What Caused the Great Depression? Herbert Hoover and Labor Shedding

Lee E. Ohanian
UCLA and Federal Reserve Bank of Minneapolis

November, 2006 - Preliminary and Incomplete

Abstract

This paper addresses the factors that started the Great Depression. Using monthly data from the industrial sector, I find that there are three significant puzzles about the start of the Depression.. The first is that the Depression was immediately severe, but that standard business cycle shocks, including monetary shocks, productivity shocks, banking shocks, and other shocks were either small or non-operative before the Depression through mid-1931. The second puzzle is that industrial hours worked fell more than implied by changes in productivity or the real wage. The third puzzle is that both nominal and real wages in the industrial sector rose during the early stages of the Depression. The most promising theory for understanding these puzzles is a change in labor policies that reduced the ability of firms to suppress unionization.

I thank Hal Cole, Monique Ebel, David Levine, Ed Prescott, Albrecht Ritschl, and Jean-Laurent Rosenthal for discussions.

1 Introduction

There are two phases of the Great Depression: the Great Contraction, from 1929 to 1933, and the failure to recover, from 1933 to 1939. This paper addresses the start of the Contraction and aims to understand why the Contraction started and why it was so immediately severe; hours worked in the industrial sector fell 42 percent between October, 1929 and August, 1931. To address these questions, I analyze the highest frequency data available,

which is monthly output and hours data from the industrial sector. I find that there are three significant puzzles about the start of the Contraction and the subsequent two years of the Contraction.

The first puzzle is that the Contraction was immediately severe, but standard business cycle shocks - monetary shocks, bank panics, productivity shocks, tax shocks - are either small or non-operative prior to the Depression up through Fall, 1931, when industrial hours are down by 42 percent. I refer to this large decline in hours worked as "labor shedding". This indicates not only that the decline in industrial output was primarily the consequence of labor shedding, rather than lower productivity, but also indicates that an alternative theory is required to understand the first two years of the contraction. Any theory will likely have an industrial genesis, because agricultural hours and output - which are typically procyclical and more volatile than the non-farm economy - were on trend during this period.

I next document a number of facts about labor, productivity, and wages in the industrial sector that guide the development of such a theory. One fact is that union wages rose considerably prior to the Contraction, while non-union wages were stagnant during the 1920s. Stagnant non-union wages are particularly striking given that output per hour rose nearly 30 percent in the 1920s. A third fact - and the second puzzle about the Contraction - is that labor shedding occurred despite no significant changes in the standard labor demand shifters, productivity and relative price; hours worked fell significantly more than implied by the standard labor demand relation. A fourth fact - and a third puzzle - is that both nominal and real industrial wages *rose* during the early stages of the Contraction.

I find that the most promising theory for understanding the large decline in industrial labor demand and the coincidence of high wages is based on changes in labor policies that occurred at the end of the 1920s. Court decisions in 1928 and 1929 laid the groundwork for establishing independent unions and substantially increased acceptable labor practices, including picketing and other strike activities, and President Hoover met with the CEO's of major industrial firms in November, 1929, and advised them to keep wages high to help him keep the "industrial peace" with labor. These court and executive policy changes prevented firms from suppressing union organization with the procedures they successfully used in the 1920s.

I develop a model of this change in labor policies. Before the policy shift, industry can suppress unionization. After Hoover's policy shift, industrial firms can be unionized, and the probability of unionization depends on the wage rate and the relative size of the firm. This latter feature follows from the fact that unions targeted large firms. This relative size effect leads to a *shift* in industrial labor demand, rather than just moving along the labor demand curve in response to higher wages. The model also includes a farm

sector in which there is no unionization. The distortion in the industrial sector leads to a relative expansion in the farm sector.

Section 2 summarizes the data during the first two years of the Depression, focusing on the immediate severity of the Contraction. Section 3 presents data on various shocks before and during the early stages of the depression and shows that these standard shocks were non-operative during these periods. Section 4 evaluates various factors that could have depressed labor, and shows that any candidate theory needs to account for a large negative shift in labor demand, and higher nominal and real manufacturing wages. Section 5 describes labor markets and unionization policies in the 1920s and after 1929. Section 6 describes a model economy that captures this policy change. Section 7 summarizes evidence supporting the theory. Section 8 discusses the continuation of the Depression after Hoover's influence on the industrial sector is over. Section 9 describes related literature. Section 10 presents a summary and conclusion.

2 The Great Contraction

This section summarizes the data, with a focus on identifying when the Contraction began and documenting the initial severity through late 1931. I begin by presenting monthly, seasonally adjusted data on industrial employment and production (IP) between early 1929, before the Depression, and the trough of the industrial depression, which is July 1932. Figure 1 shows industrial production (IP), and Figure 2 shows industrial hours worked. (Industrial production is from the Board of Governors, and manufacturing hours are from the NBER macro history database). These data clearly show that industrial depression begins abruptly in late 1929, and is immediately severe. IP and industrial hours were about 4 percent and 2 percent higher in October relative to January, and were about 2-3 percent below their summertime peaks.

Between October 1929 and June 1930, IP and hours are down about 15 percent and 22 percent, respectively. By August 1931, they are down about 34 percent and 40 percent, respectively. I will refer to this period as the first phase of the Contraction. There are three reasons I choose to focus on the first phase. One is that the factors that are cited in different literatures for causing the Depression - including monetary decline and productivity decline - are primarily operative *after* the first phase. Another is that the influence of President Hoover on industrial firm decisions, which I will describe later, begins in late 1929 and ends in late 1931. A third reason is that the first phase of the Depression is relatively understudied. Using high frequency data from the first phase will permit me to focus on the factors that started the

Depression, including the potential contribution of Hoover. I now examine these data in more detail.

I now examine data at the industry level. The immediacy and severity of the industrial Depression occurs throughout most industrial sector. Tables 1 - 4 show economic activity in various industrial sectors, including aggregate business fixed investment (source: NBER macro history database), hours worked, and output in several manufacturing and mining industries. In each industry, output and employment begin declining significantly after October 1929. While the industrial sector immediately and significantly declined after Fall, 1929, the farm sector, which accounted for about the same fraction of total employment as manufacturing at the time, did not decline. Farm hours are nearly two percent above trend in 1931.

The substantial difference between the manufacturing sector and farm sector differs considerably from a typical recession, in which both manufacturing and farming decline. A successful theory of the depression must not only account for why the manufacturing sector declined so quickly and significantly, but why the farm sector did not contract. I conclude from this analysis that the initiating shock of the Great Depression was primarily, or perhaps exclusively, an industrial shock.

3 Examining the Standard Shocks

I now examine the possible contributions of various shocks. The most plausible candidates will be those that would have a very large impact on the industrial sector, and that would occur during or before the Fall of 1929. I begin by considering the shocks that are emphasized in the literature.

3.1 Monetary Shocks

Monetary policy shocks are a frequently-cited shock (e.g. Friedman and Schwartz). But the very large monetary contraction that Friedman and Schwartz focus on occurs *after* the first phase (after August, 1931). Before the Depression (between late 1927 and the fall of 1929), there was a small monetary contraction (all data are from Friedman and Schwartz). The monetary base is about 3 percent lower between Dec. 1927 and October 1929, while M2, the aggregate studied by Friedman and Schwartz, is about 7 percent higher between these dates. These changes do not appear to be able to account for the large and immediate decline in hours worked. For example, the monetary base declines nearly 17 percent during the deflation of 1921-22, and real GNP declines about 4 percent.

Some economists estimate monetary policy shocks using short-term interest rates, such as the Federal Funds rate (e.g. Christiano and Eichenbaum), rather than monetary aggregates. Standard practice in this literature is to estimate the policy shock as the residual from a forecasting model for the interest rate. I measure the pre-Contraction interest rate shock simply as the accumulated change in interest rates using the Federal Reserve discount rate (there was no Fed Funds market at this time).

The discount rate is 4 percent in mid-1927, followed by a brief decline to 3.5 percent in late 1927, it is 5 percent in all Reserve Banks by May 1929, and remains at 5 percent until November, 1929, when it starts to decline. (The New York Fed discount rate was briefly at 6 percent, from August, 1929 through October 1929). The data are from Banking and Monetary Statistics, Board of Governors of the Federal Reserve System, 1943. Estimating the policy shock as the residual from a forecasting equation, as in the literature, would result in a smaller shock, since some of this interest rate increase would be accounted for by the predictive variables.

To assess the contribution of the increase in the discount rate, I fed in the sequence of changes in the New York Fed discount rate from 1927 through 1932 into the model of Christiano, Eichenbaum, and Evans (2003), with the exception that I mechanically doubled the impact of the shock in their model to account for the possibility that monetary shocks had larger effects during this period. Figure 3 shows the time path of actual real GNP, and the time path of output from the model. The decline generated by the model is clearly much smaller than that in the data (although the model does generate a modest decline beginning in the summer of 1929 that is also observed in the data). The output decline in the model is small because the shock is small. I am unaware of a model in which a monetary shock of this size, and its timing, could account for the very rapid and large change in hours worked that occurs beginning in late 1929.

Regarding monetary factors after October 1929, the monetary base begins to increase and the Federal Reserve discount rate declines. Regarding higher order monetary aggregates, M2 falls by much less in the first phase. Specifically, M2 falls about 2 percent between November 1929 and December 1930, compared to a 34 percent decline in industrial hours. M2 is down about eight percent in August, 1931 compared to the 42 percent decline in industrial hours.

This means that about 80 percent of the decline in M2 occurs after this first phase of the Contraction; M2 is about 34 percent lower in March, 1933 relative to November 1929. Friedman and Schwartz attribute the large M2 decline during the second phase to a large increase in the currency-deposit ratio that they further attribute to bank runs. And the currency-deposit ratio also changes modestly during the first phase, increasing from .094 in

November 1929 to .112 in August 1931. But by March, 1933, the C/D ratio is .225, an increase of over 100 percent relative to November 1929 (or to August 1931). These data show that the large monetary declines emphasized by a number of economists occur largely after August, 1931. The monetary changes up through August 1931 do not appear to be sufficiently large to account for the Contraction.

There are two other facts that suggest monetary factors are not central for understanding the Contraction before August, 1931. One is that the money supply relative to real economic activity is high. Specifically, the money-output ratio rises considerably during the first phase, whereas it declined during previous downturns. The other fact is that agricultural hours worked is on or above trend. This is unusual during a downturn, because the farm sector is procyclical, and was considered to be very sensitive to monetary and financial shocks during the early 20th century.

3.2 Banking Shocks

Friedman and Schwartz (1963) and Christiano et al (2003) cite banking panics as a key factor for why the currency/deposit (C/D) ratio rose so much during the Depression, and how the higher C/D ratio drove down broad monetary aggregates. Bernanke (1983) also cites banking panics as an important factor, arguing that bank closings reduced financial intermediation services. However, the major impact of banking panics, operating through the C/D ratio and through other channels, was after August, 1931. The modest increase in the C/D ratio noted above is consistent with the view that the bank closings that occurred in the first phase tended to be small and did not have large aggregate effects. Moreover, the earliest banking episodes occurred after the Contraction was well under way. For example, the first banking episode discussed in Friedman and Schwartz involved a marginal bank in Nashville, Tennessee, Caldwell and Company, which closed in the Fall of 1930, a year after the Contraction began. Industrial hours had already declined by 35 percent by this time.

In addition, the most detailed studies of Depression bank closings (e.g. Elmus Wicker) conclude that those during the first phase of the contraction did not have quantitatively significant aggregate effects. Specifically, the impact of Caldwell and Company was felt largely in the Nashville area and was transient. Interest rates in Nashville rose about 100 basis points for 1 month, and then declined to their previous level. New York money market interest rates were unchanged. Moreover, the closing of Caldwell and Company was not a panic-based run on an efficient bank. Instead, Caldwell had a portfolio of non-performing loans, and the Tennessee banking regulator had suggested

for some time that Caldwell should have been closed.

There are two other facts that are consistent with bank closings and related banking issues having relatively modest effects in the first phase. One is that premia between risky securities (Baa) and short-term government securities rose relatively little during the first phase. Some economists argue that changes in risk premia are useful indicators of exogenous financial shocks, but these spreads rose less in the first phase of the Contraction than during a typical post-World War II recession. Specifically, the average increase in the spread between Baa yields and short-term Treasuries during a post-World War II recession is more than 200 basis points. Between November, 1929 and August, 1931, this spread rises by 165 basis points, with industrial hours down 42 percent. (Early data are from Banking and Monetary Statistics, Board of Governors, 1943).

3.3 Productivity Shocks

Cole and Ohanian (1999), Chari, Kehoe, and McGrattan (2006), and others have cited productivity as an important factor in the Contraction. However, this factor also is primarily operative after the first stage. Table 5 shows industrial labor productivity and the standard Solow measure of total factor productivity relative to trend. I use factor shares of 0.3 for capital and 0.7 for labor to measure TFP. Output is industrial production, and the inputs are manufacturing hours and the capital stock from the manufacturing sector (source: BEA) interpolated to the monthly frequency. The standard measure of industrial TFP declines by about 5 percent relative trend in December 1929, and then remains around 5 percent below trend through August, 1931. Assuming that the Solow measure is a useful measure of TFP for this period, this indicates what appears to be a permanent five percent decline occurring at the end of 1929 and continuing thereafter. Output per hour rises significantly throughout the first phase. I include this measure since there is some evidence of reduced capital utilization during the Depression, and output per hour may be a better measure of TFP in this circumstance than the standard Solow measure.[1]

There are two interpretations of these data, depending on assumptions regarding utilization. If utilization changes are small, then the one-time five percent drop in standard TFP can account for some of the decline in output and hours in the industrial sector in late 1929, but it cannot account for the very large and continuing decline after that. Alternatively, if capital utilization changes were significant, then TFP was not a major factor, as

[1]The Kydland-Prescott workweek technology will be used to provide an alternative measure of TFP with variable capacity.

the decline in capacity would result in TFP being close to or above trend throughout this period.

3.4 Taxes and Other Shocks

I have also examined changes in labor and capital income taxes. Joines (1981) reports average marginal tax rates on capital and labor income, both of these tax rates are unchanged through the first phase of the Depression (the average labor tax is four percent, and the average capital tax rate is about 25 percent). These tax rates do rise appreciably until the New Deal. Some economists have argued international factors, such as tariffs were important. The average tariff rate, measured as import duties divided by imports, rise from about 40 percent to 53 percent in the first phase of the Depression. Because the U.S. economy was nearly closed prior to the Depression, tariffs are unlikely to be a major factor. Crucini and Kahn (1996) quantify the contribution of tariffs and find them to be small, accounting for about 1-2 percentage points of the output decline.

I conclude from this examination that at a minimum it is a challenge for these standard shocks to plausibly account for the immediate and large decline in industrial hours that began in the Fall of 1929.

4 Understanding Labor Shedding

The key to understanding the Contraction is to account for the very substantial and rapid labor shedding that began in late 1929, in the absence of large negative shocks to money, productivity, or other standard factors. I organize this analysis by evaluating whether labor shedding is most plausibly accounted for by shifts in labor supply, or shifts in labor demand.

4.1 Labor Supply

I will assess the possible contribution of labor supply using the standard first order condition in the growth model that governs household time allocation. Recent studies of the Depression, including Cole and Ohanian (2001) Mulligan (2003) and Chari, Kehoe, and McGrattan (2006) show that there is a large difference between the left and right hand sides of this condition when it is calibrated to account for these variables just prior to the Contraction. Specifically, given preference parameters, and given actual consumption, labor, and the real wage, the parameterized marginal rate of substitution between consumption and leisure is about 70 percent below the real wage during

the contraction. This suggests that given the real wage, households should have been working more and consuming more.

There are two interpretations of the gap in this first order condition. One is that the first order condition is indeed satisfied, which means that some factor either significantly increased the marginal rate of substitution between consumption and leisure relative to the pre-Contraction calibration, and/or significantly decreased the return to working relative to the measured real wage. The alternative interpretation of the gap in this first order condition is that the condition is not satisfied, implying that households were constrained in selling their labor services to employers. This alternative interpretation implies that some factor(s) were impeding the normal operation of the labor market.

The evidence favors the alternative interpretation that households were unable to equate the marginal rate of substitution to the real wage. Simon (2001, Journal of Economic History) collects data on the supply price of labor from "help offered" ads, and shows that the supply price of labor is 40% below the wage paid, adjusting for selection and quality. In contrast, the supply price of labor and the wage were very similar before the Depression. This 40% difference between the supply price and the wage is fairly similar to the decline in wages and prices in the farm sector, in which hours worked was much higher. In addition, surveys show that non-farm entrepreneurs worked many more hours than employees in the Depression. These data are hard to reconcile with the interpretation that households equated the MRS to the real wage.

Abstracting from these data, I am unaware of direct evidence supporting the view that households satisfied this first order condition. This would require demonstrating substantial increases either in the marginal rate of substitution between consumption and leisure or that the perceived return to working was substantially less than the measured real wage. Possible factors shifting the real wage include a large increase in taxes, but taxes did not rise in the early 1930s. Factors affecting the MRS include preference shocks, shifts in the price of non-market time, such as changes in the value of home production, changes to the return to human capital investment, or changes to the return to search. I am unaware of any direct evidence indicating rapid, large changes to preferences or the return to human capital accumulation that would be sufficiently large to account for the decline in market hours. The Lucas-Rapping theory is consistent with a perceived increase in the value of search, but a challenge for the search explanation, in addition to the evidence of Simon (2001), is that hours worked for entrepreneurs, including those in the farm sector, were much higher than in the industrial sector, despite the fact that the return to working in the farm sector fell considerably more than in manufacturing.

Figure 4 shows monthly nominal and real average hourly earnings in manufacturing. (The data are originally from Beney (1936), and modified by Hanes, (1996), Journal of Economic History). The data show that both real and nominal wages rose in the early stages of the Depression. In contrast, the real wage in farming and the relative price of farming - a sector with hours worked on or above trend - both declined around 40 percent. I therefore assume that households were constrained in selling labor services to the manufacturing sector in the first phase.

4.2 Labor Demand

My assumption of corner labor supply for households means that a change in labor demand is key for understanding labor shedding. The standard labor demand condition implies that the real wage is equal to the value of the marginal product:

$$pF_h = w,$$

where p is the relative price of output, F_h is the marginal product of labor, and w is the real wage. The standard factors that change labor demand are therefore changes in the relative price or productivity. However, neither of these factors can account for labor shedding. Specifically, recall from section x that output per hour in the industrial sector (which is proportional to the marginal product of labor with Cobb-Douglas production) rose during the first stage of the Contraction. The relative price of output from the industrial/manufacturing sector did not change much; the implicit price deflator for fixed investment, for nonresidential fixed investment, and for structures investment, all relative to the GDP deflator, rise between one and two percent between 1929 and 1931. (NIPA, vol., 1, p. 148, February 1993).[2]

Regarding other factors that might affect labor demand, it is possible that manufacturers expected lower future demand for industrial output, though I am unaware of any evidence that could plausibly account for such a large and immediate decline in industrial labor beginning in the fall of 1929. In fact, inventory data suggest strong demand, not weak demand; Figure 5 shows the manufacturing inventory-sales ratio, which is available quarterly beginning in the first quarter of 1926, is at its low point just prior to the start of the Contraction.

[2]Wholesale price data are available for commodities and outputs in certain industries and in the agricultural sector. The wholesale prices of the industrial goods rise relative to those of agricultural goods. It is worth noting that many of these wholesale prices drop by more than prices for the deflators of final goods.

There is not only a large shift in labor demand, but a shift in the relationship between labor demand and the real wages. Between November, 1929 and August, 1931, output per hour rose about 10 percentage points more than the real wage in manufacturing, which means that the actual decline in labor exceeds the decline implied by the marginal product relationship. Given the real wage, standard theory implies that hours worked should have fallen about 25 percent, compared to the actual decline of 40 percent.

These data suggest an alternative theory of the shift in labor demand. Accounting for this shift thus requires a factor that reduces hours worked in the absence of a major productivity decline or relative price decline, and that also prevents the wage from falling. Thus any promising theory of labor shedding should account for:

$$\begin{aligned} pF_H &= w + x \\ w &\geq \bar{w}, \end{aligned}$$

where x is a factor that raises the opportunity cost of labor, thus reducing labor demand, and the inequality represents a floor on the industrial wage. I search for a theory that can account for these facts by first documenting labor markets during the 1920s, with a focus on labor relations and labor market policies, and I describe how those relations and policies changed in 1929.

4.3 Industrial-Labor Relations and Policies

This sub-section describes factors and summarizes data from the 1920s to understand labor markets before Hoover, the shift in labor demand that occurs in late 1929, and the apparent wage floor. I begin by summarizing 1920s labor markets and industrial labor relations.

4.3.1 Wages and Employment during the 1920s

It is widely agreed by labor historians that the 1920s was a decade in which non-union labor had little bargaining power. Bernstein (1963), which is the most detailed assessment of the 1920s labor market, titled his book the "lean years" in reference to low wage growth. There are two unusual features of wage Hanes and employment during the 1920s. Figure 6 shows real wages and output per hour between 1924 and October, 1929. Output per hour rises about 25 percent, while real wages rise about five percent. This implies that returns to capital received most of the productivity gains during this period. Not surprisingly, capital's share of income was at a maximum of

nearly 40 percent. The employment rate was low during the 1920s. The ratio of employment (including farm) to the working age population between 1900 and 1919 was between 54 and 55 percent. In 1929, this ratio was about 50 percent, which is an eight percent decline from the average of the previous 30 years (Source: Historical Statistics of the United States). Moreover, the job creation rate in manufacturing, which is procyclical, was very low in the 1920s, even lower than during the Depression. Understanding the 1920s labor markets thus requires understanding why employment was low and why wages did not grow with productivity.

Regarding wages, Bernstein and other labor historians cite firm unionization policies as a key factor contributing to low wages. Specifically, firms were able to suppress union organization using a variety of methods, including company unions and modest corporate welfare programs, claims that union leaders were communists and anarchists, and occasionally employing violent methods. Bernstein documents several union organization attempts and strike attempts during the 1920s, and shows that tactics used by firms included kidnapping union organizers, firing workers who met with organizers, evicting strikers from company-owned homes, denying medical care to striker families from company-directed health providers, arresting and prosecuting union organizers and strikers, and shooting strikers. On some occasions, firms were able to buy local police, and also hired private police forces Firm actions were often not prosecuted, but union actions often were. Union organization attempts during this period typically failed.

4.3.2 Industry's Ability to Suppress Unions and Unionization Policy

The largest organization attempt was in 1919, when over 350,000 workers struck against the steel industry for higher pay, shorter hours, and union recognition. In Pennsylvania and Indiana, state police and the National Guard beat strikers, steel firm security guards were deputized, strikers were jailed on false weapons charges and would be released if they agreed to sign a statement disavowing the union. Eight strikers were killed. Steel workers tried again in December 1919, striking for union recognition. The strike failed and several hundred strikers were deported to Russia for being "anarchists", "communists", and "agitators". Other episodes in this period include several coal strikes in West Virginia. The U.S. Army intervened on four separate occasions in West Virginia coal mines, (www.wvculture.org/history/journal_wvh/wvh50-1.html), including an armed rebellion of 10,000 coal miners in 1920 following the killing of a union organizer (www.lutins.org/labor.html). In June 1922, 36 miners were killed in a coal strike in Illinois. In 1927, picketing miners were killed in Columbine, Colorado. "Palmer raids", named after Attorney

General Palmer, were conducted in the early 1920s, and involved government agents taking labor leaders into custody, and prosecuting some of them for anarchy.

Micro evidence indicates that firms that suppressed unions were profitable. Following the failed organization attempt of steel workers, U.S. Steel profits advanced 100 percent between 1923 and 1928, but wages grew two percent. Toledo Edison profits rose 61 percent, and wage rates were unchanged over the same period. American economists discussed low wages and their implications. For example, Paul Douglas remarked:

> "Failure of real wages to advance in the 1920s was the cause of rising profits and was an appreciable factor in the extraordinary increase of stock market values." (Bernstein (1963)).

While non-union wages were roughly unchanged during the 1920s, union wages, and union wage premia, rose considerably. Table 6 shows union and non-union wages. Union wages rise about 40 percent, while non-union manufacturing wages rise about 6 percent. American economist Frederick Mills argued that the threat of unions kept non-union wages from being any lower (Bernstein (1963)).

These data suggest that there were large incentives for firms to suppress unions. Firms were able to successfully do this because of several Court decisions which limited the effectiveness of union organizing efforts. These Court decisions are described in detail in Ebel and Ritschl (2006, in progress), which I draw on here. The Sherman Act originally treated unions and firms symmetrically regarding restraint of trade, and permitted court injunctions against unions. The Clayton Act of 1914 exempted labor from some aspects of the Sherman Act, which was followed by significant union growth and a large increase in strikes. However, this was followed by key court decisions that weakened labor considerably. The Duplex Printing versus ....is considered by legal scholars to be a key case that weakened unions by significantly restricting the extent to which the Clayton Act applied to unions. American Steel Foundries versus Tri-City Central Trades further weakened labor by effectively eliminating labor's right to picket, leaving organizers only the right to distribute information, and only far from the workplace entrance. A number of Labor historians note that these decisions considerably weakened labor, particularly their ability to organize.

4.4 Policy Shifts: Court Decisions and Hoover Aid Labor

By the late 1920s, labor policy began to change. Ebel and Ritschl (2006) cite key legislation and Court decisions that overturned the previous decisions that limited labor. The key legislation was the Railway Labor Act, which made collective bargaining at the company level mandatory and provided for state arbitration in labor disputes. Railroads attempted to obviate the Act by forming company unions. When the Texas and New Orleans Railroad replaced an independent union with a company union, the independent union sued and won. The decision was upheld by District Court in 1928, the Court of Appeals in 1929, and the Supreme Court in 1930. This case overturned previous rulings that upheld employer's rights against unions, and the case paved the way for the Norris-Laguardia Act of 1932 and the Wagner (National Labor Relations) Act of 1935.

Following the stock market decline of October 1929, President Hoover brought the leaders of the major industrial firms to the White House in November, including Henry Ford, Alfred Sloan (GM), Pierre Dupont, Gerard Swope (GE). The meeting is described in Hoover's memoirs

(http://www.ecommcode.com/hoover/eBooks/pdf/FULL/B1V3_Full.pdf),

and by his advisor Lamont (1930, Journal of Business). President Hoover asked industry to maintain current wage structures, as this would help keep the industrial peace:

"Our immediate duty is to consider the human problem of unemployment...our second problem is to maintain social order and industrial peace...a fundamental view (is) that wages should be maintained for the present...the industrial representatives expressed major agreement...the same afternoon I conferred with the outstanding labor leaders and secured their adherence to the program...this required the patriotic withdrawal of some wage demands..." (Hoover, pp 43-44). Lamont noted "One of the first things which they (business leaders) did was to agree in principle to maintain the level of wages to perpetuate industrial peace". I interpret the Railway Labor Act, the court decision that strengthened this Act, and Hoover's policy as significantly changing unionization policies and the ability of unions to organize industry.

5 A Model of Labor Shedding

This section presents a model of the 1920s policy, in which firms could easily suppress unionization and thus retain profits, and the shift to the 1929 policy in which court decisions and Hoover's policy permitted unions. The goal is

to develop a theory that captures the essential features that existed at that time: the pre-1929 market power of firms in the labor market, the 1929 policy shift that reduces that market power, the very different patterns of activity between the farm sector, and the apparent rationing of labor in manufacturing and a manufacturing wage floor.

5.1 The Environment and the Pre-Hoover Policy

There is a representative family with a unit measure of household members. Some family members are endowed with skills to work either in the industrial sector or the farm sector. The family's time endowment for those with skills to work in either sector is T_m. Other family members are only able to work in the farm sector, and the family's time endowment for pure farm work is T_f. The sum of these two time endowments is normalized to 1.

Preferences are given by:

$$\max E \sum_{t=0}^{\infty} \beta^t \left\{ \log(c_t) + \phi[\log(T_m - h_{mt}) + \log(T_f - h_{ft})] \right\},$$

subject to:

$$w_{ft} h_{ft} + w_{mt} h_{mt} + r_t k_t + (1 - \delta) k_t + \Pi_t \geq c_t + k_{t+1}$$

$$1 \geq T_m + T_f$$

The objects in the budget constraint are standard, though note that Π is profits from the manufacturing sector, which will be described below. The technology for the manufacturing sector is given by:

$$Y_{mt} = A_m F(K_{mt}, H_{mt})$$

The technology for the farm sector is given by:

$$Y_{ft} = A_f F(K_{ft}, H_{ft})$$

The final good is the numeraire. Final output is divided between consumption and investment and is given by:

$$Y_t = G(Y_{mt}, Y_{ft}) = C_t + X_t$$

where $F()$ and $G()$ are both linear homogeneous.
The law of motion for capital is:

$$\begin{aligned} K_{mt+1} &= (1-\delta)K_{mt} + X_{mt} \\ K_{ft+1} &= (1-\delta)K_{ft} + X_{ft} \\ X_t &= X_{mt} + X_{ft} \end{aligned}$$

There is a representative, competitive producer of farm output whose profit maximization problem is given by:

$$\max\{p_f A_f K_{ft}^{\theta} H_{ft}^{1-\theta} - w_{ft} H_{ft} - r_t k_{ft}\}$$

I now describe the maximization problem for an industrial firm. The model should be consistent with the large gap between the marginal product of labor and the real wage during the 1920s, and the low employment rate of the 1920s. I assume that the observed gap between the real wage and labor productivity during the 1920s is due to firms having market power in the labor market, and that the source of this market power is the ability of industrial firms to suppress unionization. This market power assumption suggests that firms were not price-takers in the labor market, and this is consistent with analyses of 1920 labor markets that describe industrial labor markets as uncompetitive. I assume that firms have access to a union suppression technology that can be used with a cost normalized to zero. The profit maximization problem, in which the firm chooses both hours and the wage, is given by:

$$\max_{H,w}\{p_{mt} A_m K_{mt}^{\theta} H_{mt}^{1-\theta} - w_m(H_{mt}) H_{mt} - r_t K_{mt}\},$$

Combining the first order conditions for hours and the wage yields:

$$F_{H_m} = w_m + H_m {w_{h_m}}^{3}$$

Note that in this formulation I assume that the industrial firm acts competitively in the product market. It is straightforward to relax this assumption, and this would be natural to do in a model in which there is imperfect competition among manufacturing firms for labor.

5.2 The Hoover Policy: Raising the Cost of Union Suppression

The Hoover policy shift means that manufacturing firms can no longer suppress unionization with the suppression technology that they previously used. Instead, firms are now susceptible to unionization, and they can reduce the probability of unionization by paying higher wages, and/or by reducing their relative size, as unions targeted large firms. The Hoover policy shift changes the manufacturing profit-maximization problem as follows:

$$\max_{\hat{w}_m, h_m} \{p_{mt} A_m K_{mt}^{\theta} h_{mt}^{1-\theta} - \hat{w}_m h_m - r_t K_{mt}\} \pi(\hat{w}_m - w_m(h_m), \frac{h_m}{H_m})$$

subject to:

$$\hat{w}_m \geq w_m(h_m),$$

where π is the probability that the firm is not unionized. Note that $\pi_w > 0, \pi_h < 0$.

In the absence of unionization, the firm retains profits. If unionized, I assume that profits are zero[4]. Assuming that $\hat{w}_m > w_m$,the first order conditions for hiring manufacturing labor and choosing the wage can be combined to yield:

$$F_{Hm} = \hat{w}_m + h_m w_h - \frac{[f(h_m) - \hat{w} h_m]}{h_m} \frac{h_m}{H_m} (\frac{\pi_h}{\pi}).$$

The Hoover policy affects the firm hiring decision in several ways. Assuming that $\pi_w(w(h))$ is sufficiently high, then the probability of unionization leads to a higher wage, $\hat{w} > w(h)$. The second term in the equation is also affected, as $\hat{w} > w(h)$ implies that households will be at a corner in terms of selling their industrial labor services, which implies:

$$\frac{\phi}{u_c(T_m - h_m)} < \hat{w}_m,$$

and also implies:

[4]I assume that profits are zero if the firm is unionized because it simplifies the first order conditions. It can easily be relaxed.

$$h_m w_H = 0 \quad ; \hat{w}_m > w(h_m)$$

Thus, paying a wage premium rations labor in the industrial sector, and this rationing reduces the opportunity cost of industrial labor by lowering the value of the term $h_m w_H$. The adoption of the Hoover policy adds a third term to the labor demand relation, $\frac{[f(h)-\hat{w}h]}{h}\frac{h}{H}(\frac{\pi_H}{\pi})$. This is a negative labor demand shifter, as π_h is negative (larger relative size increases the probability of unionization). The third term is profits relative to labor input, scaled by relative size $\frac{[f(h)-\hat{w}h]}{h}\frac{h}{H}$, and multiplied by the elasticity of the unionization probability with respect to size $(\frac{\pi_H}{\pi})$. This term implies that at the margin, firms could raise profits by increasing labor, but that this larger size reduces the probability that the firm will retain those profits.

5.3 Discussion and Quantitative Analysis

A theory of the first two years of the Contraction needs to account for the negative shift in labor demand, and the failure of industrial wages to decline. In the language previously adopted, the theory should provide an interpretation of the "x" factor, and account for the manufacturing wage floor:

$$\begin{aligned} pF_H &= w + x \\ w &\geq \bar{w}, \end{aligned}$$

The "Hoover theory" developed here does that by restricting the allowable technologies that firms can use to suppress unionization. The Hoover policy (and Court decisions) prevent the lowest cost suppression technologies from being used. Instead, firms can suppress unions by paying a wage premium, which is consistent with the apparent manufacturing wage floor observed through late 1931. The Hoover model accounts for the "x" factor by taking into account that unions targeted the largest firms. This size factor adds an additional term to the firm labor demand relation - $\frac{[f(h)-\hat{w}h]}{h}\frac{h}{H}(\frac{\pi_H}{\pi})$ - which tends to reduce labor, as larger size increases the probability of unionization. The model is qualitatively consistent with the early stages of the Contraction, and also presages actual changes in unionization, wages, and strikes. By the mid-1930s, unionization jumped to about 37 percent of the workforce, and unionization was probably much higher in the industrial sector. Unions clearly targeted large firms, including G.M and U.S. Steel. Unions used the "sit-down" strike, in which workers forcibly occupied factories and prevented

production. The sit-down strike was permitted by government. G.M workers took over the Fisher auto body plants and shut down production of this component. This briefly reduced G.M. output to zero. G.M. settled the strike by recognizing the union and raising wages. The threat of a sit-down strike led U.S. Steel, which was one of the most effective union suppressors in the 1920s, to settle with its workers and recognize the steel union following the G.M. sit-down strike.

Quantitative analysis (to be added) will address the following issues. Wage changes by firm size can be potentially used to understand the shape of the π function. The model can be used to assess the impact of the policy on the farm sector, as the hours worked remained on trend, despite a 40 percent drop in the relative price of farm product.

5.4 Hoover vs. Roosevelt: Comparing the Hoover Policy with Roosevelt's NIRA

Both Hoover, and Franklin Roosevelt - through the National Industrial Recovery Act (NIRA) - adopted policies that fostered high wages and lower employment. I now address what may have happened had Hoover adopted the NIRA plan of raising wages. The NIRA set up industry cartels that permitted industry collusion through "Codes of Fair Competition" that defined minimum prices and acceptable trade practices provided that industry negotiated high wages with labor. Cole and Ohanian (2004) develop an insider-outsider model of this policy, and assess the policy's consequences for employment, wages, prices, and output. In this model, insiders can choose to add workers, or reduce their number through a lottery. There is also exogenous decay rate of insiders. Cole and Ohanian showed that the adoption of the NIRA at a high level of employment does not immediately reduce employment. Since the optimal size of the insiders is less than full employment, the optimal choice for the insiders is to simply wait for worker attrition to reduce the number of insiders to the optimal size. This is because the insider-outside policy does not maximize rents, but rather maximizes rents per worker. Thus cutting employment to the optimal size means throwing away rents. Thus, with a low attrition rate, the impact of thr NIRA adopted at full employment reduces employment and output gradually. In contrast, equation x shows that Hoover's policy immediately reduces employment. The economic forces driving each policy are thus quite different, and this generates a significant difference between initial impact of Hoover's policy vs. the NIRA. This suggests that the adoption of the NIRA in late 1929 would not have

created the rapid labor shedding observed under the Hoover policy.[5]

6 Evidence: The Impact of Hoover's Meetings

Following their November 1929 meeting with Hoover, major industrial firms followed Hoover's recommendation to keep wages at current levels. A number of manufacturers, including Ford, even raised wages following this meeting (Henry Ford Memoirs). Moreover, labor shedding begins after the Hoover meeting. Many industrial firms kept their wage structures intact up until mid-late 1931, particularly the largest firms, who publicly advertised their cooperation with the President's wage request. Hoover indicated in his memoirs that he was very pleased with the wage policies of the major manufacturers, but that wage compliance was not as high among small firms.

Hoover wrote that he was satisfied that his "Wage agreement held up fairly well, and most of the non-union employers complied" (Hoover memoirs, p. 45). Hoover also prominently cited the approval of his policies by labor: "On October 6, 1930, William Green, president of the AFL said "the President suggested that peace be preserved in industry and that wages be maintained. The great influence which he exercised upon that occasion served to maintain wage standards...we appreciate the value of the service the President rendered to the wage earners of the country" (Hoover memoirs, p. 46) Labor leaders also appreciated Hoover's immigration policies, Hoover also noted: "In the 1931 convention of the American Federation of Labor, the executive council in its report again expressed appreciation of my efforts and of the substantial success. The AFL report noted "in the full year of 1930 there were only seven firms per hundred firms that had cut wages." (Hoover memoirs, page 46).

Shister (1947, AER) documents wage changes in manufacturing across firms of different sizes, and reports that wages were maintained by the largest firms, but that smaller firms cut wages significantly. This pattern of large wage cuts by smaller firms is consistent with the theory outlined above in which union organizers tended to target larger firms. In addition, Beney (1936) presents data on manufacturing and industrial wages, including railroad wages across a number of specific railroad occupations (e.g. passenger conductor, freight conductor, passenger engineer, freight engineer, etc.), as well as the average railroad wage. The railroad sector was highly regulated

[5] This does not imply that the NIRA is a beneficial policy, rather that it would have more gradual negative effects on economic activity than the policy adopted by Hoover.

by the government, and labor relations were governed by the Railway Labor Act of 1928, which provided substantial organizing power to independent unions and established collective bargaining in the industry. It is striking that changes in railroad wages were very similar to changes in wages in other manufacturing sectors. This is consistent with Hoover's conclusion that wage policies for union and non-union employees were similar.

The timing of industrial labor shedding, and the apparent industrial wage floor, coincides with Hoover's meeting with industry in November, 1929. Moreover, there are large changes in wages after a second industrial meeting with Hoover in late 1931 that are consistent with the breakdown of Hoover's influence on industry after that. Gerard Swope, who was the president of General Electric, met with Hoover in September, 1931. Swope indicated that industry has maintained wages as Hoover requested, and that industry required a policy that "stabilized industry". Swope asked Hoover to support the "Swope Plan", which was a blueprint for several elements of President Roosevelt's 1933 NIRA. The Swope Plan included policies to cartelize industry through trade associations with minimum prices, and regulation of output and competitive practices, all of which were central features of the NIRA. The plan had broad and diverse support, including other industrialists, such as Henry Harriman, the president of the U.S. Chamber of Commerce, Bernar Baruch, the President of Columbia University, the Dean of Harvard Business School, and legislators, including influential Senator Robert LaFollette. Hoover refused to support this plan, stating that it was "the most gigantic proposal of monopoly ever made in history." (Hawley, 1976)

Hoover's very negative reaction to the Swope plan was likely a surprise to industry. Himmelberg (1973) describes in detail how Hoover was very similar to Roosevelt, who adopted the NIRA. Throughout Hoover's tenure in government, particularly as Commerce Secretary in the 1920s under Coolidge, Hoover was a strong supporter and developer of industrial trade associations that fostered industrial cooperation to share information and avoid "dog-eat-dog competition".

Hoover's rejection of the Swope Plan, and more generally his refusal to consider other collusive plans, likely resolved any uncertainty that industry had about Hoover's intention to directly raise industry profits. The meeting ended with Swope advising Hoover that if there was no support for the Swope plan or an alternative plan, then industry would support the Democratic Presidential nominee in the 1932 election. Following this meeting, nominal manufacturing wages begin to fall substantially. Figure 7 shows nominal wages were starting to decline a bit in 1931, falling about three percent between January 1931 and August 1931. After this, nominal wages decline 10 percent over the following 11 months.

Hoover attempted to reduce "destructive competition" in natural resource

industries in 1932, but it was much too little and too narrow to reverse Hoover's standing with the industrial sector. Roosevelt was elected in November, 1932, and pushed the NIRA through Congress in March, 1933. The NIRA raised prices and wages in industrial sectors by about 25 percent.

7 After August, 1931

To Be Added...

8 Related Literature

A number of scholars have cited several of Hoover's policies, including the wage maintenance policy, as contributing to the Depression, though this older work is very different in methodology and data analysis. Recent discussions of Hoover's wage policy include Cole and Ohanian (2000, NBER Macro Annual), which cited the policy and its possible impact in the industrial sector. In independent and complementary work, Ebel and Ritschl (2006, in progress), discuss in detail the court decisions affecting labor that occurred between the adoption of the Sherman Act and the Depression, and also include a discussion of Hoover's wage maintenance policy. Their analysis differs considerably from this one. Ebel and Ritschl develop a bargaining model along the lines of Mortensen and Pissarides. In their analysis, they assume that prior to the Depression, there was individual bargaining, and after 1929, there was collective bargaining. In their model, the shift from individual bargaining to collective bargaining leads to much lower labor. Their conclusion that changes in labor market policies are central for understanding the Depression is similar to my findings.

9 Summary and Conclusion

Industrial hours and output declined enormously in the first two years of the Depression, before the large declines in monetary aggregates, banking panics, and productivity declines. This suggests that the monetary factors cited by Friedman and Schwartz, and Lucas and Rapping, the productivity factors cited by Cole and Ohanian and Chari, Kehoe, and McGrattan, and the banking factors cited by Bernanke, were not the central factors accounting for the decline through August, 1931, but rather were factors later during the Depression.

Understanding the first phase of the industrial Depression thus requires developing a theory for why industry shed so much labor despite no large productivity decline, and why industrial wages did not fall. A promising theory is a change in labor and unionization policies associated with Herbert Hoover that made industry susceptible to unionization. The central construct of the theory is a probability of unionization that is increasing in size, and declining in the wage. This is consistent with actual labor market changes of the mid-1930s when unionization increased substantially and the largest firms, including GM, were hit with sit-down strikes, in which workers forcibly occupied factories, prevented production.

The discussion here is related to Cole and Ohanian's analysis of the recovery failure. Specifically, the first phase of the Contraction is about a large drop in labor - not productivity, with no large negative shocks; the recovery failure was a failure of labor to recover, with productivity returning to trend and no large negative shocks. Cole and Ohanian (2004) attribute the recovery failure to unionization policies that fostered high wages and low employment. The analysis here suggests a very similar theme: Hoover's labor policy shift increased the probability of unionization, and this reduced employment and raised wages.

Table 1
Quarterly Fixed Investment (1929:3=100)

Date	Investment in Structures
1929:1	106.3
1929:2	98.3
1929:4	86.1
1930:1	80.1
1930:3	73.1
1931:1	52.8
1931:3	46.3

Table 2
Monthly Hours Worked Relative to 1929 Peak

Date	Mfg (Jul)	Autos (Feb)	Chem (Aug)	Shoes (July)
1929:1	95.9	98.2	91.7	102.2
1929:7	100.0	97.5	98.5	100
1929:9	97.7	89.1	97.6	104.2
1929:10	97.6	84.7	97.7	105.5
1929:12	89.5	55.3	87.0	103.9
1930:6	78.4	56.6	80.9	86.5
1930:12	65.5	47.1	72.2	75.8
1931:8	58.2	45.9	72.2	93.8

Table 3
Industrial Production
July, 1929 = 100

Date	Tobacco	Paper	Machine
Jan 1929	97.8	99.1	85.5
Sept 1929	92.7	101.3	98.6
Oct 1929	95.0	101.6	95.7
Dec 1929	90.4	100.5	87.7
June 1930	89.4	93.5	75.4
Dec 1930	95.0	87.4	57.3
Aug 1931	88.7	86.4	46.4

Table 4
Industrial Production

Date	Metals (July=100)	Fuel (July=100)	Dur Mfg (July=100)
Jan 1929	88.6	98.6	92.7
Sept 1929	92.6	98.5	95.6
Oct 1929	88.6	97.6	92.9
Dec 1929	71.8	94.0	78.4
June 1930	69.8	87.7	72.5
Dec 1930	50.3	78.7	54.5
Aug 1931	37.6	71.9	44.0

Table 5
Industrial TFP & Output per Hour Relative to Trend
(Nov, 1929 = 100)

Date	TFP	Y/H
Dec 1929	95.5	98.0
Mar 1930	96.4	101.4
June 1930	98.0	105.2
Dec 1930	93.0	104.9
Mar 1931	97.2	110.4
June 1931	96.0	111.0
Aug 1931	94.7	110.7

Table 6
Union and Non-Union Wages

Source: Hanes, 1996 and Historical Statistics of United States

Date	Union	Non-Union
1926	101.0	48.4
1925	98.9	49.3
1924	97.0	50.2
1923	91.3	49.1
1922	87.3	44.3
1921	92.1	46.7
1920	88.4	56.1
1919	70.4	44.8

Figure 1 - U.S. Industrial Production
January, 1929 - August, 1931

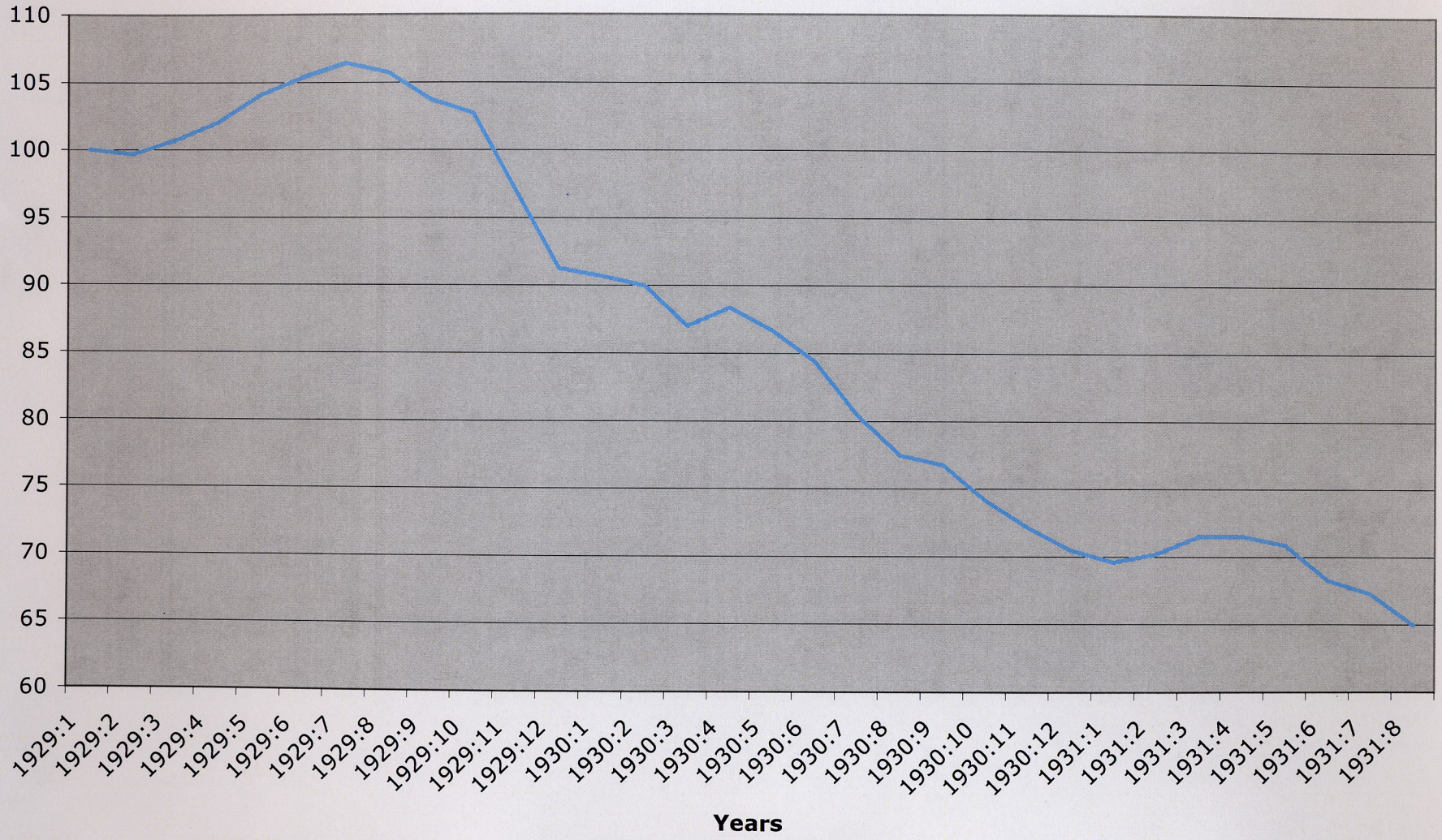
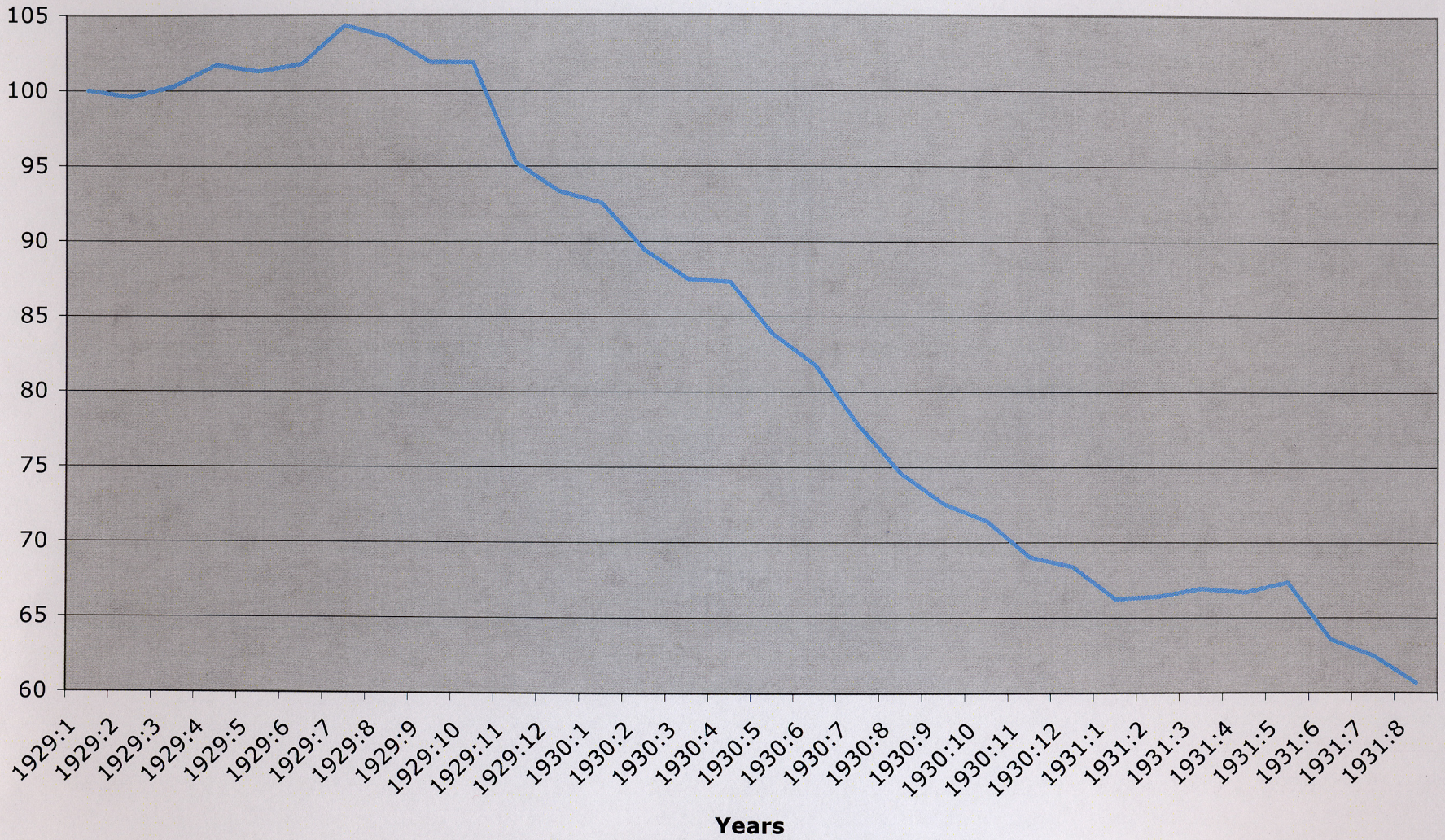


Figure 2 - U.S. Manufacturing Hours
January, 1929 - August, 1931



Real GNP (solid) and GNP from Christiano et al Model in Response to Monetary Policy Shocks (dash), 1928-33

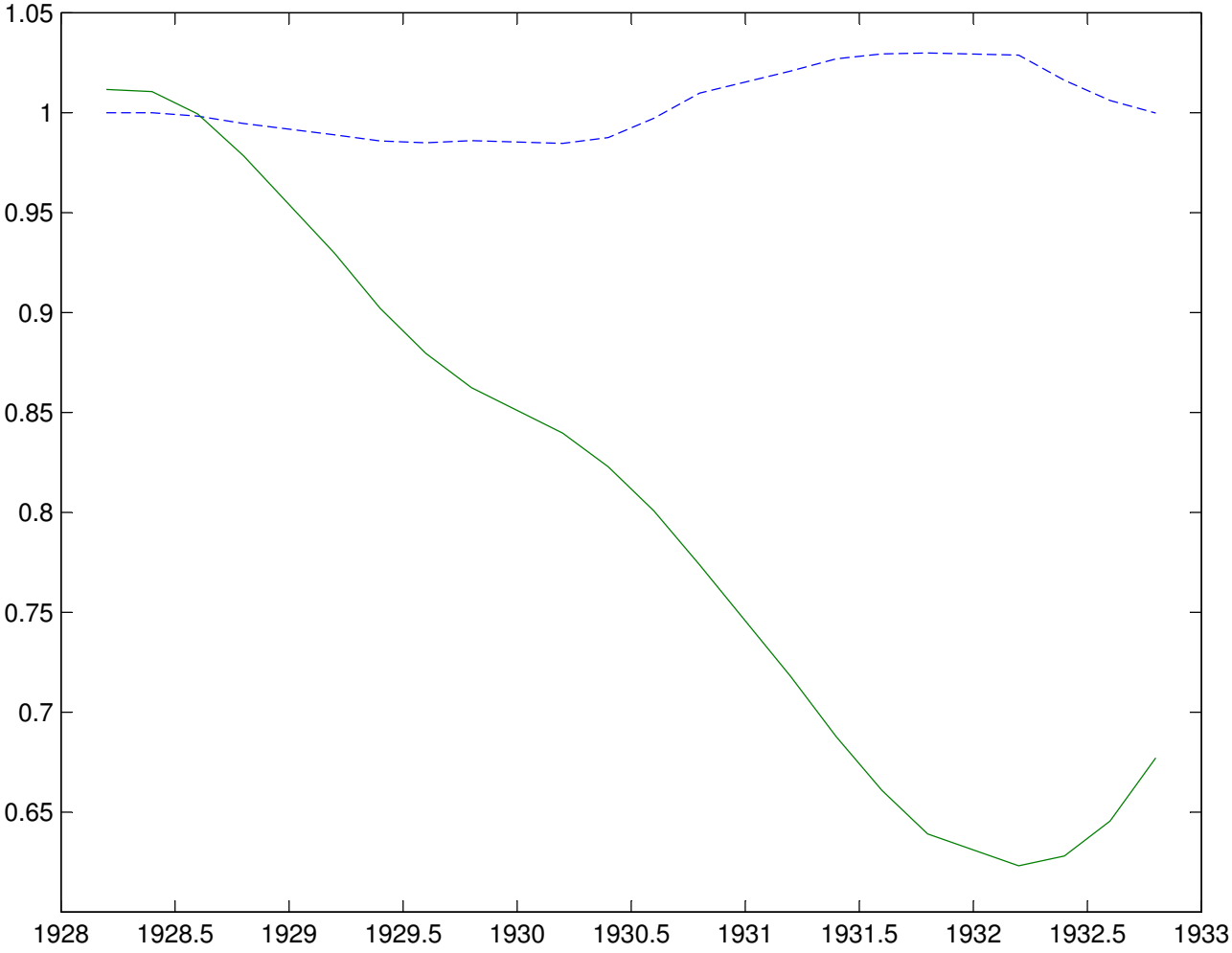


Figure 4 - Mfg Nominal and Real Wages
May, 1924 - August, 1931

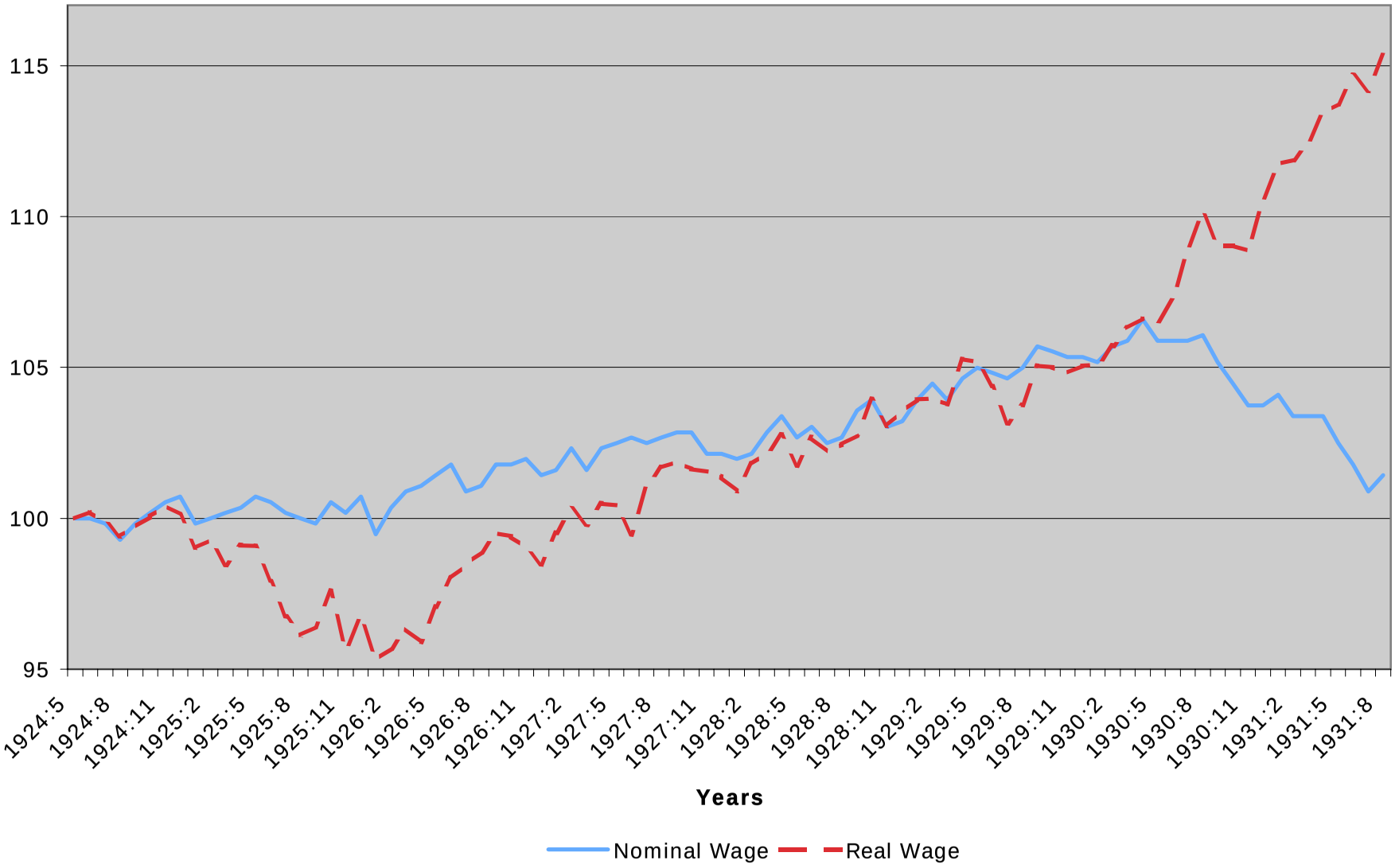


Figure 5 - Quarterly Mfg Inventories to Industrial Production
1929: 1 - 1929:3

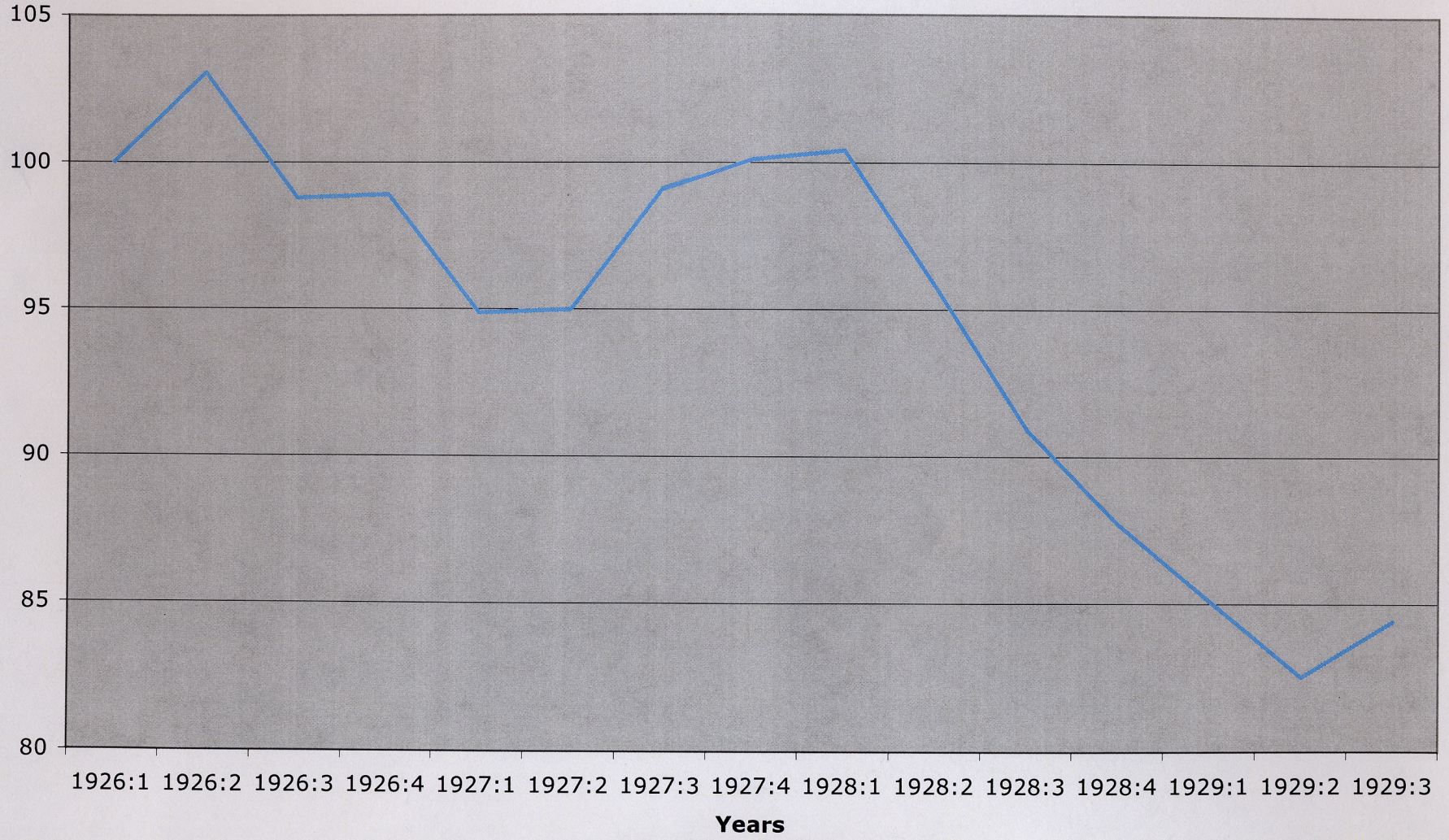


Figure 6 - Mfg Output per Hour and Mfg Real Wage
May, 1924 - August, 1931

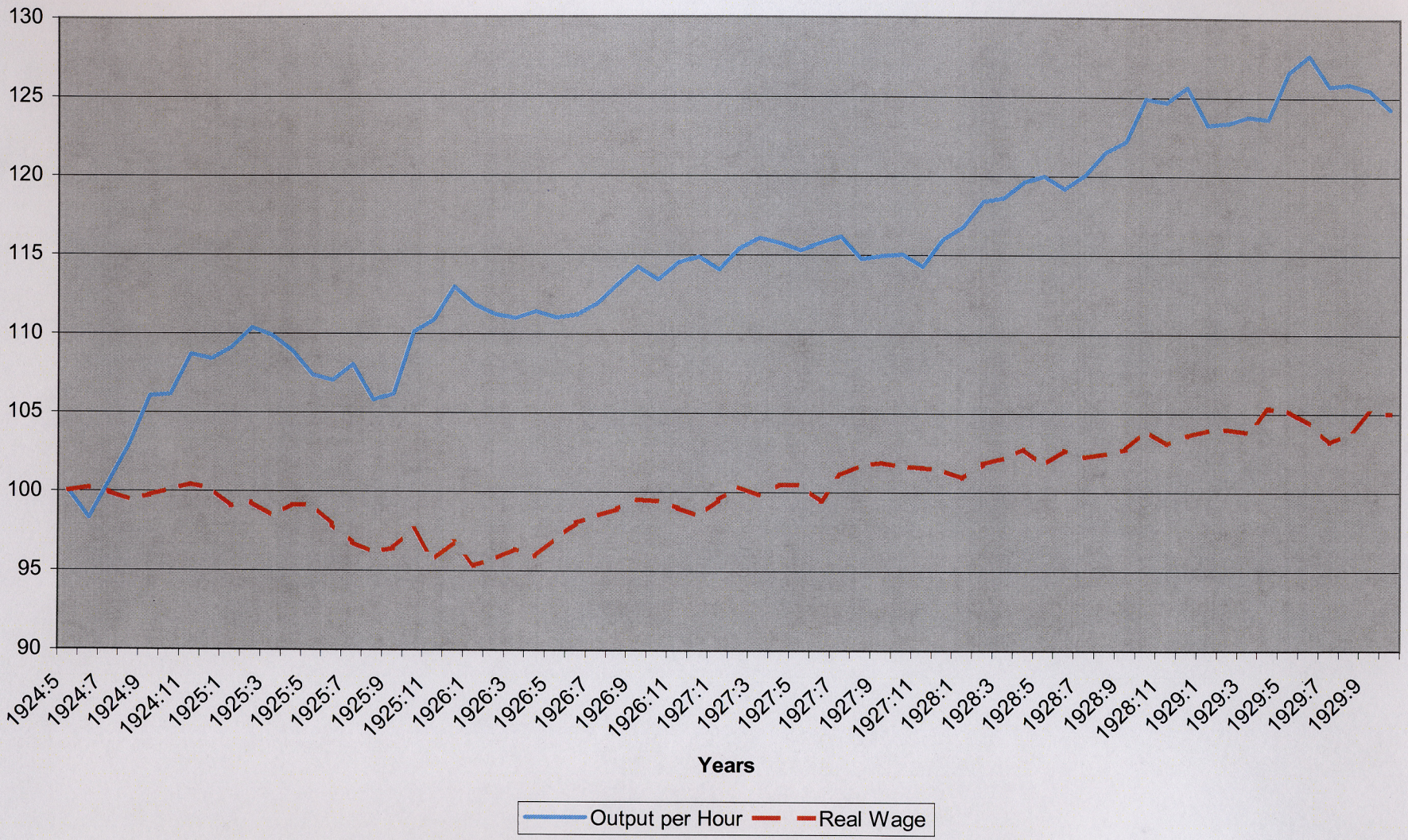


Figure 7 - Mfg Nominal and Real Wages