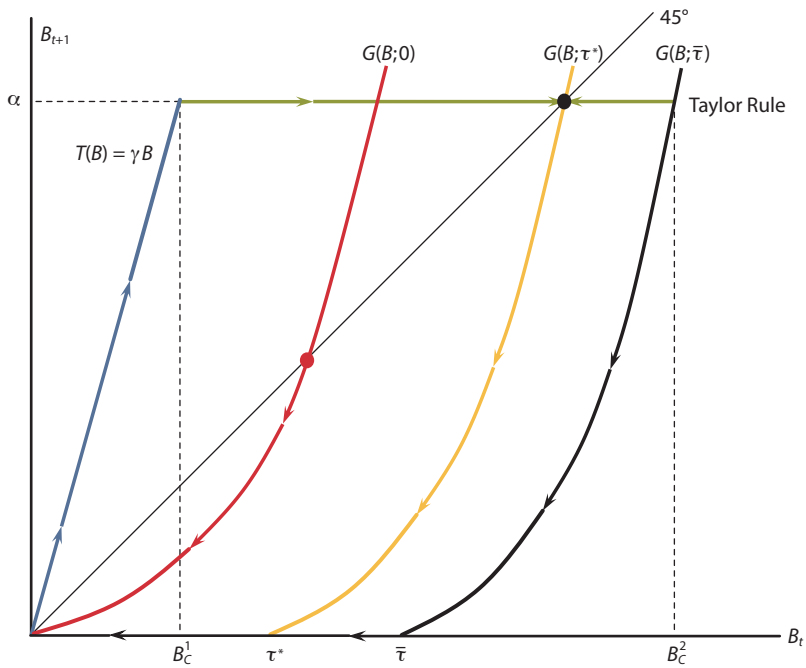
If the household's intertemporal elasticity of substitution (IES) is not too low, we can solve equation (20) explicitly for B_{t+1} to obtain

(21) $$B_{t+1} = G(B_t;T).$$

Equation (21) says that the properties of the law of motion G depend on the shape of the policy rule T. Fiscal policy can influence expectations by choosing the "right" policy rule.

Figure 1

Bubble Management



For each policy rule $T(B)$, an equilibrium is a nonnegative solution sequence (B_t) to equation (20). Figure 1 describes some of the more interesting equilibria. To compare those with a no-bubble or no-liquidity outcome, we define the autarky yield

$$\bar{R} = \frac{u'(1+\alpha)}{\beta u'(1-\alpha)}, \tag{22}$$

which is the reciprocal of the implied security price $\bar{q}$ when all credit markets are closed.

Passive Policies

Passive policies keep taxes constant for any value of public debt. Figure 1 shows how these policies fail. Keeping the tax $T(B)$ constant anywhere in the interval $[0,\bar{\tau}]$, even if we choose the optimal value $\tau^* = \alpha(1 - \beta)$, leaves the value B of debt indeterminate in the interval $[0,\alpha]$ and thus open to shocks in expectations. The price of debt is also indeterminate in the interval $[\beta,1/\bar{R}]$. Equilibria are indeterminate under the red law of motion in Figure 1, the black law of motion in the same figure, and anything in between.

The red law describes the bubble dynamics of an overlapping-generations economy analyzed by Tirole (1985) and many others. It has the shape of a "reflected" offer curve—that is, the mirror image, relative to the vertical axis, of an offer curve for a two-period-lived house-

hold with life cycle endowment $(1 + \alpha, 1 - \alpha)$. Offer curves are monotone if the IES is high enough to make current and future consumption gross substitutes, but can bend backward for low IES values. Bubble equilibria of this type exist if the no-bubble allocation is dynamically inefficient—that is, if

$$\bar{R} < 1.$$

It is worthwhile to reiterate here the argument by Farhi and Tirole (2012, p. 692) that "dynamic inefficiency is a sufficient but not necessary condition for the possibility of bubbles." Figure 1 clearly shows that public debt can command a positive price in a dynamically efficient but illiquid economy, as shown by the black and yellow laws of motion depicted in Figure 1. Under those laws, there is private demand for public debt, but it remains fragile and sensitive to adverse expectations.

The Optimal Taylor Rule

How do we keep the price of debt unchanged at the value $q = \beta$, which corresponds to the socially desirable allocation,

$$\left(c_t^H, c_t^L\right) = (1,1),$$

for all time? Equation (20) suggests a tax policy that will deliver tomorrow the optimum amount of liquidity $B^* = \alpha$ for any value B_t today that is sufficiently close to the optimum liquidity B^*.

The literature on equilibrium selection[4] provides some guidance on how to proceed. The key idea is to put $B_{t+1} = \alpha$ in equation (20) and solve for $T(B)$—that is, to choose the policy rule:

$$T(B) = B - \frac{\alpha\beta u'(1)}{u'(1+\alpha-B)}. \tag{23}$$

Under this policy, the law of motion G becomes the horizontal green line in Figure 1. For example, in the neighborhood of $(T(B^*), B^*) = (\alpha(1-\beta), \alpha)$, equation (23) takes the linear form of Taylor rules:

$$T(B) - \tau^* = (1+\lambda)\left(B - B^*\right), \tag{24}$$

where

$$\lambda := \alpha\beta\gamma \tag{25a}$$

and

$$\gamma := -u''(1)/u'(1). \tag{25b}$$

The parameter λ depends on the amplitude of income shocks and on the reciprocal of the IES.

More generally, the nonlinear policy rule in equation (23) always keeps the public debt at the right value by raising taxes whenever public debt exceeds the target value α and lowering taxes in the reverse event. Liquidity is maintained at its socially desirable value B^* by manipulating primary budget surpluses or, equivalently, taxes on buyers of public debt. These taxes exert an income effect on the private demand for public debt. When debt exceeds B^*, for example, taxes increase more than proportionally to contain the burgeoning demand for liquidity and future consumption. Overreactions to public debt movements are particularly appropriate when the income effect is small because investors are wealthy (high value of the parameter α) or when the propensity to save out of income is relatively insensitive (large value of β or low IES).

Any change in taxes will also feed into the bond price q and the interest rate r, which satisfies

$$1/q = 1 + r. \tag{26}$$

In particular, tax increases on high-income investors at time t will raise their income growth from t to $t+1$, reduce liquidity demand, and exert upward pressure on yields. Policy rules such as equation (23) then can be expressed in terms of real yields if we rewrite equation (20) in the form

$$1 + r_t = u'(1 + \alpha - B_t)/[\alpha\beta u'(1)], \tag{27}$$

with B_{t+1} set to equal its optimum value α. Expanding (r_t, B_t) about the optimal value $(1/\beta - 1, \alpha)$ we obtain a linearized interest rate rule that is completely equivalent to equation (24)—that is,

$$\frac{r_t - r^*}{r^*} = \frac{\gamma}{1-\beta}\left(B_t - B^*\right). \tag{28}$$

The government raises yields to curb excessive demand for liquidity and lowers yields to prop up liquidity.

Since taxes cannot exceed the modest upper bound $\bar{\tau}$, we need to know the range over which the linear approximation (28) is really useful. Clearly, the Taylor rule applies if $T(B)$ is in the interval $[0, \bar{\tau}]$—that is, whenever B lies in the interval $[B_C^1, B_C^2]$ around the target value B^*, where

$$B_C^1 := B^* - \frac{\tau^*}{1+\lambda} < B_C^2 := B^* + \frac{\bar{\tau} - \tau^*}{1+\lambda}.$$

If public debt falls below B_C^1, a linear tax rule with a sufficiently strong reaction coefficient will prevent the bubble from bursting, as shown in the blue law of motion in Figure 1. When the economy is dynamically inefficient, this policy amounts to a negative tax on investors—that is, on people with *temporarily* high incomes (type-2 households in periods $t = 0,2,\ldots$ and type-1 households in periods $t = 1,3,\ldots$). This subsidy will boost investors' flagging demand for assets and help achieve an optimum allocation of resources. As shown earlier, dynamic

efficiency is a sufficient but not necessary motive for the holding of public debt. Illiquidity turns out to be a sufficiently plausible motive for policy intervention even in dynamically efficient economies.

CONCLUSION

This article reviews what we know about public debt management in economies facing severe liquidity shortages. In these circumstances, public liabilities are a substitute for private ones, and the demand for public debt has little or nothing to do with the dynamic inefficiency conditions identified in Tirole (1985): low interest rates relative to growth rates.

Even as it provides needed liquidity, public debt remains a bubble whose price is sensitive to forecasts of its future value. Thoughtful fiscal policy in these circumstances must balance the liquidity needs of the private sector against adverse expectations of devalued debt. A good way to strike this balance is to tie private sector demand for new debt with the current value of maturing debt. The necessary link is provided through a Taylor rule for public debt that acts as an automatic stabilizer on investor demand for public debt. It raises taxes on investors whenever debt exceeds the socially optimal amount of liquidity and lowers taxes in the reverse situation. The extent to which taxes should overreact to debt depends on structural parameters, notably the amplitude of income fluctuations and the intertemporal elasticity of substitution.

It is worth noting that fiscal policy is most successful in providing the optimum amount of liquidity when public debt and private debt are perfect substitutes, as we have assumed throughout this article. It would be interesting to see how this rule would work in an environment with uninsurable idiosyncratic income uncertainty in which public debt must replace not just one missing credit market but as many markets as there are idiosyncratic income states.

A related and perhaps weightier issue is averting financial distress in the first place. Liquidity provision seems to be the flip side of last-resort lending.[5] It seems natural to design policies that would pursue both goals simultaneously. ■

NOTES

1 A typical modern example is Farhi and Tirole (2012), who analyze how "outside liquidity" complements private liquidity for financially constrained firms. Non-liquidity aspects of public debt include tax smoothing (Barro, 1979) and improving intertemporal allocations in life cycle economies (Auerbach and Kotlikoff, 1987).

2 This simple deterministic endowment process is the degenerate case of a stochastic economy with two Markovian states and a zero probability of remaining in the same state. Markovian endowments with two states are a straightforward extension.

3 Martin (2006), Ennis and Keister (2010), and Humphrey (2010) discuss policies that prevent large reductions in privately issued liquidity.

4 See Antinolfi, Azariadis, and Bullard (2007) and Pintus (2008) for recent examples of this procedure applied to monetary policy. Earlier examples include Grandmont (1986) and Woodford (1986).

5 See Bordo (1989) for an insightful historical review of last-resort lending.

REFERENCES

Alvarez, Fernando and Jermann, Urban J. "Efficiency, Equilibrium, and Asset Pricing with Risk of Default." *Econometrica*, July 2000, *68*(4), pp. 775-97; http://dx.doi.org/10.1111/1468-0262.00137.

Antinolfi, Gaetano; Azariadis, Costas and Bullard, James. "Monetary Policy as Equilibrium Selection." Federal Reserve Bank of St. Louis *Review*, July/August 2007, *89*(4), pp. 331-41; https://research.stlouisfed.org/publications/review/07/07/Antinolfi.pdf.

Auerbach, Alan J. and Kotlikoff, Laurence J. *Dynamic Fiscal Policy*. Cambridge, MA: Cambridge University Press, 1987.

Bagehot, Walter. *Lombard Street: A Description of the Money Market*. London: Henry S. King, 1873.

Barro, Robert J. "On the Determination of the Public Debt." *Journal of Political Economy*, October 1979, 87(5 Part 1), pp. 940-71; http://dx.doi.org/10.1086/260807.

Bernanke, Ben S. and Gertler, Mark. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review*, March 1989, *79*(1), pp. 14-31.

Bordo, Michael D. "The Lender of Last Resort: Some Historical Insights." NBER Working Paper No. 3011, National Bureau of Economic Research, June 1989; http://www.nber.org/papers/w3011.pdf.

Bulow, Jeremy and Rogoff, Kenneth. "Sovereign Debt: Is to Forgive to Forget?" *American Economic Review*, March 1989, *79*(1), pp. 43-50.

Ennis, Huberto M. and Keister, Todd. "Banking Panics and Policy Responses." *Journal of Monetary Economics*, May 2010, *57*(4), pp. 404-19; http://dx.doi.org/10.1016/j.jmoneco.2010.04.005.

Farhi, Emmanuel and Tirole, Jean. "Bubbly Liquidity." *Review of Economic Studies*, April 2012, *79*(2), pp. 678-706; http://dx.doi.org/10.1093/restud/rdr039.

Grandmont, Jean-Michel. "Stabilizing Competitive Business Cycles." *Journal of Economic Theory*, October 1986, *40*(1), pp. 57-76; http://dx.doi.org/10.1016/0022-0531(86)90007-4.

Hellwig, Christian and Lorenzoni, Guido. "Bubbles and Self-Enforcing Debt." *Econometrica*, July 2009, *77*(4), pp. 1137-64; http://dx.doi.org/10.3982/ECTA6754.

Humphrey, Thomas M. "Lender of Last Resort: What It Is, Whence It Came, and Why the Fed Isn't It." *Cato Journal*, Spring/Summer 2010, *30*(2), pp. 333-64.

Kehoe, Timothy J. and Levine, David K. "Debt-Constrained Asset Markets." *Review of Economic Studies*, 1993, *60*(4), pp. 865-88; http://dx.doi.org/10.2307/2298103.

Kiyotaki, Nobuhiro and Moore, John. "Credit Cycles," *Journal of Political Economy*, April 1997, *105*(2), pp. 211-48; http://dx.doi.org/10.1086/262072.

Kocherlakota, Narayana R. "Bursting Bubbles: Consequences and Cures." Presented at the Macroeconomic and Policy Challenges Following Financial Meltdowns Conference, International Monetary Fund, Washington, DC, April 3, 2009; https://www.imf.org/external/np/seminars/eng/2009/macro/pdf/nk.pdf.

Martin, Antoine. "Liquidity Provision vs. Deposit Insurance: Preventing Bank Panics without Moral Hazard." *Economic Theory*, May 2006, *28*(1), pp. 197-211; http://dx.doi.org/10.1007/s00199-005-0613-x.

Pintus, Patrick A. "Laffer Traps and Monetary Policy." Federal Reserve Bank of St. Louis *Review*, May/June 2008, *90*(3 Part 1), pp. 165-74; https://research.stlouisfed.org/publications/review/08/05/Pintus.pdf.

Thornton, Henry. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*. London: J. Hatchard, 1802.

Tirole, Jean. "Asset Bubbles and Overlapping Generations." *Econometrica*, November 1985, *53*(5), pp. 1499-528; http://dx.doi.org/10.2307/1913232.

Woodford, Michael. "Stationary Sunspot Equilibria in a Finance Constrained Economy." *Journal of Economic Theory*, October 1986, *40*(1), pp. 128-37; http://dx.doi.org/10.1016/0022-0531(86)90011-6.

Monetary Policy in an Oil-Exporting Economy

Franz Hamann, Jesús Bejarano, Diego Rodríguez, and Paulina Restrepo-Echavarría

The sudden collapse of oil prices poses a challenge to inflation-targeting central banks in oil-exporting economies. In this article, the authors illustrate this challenge and conduct a quantitative assessment of the impact of changes in oil prices in a small open economy in which oil represents an important fraction of its exports. They build a monetary, three-sector, dynamic stochastic general equilibrium model and estimate it for the Colombian economy. They model the oil sector as an optimal resource extracting problem and show that in oil-exporting economies the macroeconomic effects vary according to the degree of persistence of oil price shocks. The main channels through which these shocks pass to the economy come from the real exchange rate, the country risk premium, and sluggish price adjustments. Inflation-targeting central banks in such economies face a policy dilemma: raise the policy rate to fight increased inflation coming from the exchange rate passthrough or lower it to stimulate a slowing economy. (JEL C61, E31, E37, E52, F41)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2016, *98*(3), pp. 239-61.
http://dx.doi.org/10.20955/r.2016.239-261

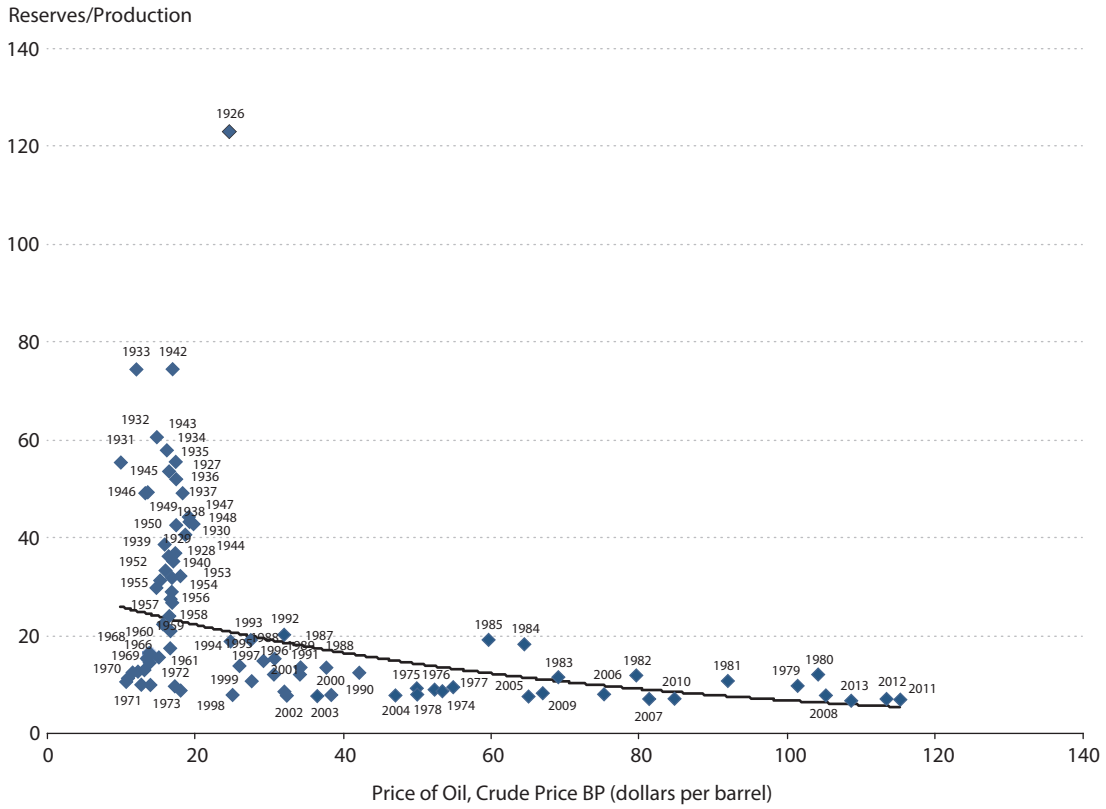
One global event shaped the economic outcomes during 2014 and 2015: the sudden collapse of world oil prices. This event has been a source of instability in global financial markets, especially in emerging economies such as Russia, Brazil, Venezuela, Ecuador, and Colombia. This sudden collapse in oil prices has caught the attention of policymakers and academics as the macroeconomic consequences may be significant.

An analysis of the implications of such a collapse for monetary policy in small oil-exporting economies is needed for several reasons. First, the oil price shock is large and to some extent occurred earlier than expected. Oil prices increased steadily after 2009 from $35 (U.S. dollars) per barrel to levels surpassing $100 per barrel. In the last quarter of 2014, oil prices fell by 38 percent and country risk spreads and interest rates in oil-exporting economies jumped.

Franz Hamann is director and Jesús Bejarano and Diego Rodríguez are chiefs in the macroeconomic modeling division at Banco de la República de Colombia. Paulina Restrepo-Echavarría is an economist at the Federal Reserve Bank of St. Louis. The authors are grateful to Enrique Mendoza, Gianluca Benigno, and Hernando Vargas for sharing their insights. The authors also thank the participants at the Closing Conference of the BIS CCA Research Network on "Incorporating Financial Stability Considerations Into Central Bank Policy Models" at the Bank of International Settlements in Mexico in January 2015; Martin Uribe for his comments to an earlier version of this work; and Paula Beltrán, Norberto Rodríguez, Rafael Hernández, and Joao Hernández for assistance. Maria A. Arias provided research assistance.

Figure 1

International Oil Prices and the Oil Reserves-to-Production Ratio: Colombia



SOURCE: From Hamann, Bejarano, and Rodríguez (2015, Figure 2, p. 3).

Second, oil production in some countries is a significant portion of gross domestic product (GDP). For instance, Colombian oil production over the past 10 years increased from 5 percent of GDP to 11 percent in 2014; the share of oil exports in GDP jumped from 3 percent in 2002 to 8 percent in 2014. In turn, fiscal revenues from oil (as a share of total public revenues) increased from under 10 percent in 2002 to close to 20 percent in 2011. Foreign direct investment (FDI) in the oil sector represented 32 percent (as a share of the total FDI in Colombia), while FDI in mining represented 17 percent in 2014. Similar patterns emerge for other oil-exporting economies.

Third, persistent swings in oil prices do impact oil activity in Colombia. Figure 1 shows the linkage between international oil prices and the ratio of oil reserves to production in Colombia. The data for the figure support the idea that as prices increase, producers extract oil from the ground and reserves fall, *ceteris paribus*. On the contrary, when prices are low, there are fewer incentives for producers to extract oil.

Oil price shocks are also related to country risk spreads,[1] capital flows, and other macroeconomic indicators at the business cycle frequency. Periods of high commodity prices have

been associated with lower spreads, capital inflows, and good macro performance, whereas the opposite is associated with periods of low prices. González, Hamann, and Rodríguez (2015) have documented some empirical regularities around *transitory* oil price shocks in Colombia.

These facts are consistent with the intuition shared by many economists who study small open economies in which resource sectors are important. Higher oil prices increase oil revenues but compress the risk premium, thereby (i) improving overall creditworthiness, (ii) creating a surge in demand for tradable and non-tradable goods, and (iii) inducing both a real exchange-rate appreciation and a shift of economic resources from the tradable sector to the non-tradable sector. Credit expands, especially in those sectors boosted by the real appreciation. Overall economic activity and demand booms move in tandem with asset prices. However, sharp oil price reversals truncate this process; resources are reallocated and asset prices and the currency collapse.

It is possible that current oil prices will remain low not just for a few quarters but for several years to come. Long-lasting changes in global conditions pose a different challenge for central banks in small open and commodity-dependent economies. Permanent changes in oil prices reduce permanent income, affect aggregate consumption and savings decisions, and have implications for resource allocations between tradable and non-tradable sectors. Resource allocations show up in the real exchange rate, wages, and the country's long-term net foreign asset position. Monetary policy is usually set to reach goals at a one- to two-year horizon. Long-term changes from lower oil prices may have different macroeconomic consequences than temporary shocks, as stressed by Rebucci and Spatafora (2006), Kilian (2009), and Kilian, Rebucci, and Spatafora (2009).

Still, nominal adjustment of the exchange rate may continue to be important because a flexible nominal exchange rate may partially compensate for the fall in oil prices. The importance of the role of nominal stickiness in small open economy models has been emphasized by Galí and Monacelli (2005); De Paoli (2009); Benigno and De Paoli (2010); Auray, de Blas, and Eyquem (2011); Gertler and Karadi (2011); and Schmitt-Grohé and Uribe (2013), to name a few. In the presence of nominal price and/or wage rigidities, the quantities of oil produced will likely further accommodate the adjustment. For instance, gasoline and other oil derivatives are key inputs of production; should these inputs become relatively cheaper, they could ease marginal cost pressure on firms and inflation. Finally, pass through from such shocks to inflation and inflation expectations may trigger a monetary policy response, which in the presence of nominal rigidities feeds back into economic activity.

In this article, we conduct a quantitative assessment of the impact of permanent oil price changes in a small open economy in which a commodity, such as oil, represents an important share of the economy. Our analysis takes into account the central bank's policy response to such changes. Here we use a highly persistent but transitory shock as a proxy for a permanent shock. In a more general setup, presented in Hamann, Bejarano, and Rodríguez (2015), we conducted the same experiment using a permanent shock that changes the long-run net foreign asset position of the economy; the results in that article are similar to those presented here.

To understand the basic mechanisms at work, we set up a monetary policy model with three sectors: non-tradable, tradable, and oil. The non-tradable sector uses labor and an

imported intermediate good (i.e., gasoline) in the production of a final good; this sector also has monopolistic competition and sticky prices. The tradable sector is modeled as an endowment, and oil is a fully exportable output whose production is endogenous and responds to economic incentives. We model the oil sector as a resource extraction problem as in Sickles and Hartley (2001) and Pesaran (1990). The economy owns a stock of oil and extracts the optimal portion of it to sell in international competitive commodity markets. Thus, optimal extraction rules depend on the stock of oil reserves, oil prices, interest rates, the marginal costs of oil operation, and the uncertain nature of oil discoveries.

We close the nominal portion of the model assuming a total inflation-targeting central bank. Our quantitative analysis indicates that this central bank is confronted with a policy dilemma: The permanent fall in oil revenues causes a permanent fall in consumption and GDP, but the nominal depreciation drives total inflation off target, causing the bank to tighten its policy stance. We also show, however, that this dilemma arises because the tradable sector features flexible prices,[2] whereas in the non-tradable sector prices are sticky. Therefore, the dilemma disappears if the central bank is able to identify exactly where the nominal rigidities reside (i.e., the non-tradable sector) and targets non-tradable inflation.

Both the nominal and the real exchange-rate adjustments are at the core of the adjustment mechanism since this plunge in oil prices incentivizes oil firms to cut extraction and increase oil reserves, which in turn reduces the availability of tradable goods in the economy. Reduced availability causes excess demand for tradable goods; this demand is adjusted through an increase in the relative price of tradable goods to non-tradable goods.

Also, at the core of the adjustment mechanism lies the external interest rate the economy faces in international financial markets. The model predicts a protracted period of higher external interest rates because of the higher risk premium caused by lower oil prices, which is in contrast to lower risk when oil reserves are high and accumulated endogenously. The interaction of these real adjustments with nominal rigidities is interesting because the model delivers nominal exchange-rate depreciation, which passes through to total inflation. This pass through is significant. It temporally but persistently raises annual inflation well above target, causing the model's total inflation-targeting central bank to tighten monetary policy to control inflation.

The rest of the article proceeds as follows. In the next section we present a monetary policy model for an oil-exporting economy and evaluate its quantitative predictions under both a permanent and a transitory oil price shock. In our concluding section we examine the implications of our framework for monetary policy.

AN OIL-EXPORTING, SMALL OPEN MONETARY ECONOMY

Structure of the Model

The model is a three-sector (oil, tradable, and non-tradable sectors) small open economy with an incomplete foreign financial assets market populated by households, producers, and the central bank. Households (i) supply labor to firms and consume final goods, (ii) save in

the form of foreign debt, and (iii) receive the revenues from the oil sector, which decides how to extract oil optimally. The tradable output is an endowment, but non-tradable output is produced in several stages in a monopolistic competitive environment with nominal rigidities. In addition, non-tradable output production needs an imported input of production (i.e., gasoline).

Households

More formally, there is a representative household that maximizes the expected discounted utility,

$$E_0\left[\sum_{t=0}^{\infty}\beta^t \frac{\left[c_t - \frac{h_t^{\omega}}{\omega}\right]^{1-\sigma}}{1-\sigma}\right],$$

subject to

$$c_t + q_t b_t^* \left(1+r_t^*\right) + Q_{t,t+1} b_{t+1} \leq w_t h_t + q_t C(s,x) + \xi_t^N + p_t^T y^T + q_t \xi_t^X + q_t b_{t+1}^* + b_t,$$

where c_t is the consumption basket; h_t are the hours worked; w_t is the real wage; $C(s,x)$ is the revenue for supplying drilling and oil field services to the oil firm[3]; b_t^* is the real external debt expressed in terms of the foreign consumption basket; b_t is a real state-contingent domestic bond; q_t is the real exchange rate; $Q_{t,t+1}$ is the real price of the domestic bond; ξ_t^N are the profits for the non-tradable goods producers; y^T is a constant stream of income of an endowment of tradable goods (which can be consumed or exported)[4]; $p_t^T = q_t p_t^{T,*}$ with $p_t^{T,*}$ following an AR(1) process (described in the appendix); ξ_t^X are the profits from the oil firms; and r_t^* is the real interest rate this economy faces in international financial markets.

We model the external real interest rate as having two components: (i) the risk-free real interest rate and (ii) a risk component that we assume is a positive function of the deviations of the external debt-to-oil reserves ratio from its steady-state value. That is,

$$r_t^* = r_t^f + \psi\left[\exp\left(\frac{q_t b_t^*}{p_t^x s_t} - \frac{q b^*}{p^x s}\right) - 1\right],$$

where $\psi > 0$ is a parameter that determines the elasticity of the risk component to deviations of the debt-to-oil reserves ratio from its steady state, s_t is the stock of oil reserves, and r_t^f represents the risk-free real interest rate. This reduced form is similar to that presented by Neumeyer and Perri (2005). As in Neumeyer and Perry (2005), this relation is introduced not to provide a satisfactory model of country risk, but rather to show that country risk may also depend on both internal and external conditions such as the domestic debt and the international oil price, respectively. The inclusion of the oil price in this specification may amplify the effects of an oil price shock beyond the usual income effect displayed in this family of models.

Assuming that the law of one price holds for oil,

$$p_t^x = q_t p_t^{x,*},$$

where p_t^x is the real price of oil and $p_t^{x,*}$ is the real price of oil in terms of a foreign consumer price index (CPI).

To simplify the (paper and pencil) calculation of the deterministic steady state of this model, we depart from the constant elasticity of substitution specification of consumption and assume that the consumption goods basket for the representative household is a Cobb-Douglas compound of tradable and non-tradable goods as follows:

$$c_t = \left(c_t^N\right)^{\gamma}\left(c_t^T\right)^{1-\gamma},$$

where c_t^T is the consumption of tradable goods and c_t^N is the basket of differentiated non-tradable goods, which is represented by a Dixit-Stiglitz aggregator:

$$c_t^N = \left[\int_0^1 c_t^N(j)^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}}.$$

Under these assumptions, the optimal household choices of consumption, hours worked, domestic bonds, and external debt are

$$\left[c_t - \frac{h_t^{\omega}}{\omega}\right]^{-\sigma} = \lambda_t$$

$$h_t^{\omega-1} = w_t$$

$$\beta E_t \lambda_{t+1} = Q_{t,t+1}\lambda_t$$

$$q_t \lambda_t = \beta E_t q_{t+1}\left(1+r_{t+1}^*\right)\lambda_{t+1}.$$

Also, as $Q_{t,t+1}$ is the present value of the domestic state-contingent bond, it has an inverse relationship with the real interest rate:

$$Q_{t,t+1} = \frac{1}{(1+r_t)} \text{ and}$$

$$1+r_t = \frac{1+i_t}{1+E_t\pi_{t+1}}.$$

Since preferences are separable across periods, the intratemporal optimal choice can be made independently of the intertemporal optimal choice; therefore, the optimal choices of non-tradable and tradable consumption are

$$c_t^N = \frac{\gamma c_t}{p_t^N} \text{ and}$$

$$c_t^T = \frac{(1-\gamma)c_t}{p_t^T},$$

where p_t^N and p_t^T are the non-tradable and tradable prices relative to the CPI, which is

$$P_t = \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}\left(P_t^N\right)^{\gamma}\left(P_t^T\right)^{1-\gamma}.$$

The previous expression can be represented in real terms as follows:

$$1 = \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}\left(p_t^N\right)^{\gamma}\left(p_t^T\right)^{1-\gamma}.$$

The optimal choice of non-traded good variety j is

$$c_t^N(j) = \left(\frac{p_t^N(j)}{p_t^N}\right)^{-\theta} c_t^N,$$

and the non-tradable goods price level aggregator is

$$p_t^N = \left[\int_0^1 p_t^N(j)^{1-\theta}\,dj\right]^{\frac{1}{1-\theta}}. \tag{1}$$

Oil Extraction. In addition to the tradable and non-tradable sectors, there is also an oil-exporting sector in the economy. Oil activities are modeled as in Sickles and Hartley (2001). There is a representative oil-extracting firm owned by agents that decides how much oil to extract from the ground. At the beginning of any given year, the country has s units of oil reserves and x units can be extracted to be exported and sold in a competitive international oil market at the given relative price $p_t^{x,*}$, which is a stochastic variable (in units of foreign CPI). The total cost of extracting x units of oil in any year, given that there are s units of oil at the beginning of the year, is $C(s,x)$. The cost function C is decreasing in s (the total extraction cost falls the larger the oil reserves) and increasing in x (the total cost rises the higher the extraction rate). The marginal cost of an additional unit of reserves, conditioned on not extracting oil, is zero: $C_s(s,0) = 0$.

The problem of the representative oil firm is to maximize the expected discounted future stream of profits. The firm decides in each period the amount of oil to extract, x_t, and the level of future reserves, s_{t+1}. That is,

$$\max_{\{x_t, s_{t+1}\}} E_t\left\{\sum_{i=0}^{\infty}\beta^i\left[\xi_t^X\right]\right\}, \tag{2}$$

subject to

(3) $$s_{t+1} = s_t + d_t - x_t,$$

where d_t is a stochastic variable and represents oil discoveries. Profits are

(4) $$\xi_t^X = p_t^{x,*} x_t - C(x_t, s_t).$$

Optimal extraction satisfies the following conditions:

$$[x_t]: E_t\left\{p_t^{x,*} - \frac{\partial C}{\partial x_t} - \beta \Upsilon_t\right\} = 0$$

$$[s_{t+1}]: E_t\left\{-\frac{\partial C}{\partial s_{t+1}} - \Upsilon_t + \beta \Upsilon_{t+1}\right\} = 0,$$

where Υ_t is the Lagrange multiplier associated with the oil reserves accumulation equation.

The first optimality condition states that the price of oil should compensate not only today's marginal cost of extraction but also the discounted marginal value of future profits, which will depend on the stock of future reserves. The second condition states that the shadow price of existing oil reserves should be equal to the marginal cost of existing reserves and the discounted marginal value of future reserves. Note that in the steady state reserves should be constant and, therefore, the optimal rate of extraction equals the rate of discovery of new oil resources. Yet the level of reserves may be higher or lower depending on the cost structure, the random nature of discoveries, the interest rate, and the oil price.

The function we use to perform the quantitative experiments is

(5) $$C = \frac{\kappa}{2}\frac{x_t^2}{1+s_t},$$

which satisfies some restrictions commonly used in the natural resource economics literature.[5]

Both oil prices and discoveries follow these processes:

$$\log\left(p_t^{x,*}\right) = \rho_{p^x}\log\left(p_{t-1}^{x,*}\right) + \left(1-\rho_{p^x}\right)\log\left(p^{\bar{x},*}\right) + \varepsilon_t^{p^{x,*}}$$

$$\log(d_t) = \rho_d \log(d_{t-1}) + (1-\rho_d)\log(\bar{d}) + \rho^{d,p^x}\log\left(p_t^x\right) + \varepsilon_t^d,$$

where $\varepsilon_t^{p^{x,*}}$ and ε_t^d are i.i.d. $(0,\sigma^2)$.

In the "Calibration, Estimation, and Baseline Results" section, we show that discoveries are positively correlated with the international oil price.

Non-Tradable Goods Production. There is a representative firm producing a homogeneous non-tradable good in a perfectly competitive environment. The firm chooses two inputs—labor and oil—to produce the non-tradable good, which is also traded in competitive markets. The firm's objective is to minimize the total cost,

$$w_t h_t + p_t^x m_t,$$

subject to

$$y_t^N = A_t h_t^\alpha (m_t)^{1-\alpha},$$

where A_t represents the total factor productivity that follows an exogenous stochastic process, and m_t is the demand for oil from producers of non-tradable goods. Note that we have implicitly assumed that capital is fixed and equal to one unit for all t.

Under these assumptions, the firm's optimal choices of hours worked, oil, and real marginal cost are as follows:

$$w_t = \varphi_t A_t \alpha \left(\frac{m_t}{h_t}\right)^{\alpha-1}$$

$$p_t^x = \varphi_t A_t (1-\alpha) \left(\frac{m_t}{h_t}\right)^{-\alpha}$$

$$\varphi_t = A_t^{-1} \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} w_t^\alpha \left(p_t^x\right)^{1-\alpha},$$

and the homogeneous non-tradable good's price is $p_t^{NH} = \varphi_t$ because the homogeneous good is produced in a perfectly competitive environment.

Price Setting. There is a continuum of retail firms that buy the homogeneous non-tradable good from the perfectly competitive firms at p_t^{NH} and transform this homogeneous good into a differentiated variety j. Therefore, each of these firms has monopoly power in its respective variety. We assume that there is Calvo price stickiness. Each retailer receives a random signal to adjust its prices with a probability of $1 - \epsilon$, setting a price $\tilde{p}_t^N(j)$ to maximize

$$E_t \sum_{i=0}^{\infty} \epsilon^i \Lambda_{t,t+i} \left[p_t^N(j) y_{t+i}^N(j) - \varphi_{t+i} y_{t+i}^N(j) \right], \tag{6}$$

subject to

$$y_t^N(j) = \left(\frac{p_t^N(j)}{p_t^N}\right)^{-\theta} y_t^N \tag{7}$$

since the market clearing condition for each non-tradable variety holds—that is, $c_t^N(j) = y_t^N(j)$. Here $\Lambda_{t,t+i} = \frac{\beta \lambda_{t+i}}{\lambda_t}$ is the stochastic discount factor for households since they own the firm.

Therefore, the retailer's optimal price setting is represented by the first-order condition of solving (6) subject to (7), which is

$$\tilde{p}_t^N(j)=\frac{\theta}{\theta-1}\frac{\left\{E_t\sum_{i=0}^{\infty}(\epsilon\beta)^i\lambda_{t+i}\varphi_{t+i}y_{t+i}^N\left(p_{t+i}^N\right)^{\theta}\right\}}{\left\{E_t\sum_{i=0}^{\infty}(\epsilon\beta)^i\lambda_{t+i}y_{t+i}^N\left(p_{t+i}^N\right)^{\theta-1}\right\}}.$$

We assume that all retailers have the same cost structure and therefore set the same price, $\tilde{p}_t^N(j)=\tilde{p}_t^N$. By the law of large numbers, ϵ represents the fraction of retailers that keep their prices fixed and $1-\epsilon$ the fraction of retailers that reoptimize their prices by choosing $\tilde{p}_t^N$; then by using (1), the non-tradable good price index can be expressed as

$$1=\left[\epsilon\left(\frac{1}{1+\pi_t^N}\right)^{1-\theta}+(1-\epsilon)\left(\tilde{p}_t^N\right)^{1-\theta}\right]^{\frac{1}{1-\theta}},$$

which is the conventional Calvo-pricing equation for the determination of prices—in this case, the non-tradable good prices.

Central Bank. Since it is assumed that this economy has sticky prices, there is a role for monetary policy, which is characterized by the following nominal interest rate rule:

$$i_t=r_t^*+\bar{\pi}+\phi_\pi\left(\pi_t-\bar{\pi}\right),$$

where $\bar{\pi}$ is a fixed inflation target and ϕ_π is the degree of responsiveness of the central bank to deviations of inflation from its target. We use r_t^* as a proxy for a natural interest rate for a small open economy.

Market Clearing Conditions. From the household's budget constraint, it can be shown that, by using the market clearing condition for the non-tradable sector, $c_t^N=y_t^N$, and for the domestic bond market, $b_t=0$, and $c_t=p_t^Nc_t^N+p_t^Tc_t^T$, the balance of payments of the economy is

$$p_t^xm_t+p_t^Tc_t^T+q_tb_t^*\left(1+r_t^*\right)=p_t^Ty_t^T+p_t^xx_t+q_tb_{t+1}^*. \tag{8}$$

Basic Mechanisms at Work. A permanent negative oil price shock reduces disposable income permanently and causes a permanent reallocation of resources between the tradable and non-tradable sectors. Since the excess supply of tradable goods can be exported but the fall in demand for non-tradable goods is permanent, there should be a permanent real exchange-rate depreciation. Since in this model only the non-tradable sector produces goods using labor and imported inputs (i.e., gasoline) and non-tradable demand falls, the demand for these inputs also falls. Thus, employment falls and imports fall. Some of these imports are intermediate inputs used in the production of non-tradable goods. Since the price of intermediate inputs is also the price of oil, the real marginal cost is reduced, which increases the quantity demanded of that input, acting in the opposite direction to the fall in demand for non-tradable goods. On balance, one can expect that the direction of the derived demand for the imported input will be ambiguous.

A key mechanism works through the country risk premium. This premium is endogenous in the sense that it depends not only on net external debt, but also on the value of the stock of

Table 1

Estimation

	Prior			Posterior				
Parameter/Std	Distribution	Mean	Std	Mode	Std	Mean	HPD inf	HPD sup
ρ_d	β	0.500	0.150	0.3471	0.0774	0.3613	0.2243	0.4949
$\rho_{p^{x,*}}$	β	0.800	0.015	0.8633	0.0115	0.8618	0.8449	0.8812
$\rho_{p^{x,*},d}$	$\mathbb{N}$	0.000	0.150	0.2023	0.1112	0.2085	0.0204	0.3959
κ	Γ	2.000	0.250	3.7889	0.1456	3.7717	3.5778	4.0213
$\varepsilon_{p^{x,*}}$	inv Γ	0.125	inf	0.7531	0.0754	0.7608	0.6315	0.8862
ε_d	inv Γ	0.125	inf	1.0416	0.1690	1.1025	0.8193	1.3690
ε_β	inv Γ	0.125	inf	1.1169	0.0102	0.1190	0.1022	0.1352

NOTE: HPD, highest posterior density; Std, standard deviation; HPD inf, lower bound of a 90 percent HPD interval; HPD sup, upper bound of a 90 percent HPD interval.

SOURCE: From Hamann, Bejarano, and Rodríguez (2015, Table 9, p. 39).

oil. On the one hand, external debt will be higher, pushing the risk premium up. On the other hand, country risk will fall with the *value* of the stock of oil reserves, $p^x s$. A collapse in oil prices increases the risk premium. However, this effect is partially compensated by the endogenous response of the stock of oil reserves to oil prices. Reserves will increase in the future, lowering the country risk premium.

Nominal adjustment is important because there are nominal rigidities. Since prices do not adjust fully to shocks, real variables such as consumption, employment, and output adjust even further compared with a flexible price economy. Therefore, real variables in the sticky price economy are likely to be more volatile than their counterparts in the flexible price economy. However, another key aspect of the nominal adjustment of the model is the role of a flexible nominal exchange rate. First, since oil export revenues are transferred to households in domestic currency, the nominal exchange-rate depreciation partially compensates the fall in the value of exports denominated in foreign currency. The nominal exchange rate eases pressure on the household's budget constraint. Second, there is pass through from the nominal depreciation to inflation. Total inflation shoots up from the central bank's target, calling for a monetary policy response. The central bank raises the nominal interest rate, which in the presence of nominal rigidities in the non-tradable sector, amplifies the fall in economic activity.

Calibration, Estimation, and Baseline Results. The parameters of the model's oil sector block are estimated, while the parameters of the model's macro block are calibrated. For the estimation of the oil's sector block, we use annual data relevant to the Colombian economy for the period 1921-2013, including the British Petroleum (BP) crude oil price,[6] the change in oil reserves relative to total oil reserves, and the ratio of oil production to oil reserves.[7,8] Since variables such as the discovery of new oil reserves and an exogenous shock process are not observed, we use the Kalman filter and Bayesian methods to estimate the standard deviations for unobservable variables, parameter values, and exogenous shocks (Table 1). The table

Table 2

Long-Run Ratios: Model Versus Data

Relation		Model	Data
External debt/GDP	$\frac{qb^*}{y}$	–0.30	–0.30
Labor income/GDP	$\frac{wh}{y}$	0.36	0.36
Non-tradable output/ Tradable output	$\frac{p^N y^N}{p^T y^T}$	1.74	1.74
Oil reserves/Oil production	$\frac{s}{x}$	6.30	6.30

NOTE: The table shows the long-run ratios of key macro variables for the model and the Colombian economy using annual frequency data from the National Administrative Department of Statistics of Colombia (DANE).

SOURCE: From Hamann, Bejarano, and Rodríguez (2015, Table 10, p. 40).

Table 3

Key Calibrated Parameters of the Model

Parameter		Value
Inverse Frisch elasticity	ω	1.6085
Long-run productivity level	A	0.0644
Long-run tradable GDP level	y^T	1.3389
Long-run tradable foreign relative price level	p^T	0.9438
Long-run oil foreign relative price level	$\bar{p^x}$	1.6896
Long-run discoveries level	$\bar{d}$	0.2113
Interest rate to debt elasticity	ψ	0.0544
Elasticity of substitution among varieties	θ	3.3571

SOURCE: Modified from Hamann, Bejarano, and Rodríguez (2015, Table 11, p. 40).

Table 4

Other Parameters of the Model

Parameter		Value	Source
Non-tradable consumption share	γ	0.6000	DANE
Labor participation in non-tradable production function	α	0.9000	González et al. (2011)
Intertemporal elasticity of substitution	σ	4.0000	González et al. (2011)
Oil discount factor	β^{oil}	0.9661	González et al. (2011)
Long-run foreign real interest rate	$\bar{r^f}$	0.0350	González et al. (2011)

NOTE: DANE, National Administrative Department of Statistics of Colombia.

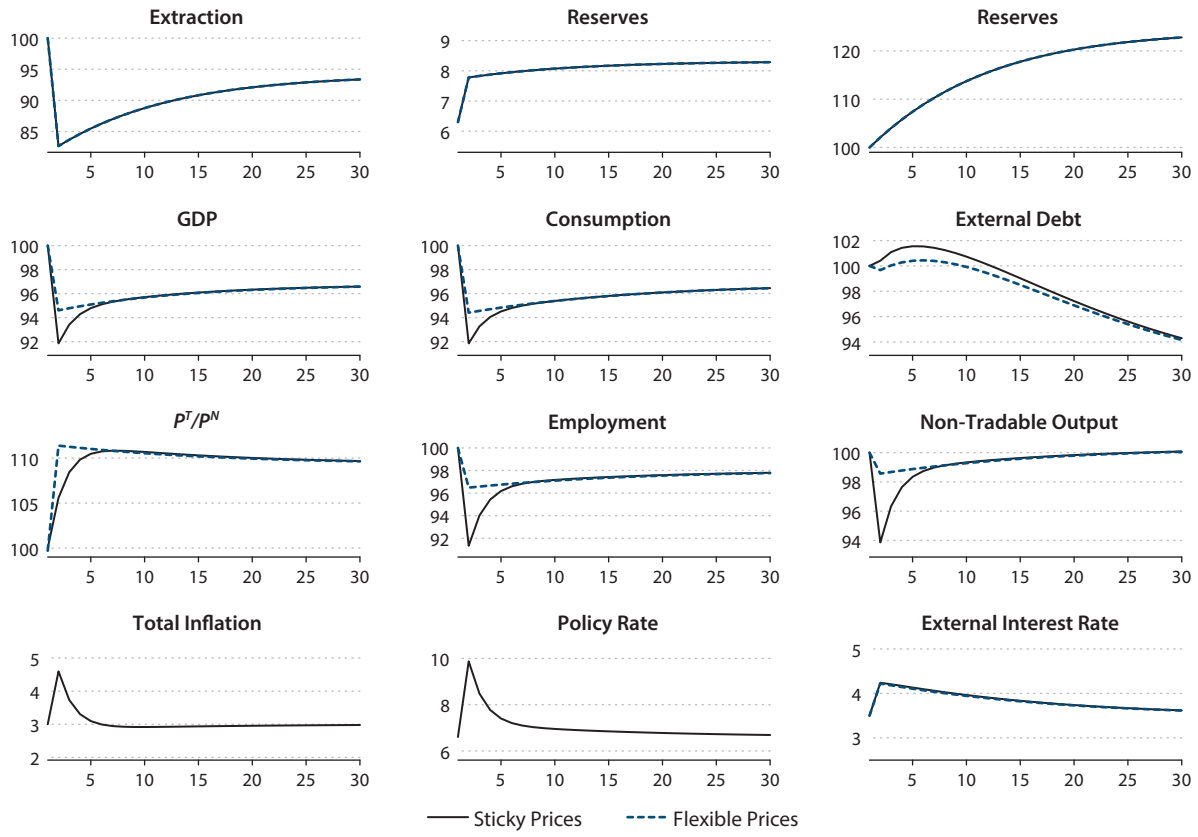
SOURCE: Modified from Hamann, Bejarano, and Rodríguez (2015, Table 12, p. 41).

reports prior and posterior distributions along with standard deviations for both parameters and shocks. The posterior distributions were computed using two Markov chains (the Markov chain Monte Carlo method) with 100,000 draws each.

Table 2 shows the long-run ratios of key macro variables for the model and the Colombian economy using annual frequency data from the National Administrative Department of Statistics of Colombia (DANE). Since the model has only labor in the non-tradable sector, we make the following calibrations: Total labor income is 60 percent of total output and non-tradable production weights 60 percent of total production. Therefore, we set the labor income share of the non-tradable sector at 0.36. Other values of the parameters used in the calibration of the model are reported in Table 3. The remaining parameters (reported in Table 4) are from previous studies of the Colombian economy.

Figure 2

Short-Run Macro Adjustment to a Permanent Fall in Oil Prices



NOTE: The inflation rate, policy (interest) rate, and external interest rate are expressed in levels; reserves are expressed in years; and the remaining variables are expressed as the deviation from their steady state, which is normalized to 100. The *x*-axis represents quarters.

SOURCE: Modified from Hamann, Bejarano, and Rodríguez (2015, Figure 11, p. 46).

Estimated Effects of Permanent Lower Oil Prices. To assess the monetary policy implications of permanent changes in oil prices, we perform an impulse response analysis by setting the persistence parameter, ρ_{p^x}, of the oil price stochastic process very close to 1. The quantitative results of the transitional dynamics exercise are reported in Figure 2. We report two cases: one with flexible prices and the other with sticky prices.

The collapse in oil prices has a large impact on the oil sector. Oil extraction is cut by nearly 20 percent, oil profits tank, and oil reserves increase by nearly 20 percent in the long run. As expected, most of the adjustment in the reaction to the permanent change in oil prices happens in the oil sector. As long as the current account is still another vehicle to smooth the permanent change in oil prices, the model predicts the current account will deteriorate slightly for a few years and then move into positive territory to eventually converge to its steady-state value, which is zero. In this model, no impatience is imposed on agents. Thus, private agents initially

borrow to mitigate the adjustment in consumption caused by the short-run negative overreaction in oil production.

Country risk increases by nearly 50 basis points on impact—a relatively small jump—to later fall back as oil reserves increase. Recall that on the one hand, external debt will be higher and lower oil prices will push the risk premium up, but on the other hand, reserves will increase in the future, lowering the country risk premium. As it turns out, with the baseline calibration the impact on country risk is small, especially in the long run.

Consumption falls on impact by 4 percent in the flexible price economy and by 6 percent in the sticky price economy. GDP also falls by similar magnitudes in both models. In the long run, consumption and GDP fall by about 4 percent. The collapse in total consumption triggers a real depreciation: Tradable consumption adjustment happens through the trade balance, while non-tradable consumption and activity tank. This fall in non-tradable consumption generates a contraction in non-tradable production and labor demand that, in turn, reduces the real wage.[9] The result is a 6 percent depreciation in the real exchange rate on impact in the flexible price economy and a smooth real depreciation in the sticky price economy. Both models predict a permanent real depreciation of around 10 percent.

The real depreciation dynamics reflect an increase in inflation in the tradable sector and a decrease in inflation in the non-tradable sector caused by the fall in the real wage and the imported input price (Figure 3). This real depreciation is consistent with nominal depreciation passing through to total inflation. In the sticky price economy, total inflation jumps 1 percent away from the inflation target, triggering a central bank response of a 3.5 percent increase in its policy rate. In the flexible price economy, these effects are smaller in magnitude.

The model also highlights a monetary policy dilemma. In this economy, the exchange-rate pass through to total inflation turns out to be high. Thus, since the central bank is assumed to target total inflation, it raises the policy rate. The policy change manages to drive inflation back to target eventually, but it does so as the oil exports and non-tradable sectors are adjusting to the new condition. A key insight from this model is that monetary policy simply cannot accommodate part of the adjustment: The economy is permanently poorer and this effect is felt in both economies—those with and without sticky prices.

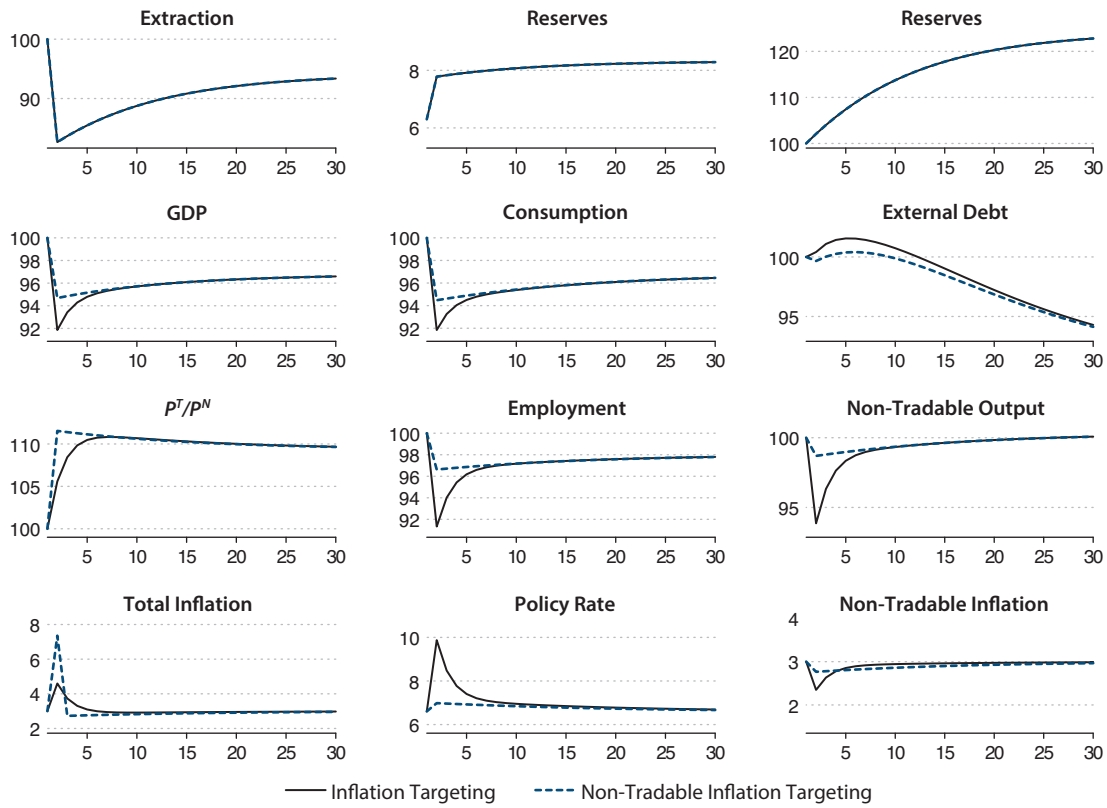
Of course, a central banker would be reluctant to raise interest rates in light of a permanent real shock with potentially large negative effects on the economy. The root of the problem is that in this economy, nominal depreciation is passing through to tradable inflation and thus driving total inflation off target. An alternative to targeting total inflation is for the central bank to target non-tradable inflation. This makes sense because in the model there is an extreme situation in which the only source of nominal rigidities resides in the non-tradable sector. Prices in the tradable sector are flexible. Thus, for the total inflation-targeting central bank, we implicitly assume that the bank ignores in which of the sectors the nominal rigidities lie.

We perform a counterfactual experiment in which we use the same shock to simulate what would have happened had the central bank targeted non-tradable inflation instead of total inflation. Figure 3 reports the results of this transitional dynamics experiment for the macroeconomic variables.

In this alternative economy, instead of hiking the policy rate the central bank barely raises it. Consumption, GDP, and employment fall by less initially, and external debt does

Figure 3

Short-Run Macro Adjustment to a Permanent Fall in Oil Prices: Total Inflation Target Versus Non-Tradable Inflation Target



NOTE: The inflation rate, policy (interest) rate, and external interest rate are expressed in levels; reserves are expressed in years; and the remaining variables are expressed as the deviation from their steady state, which is normalized to 100. The *x*-axis represents quarters.

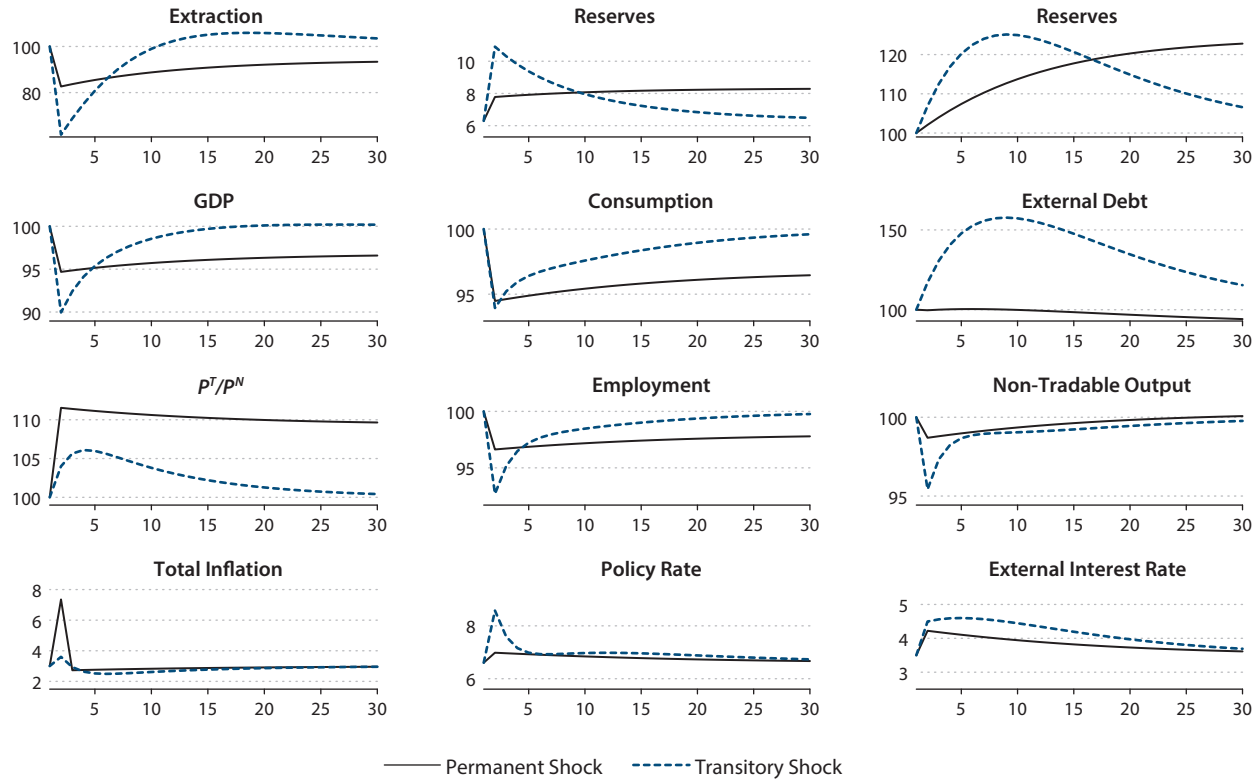
not expand as much as when policy targets total inflation. The long-run effects on both total inflation-targeting and non-tradable inflation-targeting regimes are identical. Obviously, total inflation skyrockets 400 basis points with respect to its long-run target. Once again, this intuitive result highlights the short- to medium-term policy dilemma for inflation-targeting central banks in oil-exporting economies.[10] We are aware that our analysis is limited since we are not determining the central bank's optimal policy rule.

Estimated Effects of Transitory Lower Oil Prices. To assess the monetary policy implications of transitory changes in oil prices, we perform an impulse response analysis by setting the persistence parameter, ρ_{p^x}, of the oil price stochastic process at 0.8618.[11] The quantitative results of the transitional dynamics exercise are reported in Figure 4. We compare the effects of this transitory shock with those of a permanent shock.

In the short run, the transitory collapse in oil prices has a larger impact on the oil sector than a permanent collapse. Oil extraction is cut by nearly 60 percent, oil profits tank, and oil

Figure 4

Short-Run Macro Adjustment to a Transitory Fall in Oil Prices: Total Inflation Target



NOTE: The inflation rate, policy (interest) rate, and external interest rate are expressed in levels; reserves are expressed in years; and the remaining variables are expressed as the deviation from their steady state, which is normalized to 100. The *x*-axis represents quarters.

reserves increase by nearly 11 percent in the short run, but after 30 periods they return to their long-run values.

Country risk increases by nearly 160 basis points—a huge jump—on impact as a consequence of the huge depletion in net foreign assets in the short run. Unlike the permanent fall in oil prices, the impact on foreign debt is huge as long as households can smooth consumption. Note that here consumption falls less than when there is a permanent fall in oil prices because the drop in revenues from oil is transitory as the cut in oil extraction and the collapse in oil prices are transitory.

Unlike the case of a permanent fall in oil prices, the real and nominal depreciation here are smaller. As Figure 4 shows, the central bank increases its nominal interest rate because of the natural interest rate hikes caused by the huge increase in foreign debt. This allows the central bank to keep total inflation close to its target.

CONCLUSION

In this article, we analyzed the macroeconomic consequences and the monetary policy implications of permanent changes in oil prices for a small open economy from the perspective of a dynamic stochastic general equilibrium framework. We used a quantitative approach within this framework for an economy with three sectors (tradable, non-tradable, and oil), incomplete financial assets markets, nominal rigidities, market imperfections, endogenous oil production, and different monetary policy targets.

We found that the optimal response of the oil sector in these economies was to cut extraction significantly and increase prospective long-term oil reserves. We found that long-lived lower oil prices imply a challenge for an inflation-targeting central bank. On the one hand, the permanent fall in oil revenues causes a significant and permanent fall in consumption and GDP. On the other hand, the nominal depreciation of the exchange rate drives total inflation off target, calling for the central bank to tighten its policy stance. Thus, both the nominal and the real exchange rate adjustments are at the core of the adjustment mechanism.

Finally, we also found an important role for the external interest rate faced by the economy in international financial markets. The estimated large-scale financial frictions model predicts a protracted period of higher external interest rates because of a higher risk premium. This effect, induced by larger foreign financing needs and low oil prices, dominates the effect of the lower risk induced by the higher level of future oil reserves that accumulate endogenously in the economy. The interaction of these real adjustments with nominal rigidities is important because the model delivers a nominal exchange rate depreciation, which passes through to total inflation. The pass through may be significant. It temporally but persistently raises the annual inflation well above target, calling for the model's strict inflation-targeting central bank to tighten monetary policy to control inflation. If the central bank can identify that the price stickiness resides in the non-tradable sector and chooses to target non-tradable inflation instead of total inflation, the bank cuts the policy rate. However, the resulting total inflation will be even higher in this artificial economy. ■

APPENDIXES

Appendix A: Equations of Monetary Policy with Oil Sector

(A.1) $$tb_t + q_t b_t^* = \left(1 + r_{t-1}^*\right) q_t b_{t-1}^*$$

(A.2) $$tb_t = p_t^T y_t^T + p_t^x x_t - p_t^x m_t - p_t^T c_t^T$$

(A.3) $$y_t^N = A_t h_t^\alpha \left(m_t\right)^{1-\alpha}$$

(A.4) $$\beta E_t \lambda_{t+1} = Q_{t,t+1} \lambda_t$$

(A.5) $$Q_{t,t+1} = \frac{1}{\left(1 + r_t\right)}$$

(A.6) $$\left[c_t - \frac{h_t^\omega}{\omega}\right]^{-\sigma} = \lambda_t$$

(A.7) $$r_t^* = r_t^f + \psi\left[\exp\left(\frac{q_t b_t^*}{p_t^X S_t} - \frac{qb^*}{p^X S}\right) - 1\right]$$

(A.8) $$\left[c_t - \frac{h_t^\omega}{\omega}\right]^{-\sigma} h_t^{\omega-1} = w_t^h \lambda_t$$

(A.9) $$\log\left(A_t\right) = p^A \log\left(A_{t-1}\right) + \left(1 - \rho^A\right)\log\left(\bar{A}\right) + \varepsilon_t^A$$

(A.10) $$tb_{share,t} = \frac{tb_t}{y_t}$$

(A.11) $$ca_{share,t} = tb_{share,t} - \frac{r_t^* b_t^*}{y_t}$$

(A.12) $$c_t^T = \frac{(1-\gamma) c_t}{p_t^T}$$

(A.13) $$y_t^N = c_t^N$$

(A.14) $$p_t^T = q_t\, p_t^{T,*}$$

(A.15) $$\log\left(p_t^{T,*}\right) = \rho^{p^{T,*}} \log\left(p_{t-1}^{T,*}\right) + \left(1-\rho^{p^{T,*}}\right)\log\left(p^{\overline{T},*}\right) + \varepsilon_t^{p^{T,*}}$$

(A.16) $$q_t = \frac{\left(1+r_t^*\right)E_t\,\lambda_{t+1} q_{t+1}}{\left(1+r_t\right)E_t\,\lambda_{t+1}}$$

(A.17) $$r_t^f = \rho^{r^f} r_{t-1}^f + \left(1-\rho^{r^f}\right)\overline{r}^f + \varepsilon_t^{r^f}$$

(A.18) $$\tilde{p}_t^N = \frac{num_t}{den_t}$$

(A.19) $$num_t = \frac{\theta\lambda_t\,\varphi_t\,y_t^N}{p_t^N} + \epsilon\beta num_{t+1}\left(1+\pi_{t+1}^N\right)^{\theta}$$

(A.20) $$den_t = (\theta-1)\,\lambda_t\,y_t^N + \epsilon\beta den_{t+1}\left(1+\pi_{t+1}^N\right)^{\theta-1}$$

(A.21) $$\varphi_t = A_t^{-1}\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}\,w_t^{\alpha}\left(p_t^x\right)^{1-\alpha}$$

(A.22) $$1 = \epsilon\left(\frac{1}{1+\pi_t^N}\right)^{1-\theta} + (1-\epsilon)\left(\tilde{p}_t^N\right)^{1-\theta}$$

(A.23) $$\frac{p_t^N}{p_{t-1}^N} = \frac{\left(1+\pi_t^N\right)}{\left(1+\pi_t\right)}$$

(A.24) $$y_t = p_t^N y_t^N + p_t^T y_t^T + p_t^x x_t$$

(A.25) $$i_t = r_t^* + \overline{\pi} + \phi_\pi\left(\pi_t - \overline{\pi}\right) + z_t^i$$

(A.26) $$1 = \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}\left(p_t^N\right)^{\gamma}\left(p_t^T\right)^{1-\gamma}$$

(A.27) $$z_t^i = \rho^{z^i} x_{t-1}^i + \left(1-\rho^A\right)\overline{z}^i + \varepsilon_t^{z^i}$$

(A.28) $$\left(1+r_t\right) = \frac{\left(1+i_t\right)}{\left(1+\pi_{t+1}\right)}$$

(A.29) $$c_t^N = \frac{\gamma c_t}{p_t^N}$$

(A.30) $$p_t^{x,*} = \frac{2\kappa x_t}{1+s_{t-1}} - \beta\left(2\kappa\frac{x_{t+1}}{1+s_t} - p_{t+1}^{x,*} - \kappa\left(\frac{x_{t+1}}{1+s_t}\right)^2\right)$$

(A.31) $$s_t = s_{t-1} - x_t + d_t$$

(A.32) $$\log(d_t) = \rho^d \log(d_{t-1}) + (1-\rho^d)\log(\bar{d}) + \varepsilon_t^d$$

(A.33) $$\log(y_t^T) = \rho\log(y_{t-1}^T) + (1-\rho)\log(\bar{y}^T) + \varepsilon_t^{y^T}$$

(A.34) $$\log(p_t^{x,*}) = \rho^{p^{x,*}}\log(p_{t-1}^{x,*}) + (1-\rho^{p^{x,*}})\log(p^{\bar{x},*}) + \varepsilon_t^{p^{x,*}}$$

(A.35) $$p_t^x = q_t p_t^{x,*}$$

(A.36) $$p_t^x = \varphi_t(1-\alpha)\left(\frac{m_t}{h_t}\right)^{-\alpha}$$

Appendix B: Dataset

Commercial Debt Portfolio. We used the commercial monthly real debt portfolio of the Colombian financial sector and converted it to a quarterly frequency using the value for the last month in the quarter. These data are available from 1998:Q4 to 2013:Q2.

Sectoral Commercial Debt Portfolio. We built a tradable and non-tradable commercial debt portfolio measure by adding the sectoral data. In particular, for the tradable measure, we used the commercial debt portfolio of the agriculture, fishing, mining, manufacturing, and wholesale and retail commerce sectors. For the non-tradable sector, we used these sectors: hotel and restaurant, transportation, financial intermediation, real estate, public administration, education, health, other social services, households with domestic service, and extraterritorial organs. These measures were then deflated using the CPI and were seasonally adjusted using Census X-12. These data are available from 1999:Q1 to 2013:Q2.

Oil Production. We used the monthly average of the daily crude oil production (in barrels) and averaged it for each quarter. These data are available from 1993:Q1 to 2013:Q2.

Oil Price. Quarterly prices were calculated from daily data by using an unweighted average of the daily closing spot prices for Brent Crude oil. We took the seasonally adjusted series and deflated it by the U.S. CPI. We used the cyclical component of oil prices after using a Hodrick-Prescott filter. These data are available from 1999:Q1 to 2013:Q2.

Consumption. We used disaggregated quarterly data of total private consumption from 2000:Q1 to 2013:Q2. In particular, this disaggregation divides consumption into non-

durable, durable, and semi-durable goods and services. We then approximated tradable consumption as the sum of consumption in durable and semi-durable goods and non-tradable consumption as the sum of consumption in nondurable goods and services.

Gross Fixed Capital Formation. We used disaggregated quarterly data of total gross fixed capital formation from 2000:Q1 to 2013:Q2. In particular, this disaggregation divides fixed capital formation by sector: agricultural, machinery, transportation, construction, civil project building, and services. We then approximated tradable fixed capital formation as the sum of this among the following sectors: agricultural, machinery, and transportation. We approximated non-tradable fixed capital formation as the sum of this among the following sectors: construction, civil project building, and services.

GDP. We built a measure of tradable and non-tradable GDP using sectoral data. Specifically, tradable GDP was approximated using the sum of agriculture, silviculture, hunting and fishing, mining, manufacture, air transportation, supplementary transportation services, mail and communication services, financial services to firms (excluding real estate), and total taxes. Non-tradable GDP was then computed as the difference between total and tradable GDP. We also computed a measure of tradable GDP excluding the mining sector. These data are available from 2000:Q1 to 2013:Q2.

Inflation. We built a measure of tradable and non-tradable inflation based on the CPI of the same sectoral data as those of the GDP. These CPI measures (tradable and non-tradable) were then seasonally adjusted using Census X-12 and converted to quarterly frequency by using the value for the last month in the quarter. These CPI data were then used to compute quarterly inflation. These inflation measures are available from 1999:Q2 to 2013:Q2.

Deposits. We used the quarterly savings account data starting in 1984:Q1 and ending in 2013:Q2. We then seasonally adjusted this measure using Census X-12.

Interest Rates. We used the monthly data for the interbank interest rate, the home building interest rate (different from social housing), and the corporate commercial interest rate and converted them to quarterly frequency using the value for the last month in the quarter. A measure for tradable interest rate was then approximated using the corporate commercial interest rate. The non-tradable interest rate was approximated using the home building interest rate. These data are available from 2002:Q2 to 2013:Q2.

NOTES

[1] We define the country risk premium as the difference between the risk-free interest rate and the interest rate effectively paid by debtor countries for external debt.

[2] Evidence for Colombia indicates that imported goods prices are adjusted roughly every quarter; see Bonaldi, González, and Rodríguez (2011).

[3] Since the oil firm is in a competitive international oil market, its revenues and costs are denominated in foreign currency.

[4] We make this assumption as long as the share of the non-mining exports to GDP is around 5 percent.

[5] See Pindyck (1981) for instance.

[6] The British Petroleum crude oil price series is from the BP Statistical Review of World Energy 2014. This crude price is constructed with the Brent Crude price dated over the 1984-2013 period, the Arabian Light crude price posted at Ras Tanura in the 1945-1983 period, and the U.S. average crude price over the 1861-1944 period.

[7] The data for the stock of reserves and oil production are from Colombia's National Hydrocarbons Agency.

[8] To avoid stochastic singularity in the oil's block estimation, a preferences shock is included in the model.

[9] This is a consequence of the preferences specification assumed.

[10] To check the robustness of our results, we perform an experiment with alternative specifications of the policy rule that also includes the output gap. In the first experiment, we use an output gap defined as the difference between the sticky price GDP and the flexible price GDP. In the second experiment, we use a different definition of output gap: the difference between the sticky price non-tradable output and the flexible price non-tradable output.

[11] This is the estimated value of the persistence parameter of the oil price, as reported in Table 1.

REFERENCES

Auray, Stéphanie; de Blas, Beatriz and Eyquem, Aurélien. "Ramsey Policies in a Small Open Economy with Sticky Prices and Capital." *Journal of Economic Dynamics and Control*, September 2011, *35*(9), pp. 1531-46; http://dx.doi.org/10.1016/j.jedc.2011.03.013.

Benigno, Gianluca and De Paoli, Bianca. "On the International Dimension of Fiscal Policy." *Journal of Money, Credit, and Banking*, December 2010, *42*(8), pp. 1523-42; http://dx.doi.org/10.1111/j.1538-4616.2010.00352.x.

Bonaldi, Pietro; González, Andrés and Rodríguez, Diego A. "The Importance of Nominal and Real Rigidities in Colombia: A Dynamic Stochastic General Equilibrium Approach." *Ensayos Sobre Política Económica*, December 2011, *29*(66), pp. 48-78.

De Paoli, Bianca. "Monetary Policy Under Alternative Asset Market Structures: The Case of a Small Open Economy." *Journal of Money, Credit, and Banking*, October 2009, *41*(7), pp. 1301-30; http://dx.doi.org/10.1111/j.1538-4616.2009.00257.x.

Galí, Jordi and Monacelli, Tommaso. "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies*, July 2005, *72*(3), pp. 707-34; http://dx.doi.org/10.1111/j.1467-937X.2005.00349.x.

Gertler, Mark and Karadi, Peter. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics*, January 2011, *58*(1), pp. 17-34; http://dx.doi.org/10.1016/j.jmoneco.2010.10.004.

González, Andrés; Mahadeva, Lavan; Prada, Juan D. and Rodríguez, Diego A. "Policy Analysis Tool Applied to Colombian Needs: PATACON Model Description." *Revista Ensayos Sobre Política Económica*, December 2011, *29*(66).

González, Andres; Hamann, Franz and Rodríguez, Diego A. "Macroprudential Policies in a Commodity Exporting Economy." BIS Working Papers No. 506, Bank for International Settlements, July 2015; http://www.bis.org/publ/work506.pdf.

Hamann, Franz; Bejarano, Jesus A. and Rodríguez, Diego A. "Monetary Policy Implications for an Oil-Exporting Economy of Lower Long-Run International Oil Prices." Working Paper No. 871, Borradores de Economía, Banco de la República, March 2015; http://www.banrep.gov.co/sites/default/files/publicaciones/archivos/be_871.pdf.

Kilian, Lutz. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review*, June 2009, *99*(3), pp. 1053-69; http://dx.doi.org/10.1257/aer.99.3.1053.

Kilian, Lutz; Rebucci, Alessandro and Spatafora, Nikola L. "Oil Shocks and External Balances." *Journal of International Economics*, April 2009, *77*(2), pp. 181-94; http://dx.doi.org/10.1016/j.jinteco.2009.01.001.

Neumeyer, Pablo A. and Perri, Fabrizio. "Business Cycles in Emerging Economies: The Role of Interest Rates." *Journal of Monetary Economics*, March 2005, *52*(2), pp. 345-80; http://dx.doi.org/10.1016/j.jmoneco.2004.04.011.

Pesaran, M. Hashem. "An Econometric Analysis of exploration and Extraction of Oil in the U.K. Continental Shelf." *Economic Journal*, June 1990, *100*(401), pp. 367-90; http://dx.doi.org/10.2307/2234130.

Pindyck, Robert S. "The Optimal Production of an Exhaustible Resource When Price Is Exogenous and Stochastic." *Scandinavian Journal of Economics*, Spring 1981, *83*(2), pp. 277-88; http://dx.doi.org/10.2307/3439901.

Rebucci, Alessandro and Spatafora, Nikola L. "Oil Prices and Global Imbalances," in *IMF World Economic Outlook*. Chap. 4. Washington, DC: International Monetary Fund, April 2006, pp. 71-96; https://www.imf.org/external/pubs/ft/weo/2006/01/pdf/weo0406.pdf.

Schmitt-Grohe, Stephanie and Uribe, Martin. "Downward Nominal Wage Rigidity and the Case for Temporary Inflation in the Eurozone." *Journal of Economic Perspectives*, Summer 2013, *27*(3), pp. 193-212; http://dx.doi.org/10.1257/jep.27.3.193.

Sickles, Robin C. and Hartley, Peter. "A Model of Optimal Dynamic Oil Extraction: Evidence from a Large Middle Eastern Field." *Journal of Productivity Analysis*, January 2001, *15*, pp. 59-71; http://dx.doi.org/10.1023/A:1026547923853.