Controversies about the Rise of American Inequality: A Survey

Robert J. Gordon, Northwestern University and NBER
Ian Dew-Becker, Harvard University

NBER Working Paper Submission, April 21, 2008

*We are grateful for research assistance in locating references to two Northwestern undergraduates, Bobby Krenn and Neil Sarkar, and for helpful comments from David Autor, Polly Cleveland, Xavier Gabaix, James Heckman, Steve Kaplan, Larry Katz, and Robert Topel. A much shorter version of this paper has been published as Gordon and Dew-Becker (2007).
ABSTRACT

This paper provides a comprehensive survey of seven aspects of rising inequality that are usually discussed separately: changes in labor’s share of income; inequality at the bottom of the income distribution, including labor mobility; skill-biased technical change; inequality among high incomes; consumption inequality; geographical inequality; and international differences in the income distribution, particularly at the top.

We conclude that changes in labor’s share play no role in rising inequality of labor income; by one measure labor’s income share was almost the same in 2007 as in 1950. Within the bottom 90 percent as documented by CPS data, movements in the 50-10 ratio are consistent with a role of decreased union density for men and of a decrease in the real minimum wage for women, particularly in 1980-86. There is little evidence on the effects of imports, and an ambiguous literature on immigration which implies a small overall impact on the wages of the average native American, a significant downward effect on high-school dropouts, and potentially a large impact on previous immigrants working in occupations in which immigrants specialize.

The literature on skill-biased technical change (SBTC) has been valuably enriched by a finer grid of skills, switching from a two-dimension to a three- or five-dimensional breakdown of skills. We endorse the three-way “polarization” hypothesis that seems a plausible way of explaining differentials in wage changes and also in outsourcing.

To explain increased skewness at the top, we introduce a three-way distinction between market-driven superstars where audience magnification allows a performance to reach one or ten million people, a second market-driven segment consisting of occupations like lawyers and investment bankers, and a third segment consisting of top corporate officers. Our review of the CEO debate places equal emphasis on the market in showering capital gains through stock options and an arbitrary management power hypothesis based on numerous non-market aspects of executive pay.

Data on consumption inequality are too fragile to reach firm conclusions. We introduce two new issues, disparities in the growth of price indexes and also of life expectancy between the rich and the poor. We conclude with a perspective on international differences that blends institutional and market-driven explanations.

Robert J. Gordon
Stanley G. Harris Professor in the Social Sciences
Department of Economics
Northwestern University
Evanston IL 60208-2600
rjg@northwestern.edu

Ian Dew-Becker
Graduate Student
Department of Economics
Harvard University
Cambridge MA 02138
ian.dewbecker@gmail.com
# TABLE OF CONTENTS

1. **Introduction** .......................................................... 1

2. **Facts on Labor's Share** ........................................... 3

3. **The Basics of the Bottom 90 Percent** ....................... 6
   3.1 Facts on Changes in the Bottom 90 Percent ................. 6
   3.2 Mobility across Decades and Generations ................. 8

4. **Institutions and the Bottom 90 Percent** .................. 10
   4.1 Reversing the Great Compression .......................... 1
   4.2 The Role of Unions ........................................... 11
   4.3 Imports and Globalization .................................. 12
   4.4 Immigration and Inequality .............................. 14
   4.5 The Erosion in the Real Minimum Wage ................. 15
   4.6 The Role of Progressive Taxation ....................... 16

5. **Skill-biased Technical Change** ............................. 17

6. **Increased Inequality at the Top** ........................... 19
   6.1 Making Distinctions Among the Mechanisms at the Top .. 19
   6.2 The Economics of Superstars ............................. 19
   6.3 How Much Top Income can be Traced to Particular Occupations 20
   6.4 The Conflict among Hypotheses ......................... 21
   6.5 Firm-level Models of CEO Pay ........................... 25
   6.6 Remaining Issues about High Pay ....................... 28

7. **Consumption Inequality** ....................................... 29
   7.1 Findings and Problems with the Consumer Expenditure Survey 29
   7.2 Do Price Deflators Rise Unequally? ...................... 32
   7.3 Inequality in the Growth of Life Expectancy ............. 34
   7.4 Geographical Inequality .................................... 36

8. **International Comparisons** ................................ 38
   8.1 Facts on International Differences in Inequality ....... 38
   8.2 Explanations of the Facts .................................. 40

9. **Conclusion** .......................................................... 43

References ............................................................... 47
1. Introduction

Increased American inequality, particularly increased skewness at the very top of the income distribution, has received enormous attention in the last few years. There have not only been debates about the causes of the changes in the income distribution over time, but also about the simple facts to be explained. People argue about differences between data sources, about the causes of sinking relative incomes in the middle and bottom percentiles, and have especially contentious disagreements about the interpretation of the leap of relative incomes at the top, particularly in wage and salary income.

This paper provides a comprehensive survey of seven aspects of rising inequality that are usually discussed separately: changes in labor’s share of income; inequality at the bottom of the income distribution, including labor mobility; skill-biased technical change; inequality among high incomes; consumption inequality; geographical inequality; and international differences in the income distribution, particularly at the top.¹

Our initial work on inequality (Dew-Becker and Gordon, 2005) started from an attempt to understand the differential between the growth of mean and median labor income, which we accomplished by comparing data on incomes from IRS micro tax statistics with NIPA aggregate data on income and productivity. We documented an important fact, that over the period 1966–2001 only the top 10 percent of the income distribution had real compensation growth equal to or above productivity growth. Our results explained how the growth in average compensation had kept up with productivity over time, leaving labor’s share of national income largely unchanged, while the growth rate of median income had fallen steadily behind.

This paper begins in Part 2, where we update the facts on the evolution of labor’s income share and ask whether its decline since 2001 represents a normal cyclical response or remains within the context of long-run historical movements. We also look at changes in the income share of different components of capital income.

Part 3 examines the CPS data on the income ratios of the 90th, 50th, and 10th percentiles. This section also includes a survey of the recent literature on labor mobility. Increasing inequality would not be a serious problem if there were constant churning of individuals and households across the income distribution. Yet the literature that we survey suggests that income mobility across quintiles, either over decades or across generations, has been stable and if anything may have decreased.

¹. Outside the scope of this paper are aspects of the rate of return to higher education and the evolution of the college wage premium; this is covered in the companion paper by Goldin and Katz (2007).
Part 4 covers hypotheses regarding the evolution of relative incomes in the bottom 90 percent. Our point of departure is the “The Great Compression” interpretation of Goldin and Margo (1992). Three factors worked together in the middle of the twentieth century to equalize incomes: the rise of labor unions, the decline of foreign trade, and the virtual disappearance of immigration. Increased inequality since 1970 has been linked to the reversal of each of these three factors. Here we survey some of the recent debates about the quantitative impact of unions, free trade, immigration, and in addition about the evolution since the late 1970s of the real minimum wage. We find intriguing patterns over time in the movement of the CPS 90-50 and 50-10 ratios and link these movements to particular hypotheses and the recent literature on each. We also acknowledge other research that points to a decline in tax rates on high-bracket incomes as a source of increasing inequality.

Part 5 assesses the hypothesis that the primary driver of increased inequality was skill-biased technical change (SBTC). We place this in a separate section because it is relevant both to the bottom 90 percent and top 10 percent of the income distribution. We find the standard division of the labor force into two boxes, “skilled” and “unskilled”, to be unhelpful, and we endorse the efforts by David Autor and his co-authors to broaden the number of categories from two to three or five. We summarize the convincing Autor et al. polarization hypothesis, which distinguishes between the top, middle, and bottom, and relate it to several puzzles about SBTC.

Then Part 6 tackles the most controversial issue, why have American incomes at the very top increased so much relative to incomes below the 90th percentile, and why have American top incomes increased relatively so much more than in other rich countries? In this paper we distinguish three types of top-level incomes. The first group consists of superstars, driven by the market demand for sports and entertainment stars where media have magnified the reach of the very top individuals, and where the effort of the worker is the same whether the audience is a single person or ten million people. Second, we identify a second set of market-driven incomes which do not share superstar elements of audience magnification, namely lawyers and investment bankers including hedge fund managers. Third are the CEOs who differ because their incomes are not driven by the market but by decisions of their peers, with the presumption of self-serving reciprocity. We devote particular attention to conflicting papers endorsing alternatively a market-driven and management-power-driven explanation of the explosion of CEO relative pay.

Part 7 treats several issues related to consumption inequality. We survey the literature on inequality in the growth of market consumption and find it wanting due to inadequate data. We suggest that price index bias may overstate the rise of consumption inequality due to a relative price decline for the poor due to the Wal-Mart effect and a relative price increase for the rich due to their large share of
consumption of services that perenniably experience increases in relative prices. Perhaps offsetting this is an increased skewness in health outcomes, with a startling increase in the life expectancy of the rich relative to the poor. A final dimension is geographical inequality, the increased relative per-capita income of particular bi-coastal metropolitan areas relative to the rest. We link this development of “superstar cities” to the recent literature on the locational decisions of superstar scientists.

Any paper on issues in the rise of American inequality must devote substantial attention to cross-country differences. We review basic findings that document the continuing contrast between rising inequality in the US at one extreme and Japan and continental Europe at the other, with Canada and the UK falling between. In looking for explanations, we distinguish three sets of factors. The first is a broad-ranging set of institutional differences, including on the American side the early adoption of stock options as the most important component of the pay of top executives, compared with European institutions which involve cooperative bargaining that limits management prerogatives in trade for wage moderation, and compared with Japan where a separate cultural system of rigid seniority rules and flat pay schedules was combined with the illegality until 1997 of stock options.

The second set is the market itself, which determined that stock prices and price-earnings ratios would rise by multiples between 1980 and 2000, and this then mattered more for American executive pay because of the greater role of stock options. The third set goes beyond institutions and market-driven explanations by claiming that US corporate executives choose their own pay beyond any market-driven amount, a set of hypotheses alternatively called “managerial power” or “scratch my back”.

2. Facts on Labor’s Share

A frequent problem in the discussion of the income distribution is the confusion of the evolution of labor’s share in domestic income, as shown in the top frame of Figure 1, with the distribution of wage and salary income by percentiles. There is no necessary link between labor’s share in Figure 1 and the well-being of the median American who relies almost entirely on labor (as contrasted with capital) income. For instance, it is possible, at least in principle, for there to be a rise in the labor income of top-percentile earners (consisting of salaries, bonuses, and stock options that are treated as labor income), causing an increase in labor’s income share, while the relative income of the median wage earner declines.

With that qualification in mind, we can examine the plot of labor’s share and capital’s share in Figures 1a and 1b, which have been updated to 2007:Q2. The two lines in Figure 1a provide the key results, showing that one can reach almost any conclusion about changes in labor’s share, depending on the time period examined.
The lower line is the straightforward ratio of NIPA employee compensation (including fringe benefits) in the total economy to NIPA net national factor income, that is, GNP minus consumption of fixed capital minus indirect business taxes. Here we see that labor’s income share rose from 1997 to 2001 and then declined to 2007, but overall exhibits almost no net change from 1997 to 2007. The upper line adds an estimate of the ratio of labor to capital compensation in total proprietors’ income, an important component on the income side of the NIPA.

The bottom line shows a value for labor’s share that was very similar in years when the unemployment rate was close to the natural rate or NAIRU, i.e., 1987, 1997, and 2007. The upward movement in labor’s share from 1987 to 1990 and downward movement from 1990 to 1995 is very similar to the most recent cycle, up from 1997 to 2001 and then down to 2007.

2. The sharp jump upwards in labor income and downwards in capital income in 2005:Q3 reflects the sharp downward adjustment in the NIPA estimate of domestic income due to the effects of Hurricanes Katrina and Rita.
3. The estimate of the labor component of proprietors’ income is taken from Mishel et al. (2005, Table 1.24, P. 95). The fraction of total proprietors’ income that we use from this table is linearly interpolated between the selected years that they display.
Over a longer period of time starting in 1950, the bottom line shows that the share of labor compensation has increased substantially from 65 percent to 70 percent, and that there has been a negligible increase in the alternative measure of labor’s share including proprietors’ compensation from 72 to 73 percent between 1950 and 2007. Thus, to a first approximation, we conclude that the increase in American inequality after the mid-1960s has little to do with labor’s share in domestic income. What has happened is a sharp increase in skewness within labor compensation.

Figure 1b displays the historical evolution of capital’s share. While the overall share of capital income necessarily mirrors the bottom labor’s share line in Figure 1a, there have been differences in the evolution of the various sources of capital income. In the past decade, the share of corporate profits has risen while the shares of other components fell. The 12.5 percent share of corporate profits in national income 2006:Q3 was the highest since the 12.6 percent share in 1967:Q2. The overall share of nonlabor income at 29.7 percent in 2007:Q2 was roughly the same as in mid-1997 and lower than in any quarter during the period 1950–67.
3. The Basics of the Bottom 90 Percent

Corresponding to our (2005) finding that only the top 10 percent of the income distribution enjoyed gains in real wage and salary income equal to or above the rate of productivity growth, we define the lower 90 percent of the income distribution as “the bottom” and only the top 10 percent as “the top.” The 90 percentile borderline is convenient because the widely used data from the Current Population Survey (CPS) applies mainly to the income groups at and below that borderline, since top-coding limits the ability of the CPS data to provide information on income shifts within the top 10 percent.

3.1 Facts on Changes in the Bottom 90 Percent

Much of the literature on changes below the 90th percentile places major emphasis on the exact timing of those changes. Accordingly, we begin this section by following the literature that has examined changes over time in the ratios of real incomes at the 90th, 50th and 10th percentile levels, the so-called 90-50, 90-10, and 50-10 ratios. We begin with the log ratios (normalized to 1979 = 0) for men and women together and then the ratios for men and women separately. Figure 2 displays conflicting timing, with a steady increase in the 90-50 ratio after 1979, in contrast to a rise in the 50-10 ratio during 1980–86, followed by a slow and partial reversal after
1986. The 90-10 ratio, which combines these trends, shows a distinct increase between 1980 and 1988, followed by a plateau at between 20 and 25 percent above its 1979 level.

However, the combined experience of both sexes disguises differences in the evolution of men and women. For men as shown in Figure 3 the 50-10 ratio by 1998 had returned to its 1979 level after a temporary jump in 1979–86. The 90-50 ratio increased at almost a constant rate, by 12 percentage points between 1979 and 1990 and another 11 percentage points between 1990 and 2005. The overall male 90-10 ratio is similar to that for all workers, showing an increase until 1987 followed by a relatively steady plateau.

A big surprise, at least to us, is the pattern plotted in Figure 4 for women. While the 90-50 ratio followed the same basic path as it did for men, the 50-10 ratio increased much more for women, and the increase was permanent. The greater increase in the 50-10 ratio for women is consistent with a causal role for the minimum wage; women are roughly twice as likely to be paid the minimum wage as men (see Bureau of Labor Statistics, 2006). Perhaps most surprising is that the overall increase in the 90-10 ratio for women is almost double the increase for men, by 44 log percentage points compared to 23.

Figure 3. CPS Income Ratios by Percentile for Men Only
All of the numbers cited in this section understate the overall increase in inequality, because much of that increase has occurred within the top 10 percent of the income distribution. Because of top-coding, changes within the top 10 percent cannot be examined with CPS data but rather require data that are not top-coded. As we documented using IRS data (2005, p. 113), there is an important distinction between the 90th percentile income and the average income of those within the top decile, because increased skewness within the top decile makes that average income rises faster than the income of those at the 90th percentile. We showed that the income share of the top decile increased from 27 percent in 1966 to 45 percent in 2001, and that half of this represented the increase in the 90-10 ratio discussed above. The other half represented increased skewness within the top decline, that is, gains in relative incomes in the 95th, 99th, 99.9th, and 99.99th percentiles as compared to the 90th percentile. Piketty and Saez (2003) also provide a detailed analysis of the top quantiles.

3.2 Mobility across Decades and Generations

While inequality was increasing, income mobility was simultaneously decreasing, as is evident in a large literature on income mobility. Fields and Ok (1999) provide an extensive review of the methods of measuring mobility and their theoretical foundations. Gottschalk and Danziger (1997) provide a review of the literature on earnings and family income mobility. If cross-sectional inequality rises
but income mobility does too, then these two effects may offset each other leaving welfare roughly constant. This literature investigates this proposition and rejects it for the recent history of the United States.

Bradbury and Katz (2002) study transitions between income quintiles across successive one-decade intervals and find that a worker in the top or bottom 20 percent of the income distribution has a 50 percent chance of remaining in that quintile one decade later. On the other hand, there is only a 3 percent chance somebody will move from the bottom to the top or from the top to the bottom. In contrast, they find a large amount of churning among the middle 3 quintiles, which is to be expected given the year-to-year volatility in earnings. Gottschalk and Danziger (1997) find similar results looking at two-decade spans. They also find no upward trend in mobility that would mitigate increased cross-sectional inequality. If anything, they find that mobility has decreased in the last 20 years.

Kopczuk, Saez and Song (2007) provide another valuable view using Social Security Administration (SSA) data. The SSA data have the advantage of both eliminating the recall error inherent in surveys, and representing a large panel that follows people through their entire careers. They establish a number of important facts. First, they confirm that the changes in inequality that are found in family-level IRS data, even within the top 1 percent, are replicated in individual-level data. They interpret this as evidence that assortative marriage has not exacerbated measured family-level inequality. That is, the reason that family-level income inequality has risen is not that high-earning women are more likely than in the past to marry high-earning men. Second, they confirm the result that mobility has not increased for the average worker. Third, they are able to track mobility in and out of the top 1 and 0.01 percent. People in the top 0.1 percent have approximately a 65 percent chance of remaining there in the following year, a 55 percent chance of remaining there in three years, and a 45 percent chance of remaining there in five years. When they look at the source of the top 1 percent from ten years earlier, 35 percent are drawn from the top 1 percent and another 35 percent are from the 95th through 99th percentiles. Only 10 percent come from the bottom four fifths of the distribution. These probabilities are all roughly fixed over time. Overall, the income mobility research that focuses on people over decades or more provides no evidence that mobility has increased, and weak evidence that mobility has decreased.

Another interesting issue is that of intergenerational mobility. In a completely egalitarian society, one might expect there to be little connection between a person’s income and that of their children. On the other hand, if human capital is transmitted strongly from parents to their children, then income might be persistent across generations. The literature on income mobility generally cannot distinguish these effects; rather, it can only quantify mobility across generations. Solon (1999) and Bowles and Gintis (2002) provide extensive reviews of the literature on
intergenerational mobility. Hertz (2005) studies mobility among income quintiles across generations. He confirms the result from Solon and Bowles and Gintis that the intergenerational correlation in income is approximately 0.4. Moreover, he finds that this result is largely driven by black families. A black person born in the bottom quintile of the income distribution has a 42 percent chance of staying there as an adult, as opposed to only a 17 percent chance for a white person. Overall, the results for intergenerational mobility are similar to those for individuals—people who begin life with low income are likely to stay that way, and this has changed little over time.

4. Institutions and the Bottom 90 Percent

Now we turn to the substantive issues in the evolution of the income distribution below the 90th percentile. We start with the three factors stressed as sources of Goldin and Margo’s (1992) “great compression,” namely the rise of unionization and decline of both trade and immigration that contributed to a compression of the wage distribution in the middle part of the 20th century, roughly defined as 1940 to 1970. Contributing to the post-1970 rise of inequality were the reversal of these three factors, that is, the decline of unionization, the downward pressure on low-skilled wages coming from a rising share of imports in GDP, and that coming from an increased population share of low-skilled immigration. Subsequently we go beyond a discussion of these three factors to examine the effect of changes in the real minimum wage and in the progressivity of taxation. While we do not explicitly treat the role of education in the development of the great compression and its subsequent reversal, we recognize the importance of the research by Goldin and Katz (1999) on the spread of American secondary education in the period 1910-40.

4.1 Reversing the Great Compression

Williamson (2006) broadens the historical scope of the great compression story by likening the rise of trade and immigration in the post-1970 economy with the late nineteenth and early twentieth century. In the 1870–1914 era, increasing trade and low-skilled immigration reduced the land-labor ratio in the new world and reduced the wages of unskilled labor relative to “landlords and skilled workers.” The impact of trade emphasized by Heckscher and Ohlin was reinforced by immigration.

In fact, it would seem that immigration was a much more important element in the earlier period than imports. Immigration averaged about 0.9 percent of the US population during 1900–1913, more than double the 0.4 percent for the sum of legal and illegal immigration reached in the year 2002 (Gordon, 2003, Figure 5, p. 268). In
contrast, imports were a surprisingly small 4.8 percent of GDP in 1907–11, as compared with 16.2 percent in 2005.5

Williamson (2006, Figure 1) documents a strong negative correlation between the initial 1870 real wage in a given country and the annual change in an equality index, the ratio of the unskilled labor wage to the average wage. Over 1870–1913, this index declined at 1.45 percent per year in the US, a recipient of immigrants, in contrast to an increase of about 1 percent per year in Sweden and Denmark, suppliers of emigrants. In a subsequent regression analysis, Williamson finds that immigration was a much more important contributor to changes of inequality than trade, and that several of the European countries that lost population through emigration had a reduction of inequality.

4.2 The Role of Unions

We now turn to the reversal of the great compression after 1970. The percentage of US employees in unions declined rapidly from 27 percent in 1979 to 19 percent in 1986, and then more slowly to 14 percent in 2005.6 Since the real minimum wage was declining at approximately the same time (see section 4.5), it is difficult to distinguish empirically between those two factors as a possible cause of the rise in the 50-10 income ratio plotted in Figures 2–4 above.

Card et al. (2004) present a comprehensive recent treatment of the relationship between unionization and inequality. They strongly support Freeman’s (1980) seminal paper in showing that unions tend to reduce wage inequality among men (but not women), because the inequality-increasing “between-sector” effect is smaller than the dispersion-reducing “within-sector” effects. Both these effects decline in importance when the skill composition of the labor force is taken into account. In the baseline results of these Card et al. that include corrections for the skill distribution, the decline in unionization in the US explains a relatively small 14 percent of the growth in the variance of male wages between 1973 and 2001 and none of the increased variance of female wages. Averaging these results for males and females, the decline of unionization might explain 10 percent of the increase in the 90-10 ratio.

Freeman (2008) has more recently returned to his earlier theme by arguing that the evidence is “compelling” that collective bargaining, i.e., unions, “reduce inequality of pay compared to pay in competitive markets” (p. 112). He suggests an interaction with inflation, using the example of the Italian scala mobile (indexed wages) that in an era of rapid inflation in the 1970s maintained wage increases at the top and bottom at roughly the same rate, since most of the wage increases were driven by uniform

6. Mishel et al (2007), Figure 3W, p. 182.
inflation and relatively little by unique sources of demand or supply for the top or bottom groups.

Mishel et al. (2007, Table 3.32, p. 182) quantify the union wage premium in 2005 as 28 percent for wages and 43 percent for total compensation including fringe benefits, reflecting the role of unions in achieving benefits coverage as part of their union contracts to a much greater extent than for the labor force as a whole. These authors concur with Card et al. and the previous literature that unions have the greatest effect in raising wages and benefits at the middle of the income distribution relative to the top and bottom, and that their greatest effect is on high school graduates. They suggest that the sharp decline in unionization in the 1979–86 period might explain a substantial part of the increase in the 90-50 ratio observed during that period and of the ratio in the wages of college graduates relative to high school graduates, and indeed in Figure 3 the increase in the 90-50 ratio for men is faster during 1980–85 than since 1985.

We will revisit the role of unions when comparing U. S. vs. European inequality; the stronger political power of unions in countries like Germany may lead to wage negotiations that prevent or damp the decline in the relative incomes measured by the 90-10 ratio.

4.3 Imports and Globalization

The share of nominal imports in US GDP increased from 5.4 percent in 1970 to 16.2 percent in 2005. A variety of sources, including Heshmati (2006) and Miller (2001), argues that unskilled labor embodied in imports is highly substitutable with domestic unskilled labor, and that the increase in the import share of GDP observed in the last few decades has contributed to the decline in the relative wages of unskilled workers observed since 1979.

Mishel et al. (pp. 171ff) trace multiple channels by which increased trade may increase inequality. First, even a balanced increase in trade would reduce manufacturing employment, as import-competing industries are more employment-intensive than export industries. The erosion of manufacturing jobs is exacerbated by the fact that the US has run large and increasing trade deficits over the past three decades. Second, the share of imported intermediate goods in US manufacturing has increased from 8 to 20 percent since 1979, and this outsourcing of intermediate goods production has been concentrated in the most labor-intensive processes. Third, lower prices made possible by trade have reduced the value of the marginal product of many domestic workers. Fourth, globalization has diverted investment from domestic facilities to foreign direct investment.
Feenstra and Hanson have written a series of papers on this topic (e.g. Feenstra and Hanson, 1998, 1999, and 2003). Their 2003 paper, in particular, argues on the basis of data for the 1980s that the impact of trade on inequality has been understated. They find that trade has an impact on the wage gap between high-skilled and low-skilled workers that is similar to that created by SBTC. Their case begins with the claim that the share of trade for the US has been understated due to the neglect of intermediate goods. Their figure for the share of merchandise trade in merchandise value added for the US tripled from 13.2 percent in 1913 to 35.8 percent in 1990, and presumably has increased substantially further since 1990.

Feenstra and Hanson’s results on the relative importance of trade versus SBTC are mixed. They find that for the 1979-1990 period, outsourcing drives between 15 and 25 percent of both the shift in demand towards skilled labor and the increase in the wage gap between production and non-production (relatively unskilled and skilled, respectively) workers. They do not shed much light on the impact of SBTC, which we study more below, but they consistently find that outsourcing can explain roughly one fifth of the rise in inequality during the 1980s and 1990s. During this period, they say, “outsourcing [by which they mean both trade in goods and services] contributed to changes in industry productivity and product prices that in turn mandated increases in the relative wage of skilled labor.”

In a new book, Lawrence (2008) echoes the Feenstra and Hanson results for the 1980’s, but argues that the impact of trade on inequality has been of declining importance. He cites a number of sources, Cline (1997) in particular, as confirming Feenstra and Hanson’s estimate that approximately 20 percent of the rise in inequality in the 1980’s could be explained by trade. However, he notes that over the past 25 years, there has only been a weak correlation between wages and import price pressures when the evolution of these variables is compared across industries.

Lawrence also takes a more detailed look at wage inequality within industries, finding that industries in the US (and also China and Mexico) that are most exposed to trade, and that have lost the most jobs to outsourcing, had relatively high and concentrated wages. Therefore, outsourcing, rather than lowering wages at the bottom of the income distribution, could have contributed for stagnation in the middle of the distribution.

The more important point that Lawrence makes is that the effects of trade on inequality have declined over time. As early as 1994, Wood (1995) predicted this pattern. As trade increases, the economy moves away from producing goods that

7. Feenstra and Hanson also review evidence from other research that outsourcing has had similar effects in other countries. Anderton and Brenton, 1997, study the UK; Head and Ries, 2000, find that trade with low-income countries raised the skilled wage in Japan.
compete with imports. The fact that imports seem to be now competing with relatively high paying manufacturing industries is evidence that the US has shifted away from unskilled labor-intensive production. As Chinese imports destroy the US apparel industries, other industries employing a higher fraction of skilled labor raise their share in GDP, including higher education, health care services, and exports of capital goods from such world-leading firms as Boeing, Caterpillar, and Deere. If imports of goods produced by foreign unskilled labor increase, there are fewer of these disadvantaged workers to be hurt, and wages should not be reduced to the same extent as they were in the past.

4.4 Immigration and Inequality

Annual immigration (legal and illegal) as a share of US population increased steadily from 0.13 percent in 1960 to 0.41 percent in 2002 (Gordon, 2003, p. 268). Immigration has accounted for more than half of total labor force growth in the US over the past decade (Orrenius and Zavodny, 2006). This has caused the share of foreign-born workers in the labor force to steadily grow from 5.3 percent in 1970 to 14.7 percent in 2005 (Ottaviano and Peri, 2006, p. 1). Since 1990, there have been more foreign-born workers than black workers in the US labor force. As we have seen, Williamson credits a wave of immigration in the 1870–1913 period with a large decline in the wage of unskilled workers relative to average workers. Does the contemporary evidence support a similarly large negative impact in recent decades akin to that of the late nineteenth century?

There is a large literature on the effect of immigration on wages of domestic workers. A complementary set of papers by Borjas (2003), Borjas and Katz (2005) and Borjas (2006) conclude that US domestic workers lost about 3 percent of the real value of their wages due to immigration from 1980 to 2000, and that this loss reached almost 9 percent for domestic workers without a high school degree (Borjas, 2003, Table IX, p. 1369). Their methodology is a reduced-form approach in which the wage of native workers is regressed on some measure of the share of new immigrant arrivals or foreign-born workers and other control variables.

An interesting recent paper by Orrenius and Zavodny (2006) develops data on legal immigrants from INS data and examines wage responses across three occupational groups, controlling for geographical area fixed effects and area-time interactions. For low-skilled blue-collar occupations, they find a statistically significant but very small impact of -0.1 percent on native-born wages for every 10 percent increase in the share of workers in a given occupation who are new...

8. See Kopczuk, Saez, and Song (2007), figure 10
9. For a broader view of worldwide immigration that reaches beyond the effect on the US income distribution, see Economist (2008).
immigrants. However, for professional occupations they find a significant positive effect. Thus, while their results for low-skilled workers are very small, their combined results suggest a significant contribution of immigration to increased inequality between professional wages and low-skilled blue-collar wages.

Ottaviano and Peri (2006) make a novel point about immigration. Low-skilled immigrants disproportionately take jobs and enter occupations already staffed by foreign-born workers, e.g. restaurants, construction, and landscape services, and thus their main impact is to drive down wages of foreign-born workers, not domestic workers.

The previous literature had noted the fact that among high school dropouts, wages of domestic and foreign-born workers were almost identical up to 1980, but by 2004, foreign-born workers earned 15–20 percent less. This had previously been interpreted as evidence of a declining skill level of immigrants, but Ottaviano and Peri claim that this shift is consistent with their interpretation of increased competition of immigrants with each other for job classifications in which they specialize.

The authors summarize their results by asking how much immigration accounts for the increased wage premium of college graduates vs. high-school dropouts in the 1990–2004 interval. They conclude that immigration explains only 5 percent of the increase in the college-high school dropout premium, and that immigration actually worked to reduce slightly the college-high school graduate premium. A loose end in this analysis is that they only examine the impact of immigration on native-born workers and not on the total income distribution.

Mishel et al. (2007) provide no new evidence on immigration but make two interesting comments that support the view summarized above that the effects of immigration, if any, are minor. First, they note that the unskilled did better in the 1990s than in the 1980s, even though the percentage of foreign-born workers doubled in the 1990s. Second, they note that the CPS 50-10 ratio for all workers actually declined slightly from 1989 to 2004, whereas increased pressure from immigration on the bottom deciles should have made that ratio increase. As shown in Figures 3 and 4 above, while the 50-10 ratio for men declined after 1989 back to its 1979 value, that for women did not, and this apparently permanent increase in the 50-10 ratio for women may reflect in part the impact of low-skilled immigration.

---

10. Mishel et al. (Table 3.4, p. 119) show an increase in the real wage at the 50th percentile for all workers (men and women) of 9.0 percent as compared to 12.5 percent in the 10th percentile.
4.5 The Erosion of the Real Minimum Wage

There is a contentious literature on the effects of the minimum wage on employment, but there is less evidence on its effect on wage inequality. Mishel et al. (2007, pp. 190-95) display the sharp decline in the real minimum wage (in 2005 dollars) from $7.23 to $5.09 between 1979 and 1989, then a two-step increase to $6.25 in 1997, then a further decline back to $5.15 in 2005.) The decline in the minimum wage relative to average hourly earnings over the same period is more gradual, from 45 percent in 1979 to 31 percent in 2005.

Card and DiNardo (2002) advocate the hypothesis that the erosion of the real minimum wage accounts for much of the increase of inequality as represented by the 90-10 ratio. They find an almost perfect negative correlation between the decline in the real minimum wage and the increase in the 90-10 income ratio, as most of these co-movements were concentrated in the 1980–86 period. However, as we noted in section 4.2, at the same time that the minimum wage was falling, there were other institutional changes that likely affected inequality.

As we have noted above in discussing Figure 4, the decline in the real minimum wage in 1980–86 coincides almost exactly in timing with the jump in the 50-10 ratio for women. And the fact that the 50-10 ratio increased more for women than for men is consistent with the fact that women are roughly twice as likely as men to be paid the minimum wage. Further supporting the view that the minimum wage was important and mainly mattered for women is the fact that the real minimum wage was the same in 1989 and 2005, and so was the 50-10 ratio for women.

A possible weakness in the minimum wage story is that in Figure 4 above there is no response at all in the female 50-10 ratio to the increase of the real minimum wage from $5.09 in 1989 to $6.25 in 1997 nor to its subsequent decline back to $5.15 in 2005. This raises questions as to whether the minimum wage is the single smoking gun of the large increase in 50-10 inequality in the early 1980s.

4.6 The Role of Progressive Taxation

In section 8 below, in discussing international differences in the rise of inequality, we recognize the role of institutions and social norms that limit executive compensation in some European countries. In a related hypothesis, Levy and Temin (2007) suggest that high top-bracket tax rates in the 1940s and 1950s represented a “clear signal to limit top salaries” (p. 21). They cite findings by Frydman and Saks (2007) that with year-2000 tax rates, the level of executive compensation would have been 35 percent higher in the 1950s and 1960s. Given that the ratio of CEO compensation to average worker pay has increased at least by a factor of ten (Mishel et
al. 2007, Figure 3Z, p. 203), the Frydman-Saks results explain only a small part of the puzzle of rising pay at the top.

There are several mechanisms beyond social norms by which lower top-bracket tax rates might increase top pay. Feenberg and Poterba (1993), as well as Gordon and Slemrod (1998), suggest that declining top-bracket tax rates in the Tax Reform Act of 1986 contributed to the rise in top-bracket labor income by reducing incentives to report income as corporate income (not a part of labor’s income share) and instead to report it as personal income. Another less convincing explanation is that high income individuals responded to the lower rates by reporting income that otherwise would have been postponed. This could account for shifting income from a given year to a year or two later, but not for the rise of top incomes that has persisted for three decades. Overall, we agree with the conclusion of Levy and Temin that the rise in the share of top incomes is not an artifact of changes in top-bracket tax rates but is a genuine phenomenon that we review in detail below.

5. Skill-biased Technical Change

Thus far the timing seems to support a major role in the rise of the 50-10 ratio in 1980–86 for men to unions and for women to the decline in the real minimum wage. But we have not yet found a smoking gun that explains the gradual steady increase in the 90-50 ratio for both men and women throughout the post-1979 period.

A central unifying hypothesis in the labor economics literature on inequality is the role of skill-biased technical change (SBTC) (See, e.g. Bound and Johnson, 1992; and Juhn, Murphy, and Pierce, 1991; Card and DiNardo, 2002, provide an extensive review of the literature). SBTC is based on a simple model in which two skill classes of labor are imperfect substitutes. Because both the relative wage and the relative quantity of college graduates have increased since 1970, the SBTC hypothesis concludes that there must have been a shift in the demand by employers toward more skilled workers.

One approach criticizes the SBTC hypothesis on grounds of timing. If technology means innovations in computers, these were fastest in the late 1990s but inequality grew fastest in the early 1980s. A complementary criticism is that the irregular increase of inequality, concentrated in the early 1980s, is not consistent with steady improvement of technology over the past several decades.

We are skeptical of these criticisms of the SBTC hypothesis based on timing. The slow and steady increase of the 90-50 ratio both for men and for women seems consistent with a cause that is slow and steady, such as the gradual increase in occupational skill requirements in response to steady technological change. A
consistent theme of the literature is that the 90-50 ratio has increased since the 1970s but that the 50-10 ratio has increased little if at all. We have already seen that this statement is accurate for men but not for women.

At least for men, whatever skills are favored by SBTC must refer to those of workers well above the 50th percentile. Thus what we are looking for, and what the SBTC literature has sometimes been vague about, is the nature of the skills that favor those at the 90th percentile and above but are lacking in, say, the 70th percentile and below.

Autor, Katz, and Kearney (2008), building on earlier work by Autor, Murname, and Levy (2003), adopt a simpler three-way distinction between a high group doing non-routine cognitive work (including CEOs, lawyers, investment bankers, professors, and doctors), a middle group doing routine repetitive work (bookkeepers, accountants, and some engineers and computer programmers), and a low group doing manual but interactive work (truck drivers, nurses, waiters). This distinction emphasizes that work at the top and bottom is inherently interactive and is less prone to outsourcing than the non-interactive middle jobs. SBTC has increased the demand for people in the top group.

This enrichment of the concept of SBTC helps to answer an objection we posed in our 2005 paper (p. 117), where we cited evidence showing that there was no relative increase in the starting salaries of engineering and science graduate BAs in the 1980s relative to humanities graduate BAs, and in fact the reverse was true. Further, there were no above-average wage increases for the occupational groups most directly involved with the development and use of computers, namely, “engineers” and “math/computer”. During 1979–97 fully half of the growth in the college-noncollege wage premium can be attributed to the increased relative wage of the group called “managers,” and only 17 percent to the computer-related occupational groups. The Autor et al. three way distinction would place computer programmers and many types of engineers in the middle rather than high category as jobs subject to outsourcing and not benefiting from a rapid growth of demand relative to supply.

The SBTC hypothesis is about the demand for skilled workers growing faster than the supply. Most of the above discussion is about the increase in demand, but Autor, Katz, and Kearney (2008) and other recent papers emphasize as well a slowdown in the rate of growth of the relative supply of college workers from 3.89 percent per year from 1960 to 1989 to 2.27 percent per year from 1980 to 2005. This slowdown is largely driven by the native-born, not foreign-born immigrants. They find that the slowdown in the growth of relative supply can fully explain the behavior of the college wage premium. We regard other aspects of the evolution of the college-high school wage premium as outside the scope of this paper, as it is covered by the Goldin-Katz (2007) paper.
6. Increased Inequality at the Top

We now turn to increased skewness within the top decile, e.g., the large increase in wage and salary income (not to mention capital income) at the 99.99th percentile as compared to the 90th. Our initial (2005) distinction was between superstars and CEOs, and we were perhaps the first to make that distinction. In this paper we broaden this to a three-way distinction between superstars in the sports and entertainment industries, high-paid lawyers and investment bankers who do not have the audience-magnifying properties of superstars, and the controversial additional category of CEOs and other top corporate managers.

6.1 Making Distinctions Among the Mechanisms at the Top

An important distinction between groups of high income individuals is that some incomes are “chosen by the market” and some incomes are “chosen by their peers.” In this paper we further subdivide market-driven high incomes between superstars in entertainment and sports which benefit from audience-magnification, from other market-driven professions, particularly lawyers and investment bankers, where incomes are often earned in partnerships where the total pie is created by the market. We distinguish these two classes of market-driven high income occupations from CEOs, where there is substantial ambiguity about the relationship between the pay outcome and the market.

6.2 The Economics of Superstars

The economics of superstars was invented in Rosen’s (1981) analysis. In brief, Rosen explained the extreme skewness in occupational categories dominated by superstars by particular characteristics of demand and supply. On the demand side, audiences want to see the very best talent, not the second-best, and so income is highly skewed because the ability of top superstars to fill large entertainment venues and to sell recordings is an order of magnitude higher than second-best stars. On the supply side, the performer exerts the same effort whether 10 or 10,000 witness the performance.

Our analysis of superstars recognized that increased superstar premia reflected a particular type of SBT. Rosen had suggested that a succession of innovations going back to the phonograph had increased the size of audiences who can hear a given performance and thus increased the incomes of superstars by many multiples. We suggested that the sharp increase of superstar incomes since the time of Rosen’s (1981) article reflected the further development of technology, including cable TV, rentable videotapes and DVDs, and downloadable music and video.
A second group of high-income individuals, like superstars, reflects the operation of the market. Important professions, especially top-rank lawyers and investment bankers, earn incomes that are determined by market demand for the services provided by their firms, whether an enormous law firm like Chicago’s Kirkland & Ellis or an investment bank like Goldman Sachs. These market driven professionals differ from superstars in that their product is not amplified by electronic media. They are still tied down by the need to meet in person with clients and to attend legal proceedings with adversaries, in person.

The third group includes CEOs and other top corporate officers. We distinguish CEOs because of the ample literature arguing that CEO pay is not set by the market, but rather is set by gifts from peer CEOs who sit on compensation committees, and because the bulk of CEO and top-executive compensation since the late 1990s consists of stock options which in a significant number of cases have been fraudulently manipulated.

Bebchuk and Fried (2004) provide a comprehensive set of evidence in favor of a managerial power hypothesis that drives top executive pay well above the market solution. This is supported by the empirical work of Bebchuk and Grinstein (2005), who found that the pay of the top five corporate officers in 1500 firms increased by almost twice as much between 1993 and 2003 as their explanatory variables could explain. These variables included sales, return on assets, and the market return on the firm’s stock. These authors report the striking fact that the ratio of top-five compensation to total corporate profits for their 1500 firms rose from 5.0 percent in 1993–95 to 12.8 percent in 2000–02.

6.3 How Much Top Income can be Traced to Particular Occupations?

Previous research before our 2005 paper had not made the distinction between the superstar and CEO models of top-bracket compensation. An initial question is how much of pay in the top quantiles accrues to superstars, how much to CEOs and other top corporate offices, and how much is left over to be explained by lawyers, investment bankers, and other high-income occupations.

Regarding CEOs and the other four top corporate executives whose compensation is reported to the SEC, Kaplan and Rauh (2007) show that of the 6750 executives in the 1500 firms for which they have data on adjusted gross income in 2004, about 1925 have incomes above the $3 million threshold for the top 0.01 percent. The sum of executive incomes in the top 0.01 percent is $18.5 billion, or 22 percent of the

---

11. This statement is based on the set of previous hypotheses summarized by Kaplan and Rauh (2006, pp. 4-5) and the fact that they highlight our paper aside as most closely related to theirs.
$82 billion in labor (i.e., W-2) income earned by the top 0.01 percent in our 2001 IRS data.

This percentage (22 percent) is much higher than that highlighted by Kaplan-Rauh, because they are asking a different question than we did. Our paper was explicitly about a comparison of labor income per hour with aggregate productivity, and hence our distributional analysis was based entirely on IRS microdata covering W-2 statements pertaining to employee income, not on the much broader category of Adjusted Gross Income (AGI) which includes capital and professional income and capital gains.

Because AGI income is much higher than W-2 income, the W-2 income of top corporate executives is a much smaller share of AGI (the Kaplan-Rauh result) than it is of W-2 income (our result). To put AGI in the denominator reduces the importance of the employment income of corporate executives by a factor of four; this denominator includes coupon-clipping rentiers, retired movie stars, and a large variety of non-employees who are enjoying large amounts of capital and capital gains income but who are not employees at all.

It is difficult to extract from the Kaplan-Rauh data exercise the amount of additional W-2 earnings that they have located in the categories of non-Execucomp top financial and non-financial executives, investment bankers, hedge fund investors, law partners, professional athletes, and celebrities. Lawyers’ partnership income, in particular, is included in AGI but not in W-2 income. Nevertheless, it appears that they have located high earners in the top 0.01 percentile roughly triple in number of the Execucomp top-five executives that we copied from Bebchuk-Grinstein’s count.\footnote{In Kaplan-Rauh’s Table 8a, the number of non-financial ExecuComp executives is 594 and “estimated financial top executives” (some but not all in ExecuComp) were 137, for a total of 731. Yet the total of Kaplan-Rauh’s count of individuals in the top 0.01 percentile including investment bankers and other Wall Street high earners, top-100 firm law partners, professional athletes, and top celebrities, comes to 2339, so they identified 3.2 times as many people as in the ExecuComp database.} If all of this was W-2 income, they would have accounted for 66 percent of the income of the top 0.01 percent (that is, 3.0 times the 22 percent share of corporate executives). Allowing for lawyers and others who are not paid W-2 income, we might wind up with a share of around 50 percent.

\subsection{6.4 The Conflict among Hypotheses}

As we previewed above, Bebchuk and Grinstein run regressions that can explain only about half of the increase in top-five executive pay from 1993 to 2003. They make much of the fact that the share of executive pay in total corporate profits doubled over that interval. An important criticism of their work is that executives care
about share prices, not profits. If there was a more rapid increase in share prices than in profits, i.e., in the price-earnings ratio over this period, then the ratio of executive pay to market capitalization would have increased less than the ratio of pay to profits.

However, there was only a modest 16 percent increase in the price-earnings ratio of the S&P 500 over the Bechuk-Grinstein interval of 1993–2003.\textsuperscript{13} One must look earlier to find large increases in the price-earnings ratio, e.g., a doubling between 1990 and 1999. Since price-earnings ratios clearly did increase enormously during much of the period of the big run-up in executive pay in the 1980s and 1990s, it is worth considering the simple equilibrium explanation of Gabaix and Landier (2008) that executive pay moves in proportion to market capitalization.

The Gabaix-Landier model is based on a set of theoretical distributions plus the assumption that the best CEOs manage the largest firms. Their stunning result is that the elasticity of CEO pay to aggregate stock market capitalization is predicted to be exactly unity. “The sixfold increase of CEO pay between 1980 and 2003 can be fully attributed to the six-fold increase in market capitalization of large US companies during that period.” They also derive a superstar effect, in that “a very small dispersion in CEO talent . . . justifies large pay differences.”

The first reason for skepticism is that the authors’ own data do not support their unitary elasticity. Their Figure 1 shows that over the interval 1970–2000, their preferred “JMW” index of executive compensation increased by a factor of 22 while market capitalization increased by a factor of roughly 8.\textsuperscript{14} This would be consistent with the Bechuk-Grinstein result, which is that for the quite different 1993–2003 period, identifiable factors like sales, profits, and stock prices can explain only about half of the increase in executive compensation.

More troubling is a second problem that the results driving the Gabaix-Landier hypothesis do not hold prior to 1970. Frydman and Saks (2007) study executive compensation going back to 1936. They run regressions similar to those of Gabaix and Landier to find the elasticity of CEO compensation to both own-firm and average-firm size. They confirm the result from Gabaix and Landier (2008) and Rosen (1992) that the cross-sectional elasticity of CEO pay with respect to firm size is 0.3. And they reject the hypothesis that there is a unit elasticity between CEO pay and average firm size. They find an elasticity closer to 0.1 for most of their sample, both including and excluding the Great Depression and World War II. They also find that the Gabaix-Landier result is not robust to measuring firm size in terms of sales rather than market

\textsuperscript{13} The increase in the price-earnings ratio is the inverse of the decline in the earnings-to-price ratio on the S&P500 from Economic Report of the President 2008, Table B-95.

\textsuperscript{14} Reading points off Figure 1, the 1970 and 2000 observations for the JMW index are 0.5 and 11, respectively, while the numbers for market capitalization are 0.8 and 6.7.
capitalization. When sales are used to proxy for firm size, the coefficient for 1976–2005 rises from 0.94 to 2.65, but the coefficient for 1946–1975 only inches up to 0.16.

We also tried to replicate Gabaix and Landier’s result for 1970–2005. Rather than simply taking the entire period as one regression sample, we ran rolling 20-year regressions. The rolling regression is a technique that allows us to track the evolution of the estimate of the compensation-firm size elasticity over time. Figure 5 plots estimates from the rolling regressions. The standard errors are large due to the small sample sizes and the use of robust standard errors. However, it is clear that there is a large change in the point estimate over time. The estimated elasticity rises from approximately 0.5 to 1.5. So a unit elasticity does not seem to be an accurate description of CEO pay even for 1980–2005.

An important puzzle about CEO pay is the discrepancy between large increases in the US and smaller increases in Europe and Japan. The cross-country empirical work in the Gabaix-Landier paper is very crude. Lacking market capitalization data, they regress median executive compensation on median firm net income and find an elasticity of only about 0.33 (their figure 2). Further, the US observation lies above the plotted line by about 50 log percentage points. Interestingly, this is similar to the regression from Frydman and Saks noted above, where the coefficient on firm size is estimated to be 2.65.

![Figure 5. 20-Year Rolling Regressions of CEO Compensation on Firm Size as in Gabaix and Landier's Table II](image)

Note: The x-axis lists the final year of the regression; standard errors reported are robust.
Hall and Liebman (1998) have also developed detailed results about the elasticity of CEO pay to market value and other measures of firm size. Unfortunately the date of their paper limits their sample period to 1980-94 and excludes the late 1990s when the stock market boomed and the share of executive compensation taking the form of stock options mushroomed. They distinguish between the salary and bonus part of CEO compensation from the stock and stock options component. Their study differs from several others by including in executive compensation changes in wealth from the stock already owned by CEOs beyond that granted as stock options in a particular year. They dismiss previous studies that included only salary and bonus compensation although they point out that the small elasticity to firm value doubled from 0.12 to 0.23 between the first and last half of their sample period. Once the changes in the value of stock holdings and stock options are added in, the elasticity rises nearly monotonically from 1.2 in 1980 to 3.9 in 1994 (Hall and Liebman, 1998, Table VIII). This, however, is not directly related to the Gabaix-Landier result. The Hall-Liebman regressions are mainly used to measure the incentive structure that executives face. It is not meant to help understand why long-term increases in average firm size would lead to similar increases in average CEO pay.

Nevertheless, it is clear that the relationship between firm size and CEO pay has not been stable over time. From this fact, we might come to one of two conclusions. First, there may be no connection between firm size and CEO pay, and all theoretical attempts at justifying such a relationship are misguided. If the distribution of firm size is separate from more structural factors such as the aggregate capital stock, then we might not expect there to be any relationship between firm size and CEO pay. On the other hand, if the average firm grows at precisely the same rate as the average quantity of capital per worker—e.g. if the number of workers per firm tends to be stable—then we would expect there to be a strong correlation between average firm size and the pay of every employee.

Alternatively, we could take seriously the idea that there is a causal relationship between average firm size and CEO pay. In that case, the question we need to ask is whether a unit elasticity would be predicted by a theoretical model, or whether a lower elasticity is more plausible. Gabaix and Landier posit a firm production function of the form,

\[ Y = S + C \times S' \times T, \]

where S is the size of the firm, C is a constant, and T is CEO talent. The parameter \( \gamma \) controls the returns to scale for a given CEO. If \( \gamma = 1 \) and the size of a firm doubles and it keeps the same CEO, then its output will also double. Gabaix and Landier therefore claim \( \gamma = 1 \) implies constant returns to scale and take it as their benchmark. This leads to the result that there should be a unit elasticity between firm size and CEO pay.
The problem is that when a firm doubles in size it will also be able to afford a more talented CEO. This implies that there are in fact increasing returns to scale. The most efficient outcome in this model would be for every firm in the economy to merge into a single gigantic firm! While one might assume away mergers, possibly with firms producing differentiated goods, this seems unappealing given that firms seem willing to merge with other firms that produce different products. Moreover, even if firms cannot merge, a marginal dollar of capital should always be invested in the largest firm. That is, the return to further investment is not equal across firms. Furthermore, it is trivial for an investor to figure out which firm will provide the best return because he just has to find the biggest one. What all of this means is that $\gamma$ should actually be less than one. This implies that the greater-than-unit elasticity observed in the 1990’s cannot be explained by the Gabaix-Landier model, and the pre-1975 data seem to fit the model better.

It is clear from the literature that increasing relative executive options pay is dominated by the role of stock options, the growth in the role of stock options, and wealth changes from CEO existing stock options. Hall-Liebman (1998) propose two alternative explanations; as in common with much of the above analysis, they juxtapose a market-driven explanation against a managerial power explanation. The first is that the use of stock options has increased so dramatically because corporate boards want to increase the relationship between pay and performance. After a rise in options as a share of total compensation, if the value of compensation to the CEO is to be held constant, average pay will have to rise since more risky income has less value. A weakness in this explanation is that stock option rewards are not structured in terms of a firm’s performance relative to the overall market, but rather CEOs receive large rewards related to the aggregate market, the equity premium, and macroeconomic conditions. Their second proposed explanation, which is complementary to the first, is that boards want to increase CEO pay and choose option grants as a “less visible” method less likely to incite stockholder anger.

6.5 Firm-Level Models of CEO Pay

The Gabaix and Landier model can be thought of as a partial equilibrium model of CEO pay. It involves firms that are differentiated only by size and does not consider any sort of bargaining. However, there is an extensive literature that studies exactly how CEO compensation is set, in particular, the interactions between CEOs, boards of directors, and shareholders (see Murphy (1999), for a survey of this literature). The classic principal-agent model treats directors as chosen by

---

15. This result does not even require mergers. Even if firms could not merge, the marginal dollar invested in the stock market would always have a higher expected return the larger the firm it was invested in.

16. Note that the marginal product of capital in the Gabaix-Landier model is $dY/dS=1+CT$ when $\gamma=1$, so the marginal product is strictly increasing in CEO talent.
shareholders, and then studies the optimal contracts they set up for CEOs. This framework requires that boards have only the shareholders’ interests in mind. As we have discussed above, this assumption seems at best implausible.

Bebchuk and Fried (2004) provide a substantial amount of evidence that the principal-agent setting cannot explain the salient facts about CEO pay. They propose an alternative model in which CEOs have control over boards of directors and are mainly restricted by an “outrage constraint” where shareholders retaliate if they perceive executive compensation to be excessive. The Bebchuk-Fried model is more convincing than the principal-agent model. Specifically, they provide evidence that CEO contracts are far from optimal, that CEOs control directors, and that directors put substantial effort into disguising the size of CEO compensation packages. Their proposal has been criticized, notably in the Chicago Law Review in an article by Murphy (2002) and a response by Bebchuk, Fried, Walker (2002).

The key assumption that directors are independent turns out to be highly questionable. To start, their pay is far from negligible—an average of $152,000 per year in the top 200 firms. While directors also usually own stock in the companies they oversee, presumably the amount they stand to gain from good governance is smaller than the salary they would lose if they were not renominated. Moreover, directors receive substantial non-salary benefits in the form of perks, or in business directed to their own firms. Also, as we have noted before, if a CEO is also on the boards of any of his directors, there are ample opportunities for tit-for-tat relationships. 17

Bebchuk and Fried also provide compelling evidence that CEO contracts are in no sense optimal. They begin with the example of stock options. First, CEOs often exercise and sell their options as soon as they are vested, long before they leave the firm, thus undoing any incentive effect. Second, adept use of specialized derivatives, known as “collars”, allows CEOs to undo the incentive effect of their options and lock in a set value. Most important, options are not indexed to the market as a whole. This means that when the market goes up, CEOs earn more even if their company performs below average. On the other hand, when the market goes down, boards of directors reindex out of the money options in order to insure that they still have an incentive effect. Bebchuk and Fried argue that this asymmetry is a way of claiming that high pay is actually incentive pay. If CEO contracts were optimal, stock options would be indexed to the market instead of rewarding them for factors completely out of their control.

17. A recent report by Lublin (2008) suggests that “directors, facing unprecedented pressure from investors, lawmakers, and regulators, are increasingly cutting back the pay-setting power of CEOs.” This article is consistent with the Bebchuk view that boards have previously exerted little or no control over self-serving CEO pay awards to themselves.
Finally, Bebchuk and Fried provide ample evidence that firms work to disguise the magnitude of CEO pay. If contracts are optimal, there is no reason to try to hide what CEOs earn. On the other hand, if there is an outrage constraint, then firms have every reason to hide what their CEOs earn. This type of camouflage is available because the financial press generally only quotes annual compensation, ignoring deferred compensation and benefits. Executives often receive lifetime healthcare, company vehicles, offices, and consulting contracts. Firms often also give loans to their CEOs at well below market rates. Similarly, they will pay interest far above market rates when CEOs defer their compensation. It is difficult to imagine that these sorts of compensation could actually be more efficient than simple cash payouts. What is clear, however, is that they are very effective ways of reducing the visibility of compensation.

It is clear from the literature that increasing relative executive pay is dominated by the role of stock options, the growth in the role of stock options, and wealth changes from CEO existing stock options. Hall-Liebman (1998) propose two alternative explanations; as in common with much of the above analysis, they juxtapose a market-driven explanation against a managerial power explanation. The first is that the use of stock options has increased so dramatically because corporate boards want to increase the relationship between pay and performance. They echo the argument that optimal incentives would not be tied to market performance. Their second proposed explanation, which is complementary to the first, is that boards want to increase CEO pay and choose option grants as a “less visible” method less likely to incite stockholder anger.

Weisbach (2007) questions whether there is any way to improve CEO compensation. He notes that the corporation has been a stunningly successful organization, and that even going back to Adam Smith there have been complaints similar to those of Bebchuk and Fried. We take Brandeis’ view that sunlight is the disinfectant and electric light is the best policeman. Simply airing these sorts of problems makes them more salient in shareholders minds, as we have seen in the past few years in the popular press. Moreover, disclosure of CEO pay, as in the SEC’s 1992 and 2005 rule changes, helps investors monitor the pay of executives.

The fact that there were complaints 230 years ago about executive compensation does not mean that shareholders today should ignore what their CEOs are paid. Rather, research like that of Gabaix and Landier, Hall and Liebman, and Bebchuk and Fried is important because it tells shareholders what to expect and where their outrage constraint should be set. If we were to take a very strong view of the

18. Core, Guay, and Larcker (2007) study press coverage of CEO pay and find that the press tends to focus on CEOs with “excessive” pay rather than simply high pay, but that firms rarely react to negative coverage.
efficient markets hypothesis, we might think that shareholders are already taking into account every available piece of information on CEO pay. Given the ample evidence that in fact shareholders do not take into account all available information, we believe that the simple act of making certain irregularities and disparities more salient can improve outcomes.

6.6 Remaining Issues about High Pay

While we have questioned the usefulness of the skill-biased technical change (SBTC) hypothesis in explaining the evolution of the 90-50 and 50-10 percentile ratios, there seems little doubt that technology has been a substantial part of the story at the top. The development of electronic media must have contributed much or even most of the increased relative incomes of sports and entertainment superstars. The ability of technology to shift Wall Street from million-share trading days in the 1970s to billion-share trading days since the late 1990s must also have contributed to the multi-fold increase in real incomes of investment bankers and share traders.

Why is there any need to distinguish between CEOs and our broadly defined set of superstars? The primary distinction is that CEOs are not paid like superstars by multiplying a ticket price times numbers of tickets sold (or for investment bankers the price per deal times the number of deals). Rather, CEOs, through compensation committees and inbreeding of boards of directors, have a unique ability to control their own compensation. Furthermore, if a director approves a higher compensation package, that may subsequently lead her to receive more compensation at her own firm. 19

While the market plays some role in how CEO compensation is set, there is clearly an arbitrary component. The fraudulent repricing of stock options, which made up more than half of CEO pay in the 1990s (Bebchuk-Grinstein, Table 4), has attracted particular attention. Recently, the Wall Street Journal (Maremont and Forelle, 2006) felt obliged to develop a vocabulary primer involving such words as “backdating,” “repricing,” “reloading,” “spring loading,” “bullet-dodging,” and “exercise backdating.” The old phrase the “golden parachute” for retiring CEOs has been supplemented by the “golden handshake” for newly arriving CEOs. 20

A surprising conclusion to this section comes from the CEOs themselves, who conclude that “we are overpaid” according to a recent article (Guerrera, 2007). These surveys of CEOs are very partial, since they do not address the issue of cross-country

19. A compelling statement of the managerial power or “scratch my back” hypothesis about CEO compensation is presented by Stein (2008).
20. A summary of the current relationship between compensation consultants and boards of directors is provided by Nocera (2007).
disparities discussed below. But they do suggest that pay is not aligned with performance, and that 80 percent of outside directors felt that CEOs were “overpaid.”

### 7. Consumption Inequality

While income inequality tells us about the year-to-year distribution of economic rewards, to understand the distribution of overall welfare, income may not be the best measure. Many authors have made the point that the life-cycle/permanent income hypothesis implies that consumption may be a better measure than income of the distribution of welfare. Specifically, if people can insure effectively against transitory income shocks, then as the variance of those transitory shocks rises, measured income inequality will rise, but consumption and welfare inequality will stay fixed.

#### 7.1 Findings and Problems with the Consumer Expenditure Survey

There are numerous data sources on income, but there are few good data sources on consumption. The most widely used data set is the BLS’s Consumer Expenditure Survey (CEX), which consists of annual data beginning in 1980 with a few sporadic surveys before then. The CEX produces two surveys, both an interview survey and a diary survey. The interview survey follows households through 4 interviews, each of which cover the prior three months of expenditures. It is designed to measure large or routine expenditures, such as mortgages, utilities and car purchases.

The diary survey asks households to record all purchases over two two-week periods. It is designed to measure expenditures on non-durable goods and services. An alternative to the CEX is the Panel Study on Income Dynamics (PSID). The PSID has been following a set of 8,000 households since 1968, and it obtains very detailed measurements of income, employment and health. However, the PSID only measures food consumption, which is a serious limitation.

The first study to use the CEX to measure the distribution of economic well-being is Cutler and Katz (1991). Their work covers a variety of topics, including poverty rates, income inequality and consumption inequality. As in much of the other work we have discussed, they use the CPS to measure income. They find that the Gini coefficient for income fell from 0.379 to 0.366 between 1963 and 1980, but then jumped up to 0.397 in 1984. This result is in contrast to that of Kopczuk, Saez and Song (2007), who find that the Gini coefficient for income rose monotonically since 1953 in the Social Security Administration data. For consumption, Cutler and Katz find that the Gini coefficient fell from 0.298 to 0.285 between 1960 and 1972, but then rose to 0.314 in

Cutler and Katz (1991) report that income inequality is greater than consumption inequality, which is consistent with the permanent income hypothesis. However, they also find a nearly perfect correspondence between income and consumption inequality, implying that, at least during the 1980's, increased cross-sectional income inequality was not driven by increased transitory shocks. Or, if it was, people were not able to insure against the new shocks. Between 1980 and 1984, they report that the Gini coefficients for both income and consumption rose by exactly the same amount, 0.033. The timing of the rise in the Gini coefficient matches similar results on the 90-10 ratio evident in Figures 2, 3, and 4 discussed above.

Following Cutler and Katz (1991), more research was done during the 1990s on consumption inequality. Blundell and Preston (1998) provide a good review of the literature that examines the structure of longitudinal income data, citing principally Moffitt and Gottschalk (1995), Buchinsky and Hunt (1999) and Gittleman and Joyce (1996). All three of these papers study permanent and transitory shocks to income and all find increases in the variance of both types of shocks. They also confirm the result that income inequality rose significantly during the 1980's. Attanasio and Davis (1996) replicate the Cutler and Katz results using the PSID, showing that consumption and wage inequality have followed the same path over time. They also find that these results are confirmed when looking across education and birth year cohorts.

An important addition to the literature is Slesnick (2001), who reports that during the 1990's consumption inequality did not rise, in sharp contrast to the path of income inequality that we have examined. Krueger and Perri (2003) confirm this result and provide a model of endogenous credit markets which fit with the view that in the 1990s transitory shocks to income rose, credit markets became more developed, and consumption was smoothed out over the life cycle.

Much of the literature and the popular press now takes the Slesnick and Krueger-Perri results as stylized facts.21 However, Attanasio et al. (2006) show that we should not yet close the book on consumption inequality in the 1990s. They first provide evidence raising serious questions about the accuracy of the CEX during the 1990's. They replicate results from Battistin (2003) showing that the CEX has actually measured declining consumption during the 1990's, and McCarthy et al. (2002) show that the CEX matches the BEA’s data on personal consumption expenditures (PCE)

---

21. See, e.g. the New York Times' “Economic Scene” column: “inequality of consumption... does not show a significant upward trend.” (Cowen, 1/25/2007) “It is hard to see the effects of increasing income inequality in how people actually live.” (Postrel, 11/7/2002).
badly in both the level and the trend. Garner et al. (2003) further analyze the causes for the gap between the CEX and PCE data.

The mismatch between the CEX and PCE is worrisome because, as Attanasio et al. note, Banks and Johnson (1998) find that the UK’s Family Expenditure Survey (FES) matches their national accounts data well. So in principle, there is no reason a survey cannot match national accounts data. Moreover, Garner et al. (2003) find that the consumption category that has the largest shortfall in comparison with the PCE data are the components most widely used for consumption studies—non-durable goods and services.

The most important finding of Attanasio et al. is that the two CEX surveys give very different results on inequality. The interview survey on nondurables consumption is found by Slesnick (2001) and Krueger and Perri (2003) to exhibit no change in inequality during the 1990s. On the other hand, the diary survey shows consumption inequality rising faster than income inequality.

Attanasio et al. therefore use the findings of McCarthy et al. (2002) to determine which of the two surveys best measures each consumption category. After creating an index using a combination of data from each survey, they find, as we would expect, a change in inequality roughly equal to the average of the changes in each survey alone. Between 1990 and 2000 they find that the standard deviation of log consumption rose by 5.4 percentage points, as opposed to the 1.0 percentage point found by Krueger and Perri (2003). In the CPS, the standard deviation of log wages rose by about 4 percentage points over the same period. Thus Attanasio et al. actually find a slightly larger increase in consumption inequality than income inequality.

The literature has thus far failed to address an important issue in the inequality of consumption. The most important component of consumer expenditure is housing, yet we have no data or analysis yet on the inequality of consumption of housing quantity and quality, as measured by square feet of internal space, square feet of external space (lot size), and quality as measured by cross-sectional differences in the price per square foot. International data show that American households occupy roughly double the internal square feet of area and roughly four times the external square feet of area as Europeans (Gordon, 2008, p. 39). It seems conceivable that the consumption of housing quantity and quality by Americans is much more equal than income. This issue is ripe for further research, in particular because the benefits of purchase by Americans of large exurban houses may be partly mitigated by long commuting times.

22. Notably, Attanasio et al. (2006) and Battistin (2003) use a technique similar to that of the FES to combine data from the interview and diary portions of the CEX.
23. Non-durables consumption is a common metric in consumption studies, because it avoids the problems associated with treating durable goods as capital.
A basic issue about consumption inequality goes back to the question as to whether consumption or income best represents welfare. While most economic models make utility depend on market consumption and leisure time, it could be argued that at least some component of the portion of disposable income that is not consumed but rather is saved contributes directly to welfare. This point is most relevant to the top percentile where saving can be used in utility-enhancing activities, such as charitable and political contributions. Political lobbying persuades national or local politicians, scholarships and professorships are named, wings of museums are endowed, cocktail receptions are attended with equally rich contributors, and at the stratosphere of income, the Gates Foundation creates utility both for the Gates family and their many beneficiaries as they promote research and programs to combat African tropical diseases. None of these sources of utility for the very rich are counted as consumption in the national accounts, and indeed they are not counted as consumption at all due to top-coding in the CPS and CEX.

7.2 Do Price Deflators Rise Unequally?

Students of the rise of American inequality may be startled by the simple fact that the bottom half of the income distribution, which in the aggregate data has had an increase in real median household income since 1973 far below the rate of productivity growth, still has substantially increased its standard of living as measured by quantitative counts of possessions and housing.

Taking available data for 2001, the facts are astonishing for anyone convinced that the median or even 20th percentile American household is languishing in poverty. Fully one-third of all American households own or lease two or more cars. Two-thirds of households have a garage or carport attached to their home. Ninety-nine percent own a color TV, and 70 percent of American households own two or more color TVs. More than 90 percent own a VCR and/or DVD player. More than half in 2001 owned personal computers, most of which had access to the internet. Half owned a self-cleaning oven, 95 percent owned a frost-free refrigerator, 85 percent owned a microwave oven, one-third owned a detached freezer, more than half owned a dishwasher, and most remarkably, fully three-quarters had air conditioning, roughly 50 percent with central air conditioning and 25 percent with room individual-unit air conditioning. 24

While the official statistical reports state that the median real wage stagnated and median household income barely increased, American households in their millions increased their standard of living, both as measured by the quality of housing as reviewed in the previous section and as measured by the quantity (not to mention

the quality) of the possessions with which they stocked these housing units. How could they afford to do so? The standard consumption inequality literature reviewed above looks only at growth in nominal income corrected by inflation by a single price index, the CPI or the price index for personal consumption expenditures (PCE).

This literature has not yet absorbed the implications of research on historic price index bias. While there is as yet no systematic study of price index bias by income class, there is a prima facie case to be made that relative prices have been declining for the lower half of the income distribution and have been rising for the upper quintile and above.

The 1995-96 Boskin Commission reviewed the sources of price index bias and developed quantitative evidence on the extent of the bias for each of its components. The report is summarized and more recent developments in the price index bias literature are assessed in Gordon (2006). The most important new development is the work of Hausman and Leibtag (2003) which shows that the arrival of a Wal-Mart store in a community reduces consumer prices for food by fully 25 percent, of which 20 percent is the direct Wal-Mart impact and the other 5 percent represents the response of local stores to the Wal-Mart arrival.

Both because low-income households shop at Wal-Mart, and because they spend a larger proportion of their household income on food than high-income households, there is a prima-facie case that the retail channel involving Wal-Mart and in turn its reliance on Chinese imports for non-food products has significantly reduced the cost of living for lower-income households. As pointed out by the Boskin Commission, the CPI as official policy ignores any drop in prices when Wal-Mart arrives and a ma-pa grocery store goes out of business. The entire price reduction is treated implicitly as a “decline in service.”

At the top of the income distribution, consumption expenditures have a very different composition. Many high income households have never visited a Wal-Mart. Their expenditures exhibit a higher share of services, particularly high-end services like private secondary schools, college tuition, high-end spas, massage therapists, landscape gardeners, and other service providers whose relative prices rise steadily relative to the overall consumer price level. Because wages rise but productivity in these services typically do not increase, the rich face rising relative prices while the poor enjoy declining relative prices. So far there is no quantitative evidence on the magnitude of this effect, but it is surely important and may help to explain some of the puzzles in the contrast between the abysmal increase in median wages and household real income, in contrast to the major observed increase in the standard of living of the bottom half of the income distribution in the past half-century.
What is the evidence on a possible price index bias that might overstate the inflation rate of the poor and understate the inflation rate of the rich? The existing literature is not helpful. Hobijn and Lagakos (2005) study differences across American households to see if there are systematic relationships between inflation rates and household characteristics. They find none. However, their paper is flawed by its dependence on official CPI data, which are subject to the outlet substitution bias that does not treat the lower Wal-Mart prices as a measured decline in the price level. Their research also focuses on deviations in household inflation rates over three-month intervals rather than the long-run of several decades that matters for a consideration of consumption inequality.

Gamer, Johnson, and Kokoski (1996) address the question of inequality of inflation rates by asking whether inflation rates differ for the “poor”. As with the Hobijn and Lagakos paper, their work is handicapped by relying on official CPI data that ignores outlet substitution bias. Their study covers only 1984-94 and finds little difference for the CPI inflation rate of the aggregate population and of the “poor”, with expenditure weights based on consumer expenditure surveys of different income groups.

However, we doubt that these earlier results would hold up if extended over a longer period of time. The bottom half of the income distribution has a larger budget share for food, clothing, shelter, and a lower share for services (Garner et al. 1996). According to the deflator for personal consumption expenditures, which rose at an average annual (log) growth rate from 1979 to 2007 of 3.27 percent, several categories disproportionately consumed by the bottom half rose at growth rates less than average, including food (3.01), and clothing and shoes (0.12). The prices of the category “furniture and household equipment” fell at an annual rate 2.07 percent per year, meaning that the average good in this category in 2007 cost only 56 percent as much as in 1979 (and this does not take into account many dimensions of improved quality in these goods). The growth rate of prices of services was 4.21 percent, cumulatively an increase in the relative price level of services of 30 percent, and these are disproportionately consumed by the rich. None of these comparisons allow for outlet substitution bias, i.e., the Wal-Mart effect, not to mention the availability of less expensive pet supplies at Pet Smart, and less expensive home improvement supplies at Home Depot. Overall, we think there is a plausible case to be made that the true price paid by the bottom 50 percent of the income distribution may be rising at 0.2 to 0.3 percent slower than that paid by the top 10 percent, offsetting some of the differential growth rate of median vs. mean labor income.

7.3 Inequality in the Growth of Life Expectancy

The previous section suggested that lower inflation for the poor than for the rich may in part explain the accumulation of possessions and housing space by most
American households, including those below the median. Operating in the opposite direction, however, are recent findings that the poor are plagued by health problems and a lack of access to medical care that may account for a startling divergence in the rate of improvement of life expectancy. As summarized by Pear (2008), a recent study concludes that “growing inequalities in life expectancy mirrored trends in infant mortality and in death from heart disease and certain cancers.” This study (Singh and Siahpush, 2006) is complemented by another recent sets of results that reports that virtually all of the increases in life expectancy reported for the US population can be attributed to the highly educated (Meara, Richards, and Cutler, 2008).

The extent of the differences are startling. Dividing the US population into deciles, Singh and Siahpush conclude that the already large gap in life expectancy in 1980-82 of 2.8 years between the top and bottom deciles (75.8 vs. 73 years) had widened to 4.5 years (79.2 vs. 74.7). They comment that “life expectancy for the rich group in 1980 was higher than that for the poor group in 2000.”

Research on inequality of life expectancy has identified several sources of growing inequality which combine social environment, life-style differences, and access to medical care. First and most obvious, the poor are less likely to have health insurance and accordingly are less likely to receive preventive medicine, including checkups, screenings, diagnostic tests, and prescription drugs. Rich people have better access to high-quality medical care and better information on the latest treatments which they demand from their medical providers. But life expectancy depends not just on access to medical care but also on personal habits and the environment. The prevalence of smoking has declined faster among higher income and education groups. Lower-income people are more likely to live in unsafe neighborhoods and to engage in risky behavior and/or to eat unhealthy foods.

To convert the results of the Singh study into the metric of consumption growth, we can use the bold method developed by Nordhaus (2003). The Singh study says that the rich gained 3.4 years of life expectancy over 1980-2000, while the poor gained 1.7 years, only half as much. This implies that the average 1980-2000 gain in average life expectancy is 2.5 years, not far from the Nordhaus 1975-95 number of 2.1 years.

Nordhaus translates 2.1 years of added life expectancy over 20 years into an annual gain of 1.8 percent relative to the initial year (1975) level of consumption. By analogy we can translate gains for the rich into a 3 percent annual gain in health relative to 1980 market consumption and for the poor into a 1.5 percent gain. How large is this gap of 1.5 percent per year compared to other measures of the growth of income at the top compared to the bottom. Dew-Becker and Gordon (2005, Table 8) calculate from IRS data a real income gain of 1.30 percent for the 90th percentile as compared to an annual increase of 0.48 percent for the 20th percentile, a gap of 0.82
percent. Thus the gap in the growth of life expectancy suggests a greater gap between rich and poor than measured income itself.

We need to be cautious about this conclusion. The Nordhaus method implicitly values life at roughly three times the value of lifetime market consumption, the remainder presumably consisting of the value of leisure and of consumer surplus. But even if we cut the Nordhaus-based difference in consumption growth between rich and poor by half, there still is a major additional aspect to growing American consumption inequality of roughly the same order of magnitude as growing income inequality and that has been largely ignored thus far in the literature on inequality.

7.4 Geographical Inequality

Thus far we have interpreted the growth of inequality as applying to workers, households, or tax-payer units. Whether average per-capita incomes are becoming more or less equal across metropolitan areas is not an independent source of rising inequality. Indeed, any change in per capita incomes across metro areas reflects choices by high and low income individuals about where to work and to live.

Edward Glaeser (in Cutler, 2007, p. 7) provides a provocative description of America becoming more unequal in a geographical dimension:

“There’s a world of prosperous places, mostly on the coasts, that are driven primarily by production of new ideas. . . . Then there are vast areas of America where people are lower-middle income. They’re driving everywhere; they’re buying cheap houses and things at Wal-Mart; they’re living a relatively decent life. . . . And then there’s the third group – the truly disadvantaged in America’s inner cities, particularly those inner cities hit by the decline in transportation costs and the end of manufacturing that weren’t able to reinvent themselves around new ideas” (Cutler, 2007, p. 7)

While Glaeser’s description does not shed new light on the rise in American inequality, it nevertheless suggests an interesting issue to address as we attempt to sort out the causes of rising American inequality. The Bureau of Economic Analysis provides a long time series of per-capita income across metropolitan areas since 1969, broadly and narrowly defined. We have constructed a dataset on income per capita consisting of the broadest BEA definition of metro areas, with 23 million people located in the New York Area, 19 million in the Los Angeles area, and 10 million people located in the Chicago area. There are 37 such broadly defined “BEA economic areas” with populations of over 2 million people in 2005. The ratios of per-capita income to the US average in 1969 range from 75 per cent in Raleigh-Durham to 127 percent in New York City. The ratios in 2005 range from 80 percent in Salt Lake City to 130 percent in New York City.
We find a dominant pattern of regression to the mean in Southern cities that caught up in their ratios to the US average from 1969 to 2005 (Charlotte from 84 to 94, Nashville from 77 to 92, and Raleigh-Durham from 75 to 90). We also find examples of traditional Midwest industrial cities regressing toward the mean, notably Detroit going from 114 to 100. And defying convergence are the bicoastal cities, starting rich and becoming richer, as in Boston from 108 to 122 and Washington-Baltimore from 112 to 129.

We develop a regression equation explaining the 1969-2005 growth rate of the ratio of income per capita in a particular BEA economic area relative to the US average. The explanatory variables are a “convergence variable”, that is the level of the ratio in 1969 (we would expect this to have a negative coefficient) and a set of four regional dummy variables which divide up the 37 metro areas into four groups, “Bicoastal,” “South,” “Immigration” (the states that border Mexico including southern California), and “Central” (everything else). The regressions results yield a highly significant -1.07 coefficient on initial per-capita income, implying a strong convergence effect, but also a strongly positive coefficient on the “bicoastal” cities. Just as is implied by Glaeser’s informal characterization, the bicoastal cities have moved ahead of the rest and have defied convergence, starting with above-average ratios and in most cases increasing the ratios from 1969 to 2005. A notable exception is Los Angeles, for which the ratio fell from 119 to 99, a clear example of the negative impact of immigration on per-capita incomes.

What are the mechanisms by which some metro areas become richer? One mechanism is explored by Zucker and Darby (2008), who show that super-star scientists choose to work in proximity of each other. Metro areas that have a head start, often because of association with a premier university (Harvard or MIT in Boston, Berkeley or Stanford in the SF Bay Area), tend to attract super-star scientists through an agglomeration effect. The Zucker-Darby charts illustrate a striking preference among these scientists to live in only a few bicoastal metropolitan areas. We suggest that the superstar scientists studied by Zucker and Darby are joined by superstar lawyers and investment bankers in moving to the coasts, although in the legal profession this is movement is impeded by a strong infrastructure of law firms in Chicago.

Gyourko et al. (2006) flesh out a richer model of bicoastal separatism. They mesh the demand for housing by the upper income groups with supply constraints in bicoastal (but not Midwest) metro areas that result in high housing prices that only the most educated and richest can afford. Thus the most educated and talented wind up living in Boston, New York, and San Francisco, while the least educated are driven by high housing prices to move from the two coasts to inland areas like Idaho, Utah, and Nevada. The Gyourko paper develops the intriguing and suggestive idea of
“superstar land rents” resulting from constraints on the supply of land for residential building.

This set of papers suggests that geographical separatism is not a cause of rising American inequality but rather a result. There is no policy response that can combat the complex set of causes of increasing geographical inequality. Land use regulations are inherently local. We note in passing that land-use regulations create extreme disparities in urban house prices not just in the United States, but between capital cities and rural regions in Europe, especially between London and its suburbs and the rest of England, and similarly for Paris and the rest of France. This theme leads us now to turn to international comparisons of the increase of inequality.

8. International Comparisons

8.1 Facts on International Differences in Inequality

The stark facts about differences across countries originate in the raw data on inequality at the top as summarized by Piketty-Saez (2006) and analyzed by Atkinson (2007). Figure 6 summarizes the remarkable compilation of data by Piketty and Saez (2006) on the income shares of the top 0.1 percent over the period since 1920 in Canada, France, Japan, the UK, and the US. Here income includes labor, business, and capital income but excludes capital gains. The contrast between the US and the other countries could not be more dramatic.

The time series for the income share of the top one percent in the US traces out a distinct U-shaped pattern, with a high point of 8.2 percent reached in 1928, then a slide to a minimum of 1.9 percent reached in 1973, followed by a steady increase to 7.3 percent in the year 2000.

At the other extreme from the US is Japan, which exhibits a peak of 9.2 percent in 1938 followed by a precipitous collapse during World War II, followed by stability after 1947 at a rate close to 2.0 percent or slightly below, ending up at 1.7 percent in 1998. The other countries had widely ranging shares, but all above 5 percent, between 1920 and 1938, followed by a sharp decline to an amazing convergence in which between 1960 and 1980 all countries had a share of the top 0.1 percent close to 2 percent. During the late 1970s the share for the UK sagged to close to 1 percent and for the other countries between 1960 and 1980 was between 1.5 and 2.3 percent. The upsurge in the top income share after 1980 was led by the remarkable increase in the

25. A variety of papers containing cross-country comparisons of top incomes is in the volume edited by Atkinson and Piketty (2007). A list of citations to ten different countries is contained in Roine et al. (2007, footnote 11).
Figure 6. Share of top 1 percent in Total Income (Labor, Business, and Capital Income, excluding Capital Gains), for U. S., U. K., Canada, France, and Japan, 1920-2000

8.2 Explanations of the Facts

What hypotheses have been proposed to explain the high level and growth of inequality in the US as compared with other OECD countries? One approach taken by Mishel et al. (2007, p. 357) and others, is to cite a difference in the socio-political-economic “system” that differentiates the US from other developed countries, so-called “American exceptionalism” that dates back to the nineteenth century. In the

ratio for the US, followed by Canada and the UK which appeared to mimic the US performance with an elasticity of between 0.3 and 0.5. But the top share in France and Japan has remained amazingly stable, a conundrum of great importance in assessing the behavior of inequality on both sides of the Atlantic.

Other data applying to the entire distribution, not just the top one percent, support the view that inequality in the US is high and growing. Smeeding (2006) reports that the US Gini coefficient in 2000 was 0.368 compared to an average of 0.281 in 17 OECD nations excluding the US. The increase in the Gini coefficient between 1989 and 2000 was 0.030 points for the US and 0.017 points for the other countries. Harjes (2007) notes, however, that by his measure the Gini coefficient rose more in the UK than in the US between the late 1970s and the late 1990s, and that the increase was substantial in such large European countries as Germany and Spain, while being zero in Italy and slightly negative in France. He thus paints a picture not of a sharp contrast between the US and everyone else, but of a high degree of heterogeneity. He also points out that the size of the pie has been shrinking in Europe, with labor’s income share shrinking by 6 to 10 percent in most European countries in contrast to the stability in the US of labor’s share as exhibited in Figure 1 above.
view of Mishel et al. (2007), the market-driven “US Model” leads to more inequality, higher poverty rates, an “expensive-yet-underperforming” health care system, and jobs that require more work hours per year and fewer paid days off. This view is consistent with a view that culture and social norms matter in explaining numerous dimensions of American exceptionalism, of which income inequality is only one.26

Yet for Harjes (2007), the American exceptionalism approach misses the heterogeneity in the level and growth of inequality outside of the US. This points to the possibility of numerous models utilizing different combinations of policies and institutions. For instance, the UK pursued de-unionization in the 1980s as government policy along with privatization, helping to explain the relatively high UK ranking in measures of inequality and its change. In contrast the “consensus model” adopted in the Netherlands and to a lesser extent in Sweden, Ireland, and Germany obtained moderation in wage demands by labor in return in some cases for reduced income taxes and in other cases with the expectations that managers would avoid excess compensation increases for themselves. In Germany excessive executive compensation is mitigated by such institutional features as the two-tier company board with strong labor representation, “legal co-determination rights,” and a high tax rate on capital gains from stock options” (Ponssard, 2001).

There is a large body of research on the effects of specific institutions on inequality. One of the most interesting and influential papers in this literature is Alesina and Angeletos (2005), which argues that there may be feedback between current redistribution and preferences for future redistributive policy. When people live in a country with high levels of redistribution, they may believe that those who are rich have only become wealthy through unfair means. This reinforces the preference for redistribution. The reverse may work in situations with low redistribution. They then argue that the multiple equilibria generated by this model may explain the difference between institutions in the US and Europe. Alesina and La Ferrara (2005) provide evidence supporting the Alesina-Angeletos model.

In recent empirical work Chong and Gradstein (2007) find that there is a joint relationship between inequality and general institutional quality. Using panel data on a large set of countries over 20 years, they find that inequality drives future institutional quality and that institutions drive future inequality. This relationship holds for a variety of measures of institutions, including indexes of civil liberties, political rights, government stability, corruption, and rule of law. Chong and Gradstein confirm the result that inequality can affect subsequent institutions,

26. Others include higher US fertility and relatively low US rankings in league tables of life expectancy and math/science test scores.
previously found by Alesina and Angeletos (2005), Hoff and Stiglitz (2004) and Sonin (2003), and extend it to show that the causation also runs in the reverse direction.27

Another recent study of top incomes by Roine et al. (2007) is based on data for 16 countries extending over most of the twentieth century. Hypotheses are tested as explanations of the level and change in the income share of the top one percent and the ratio of the share of the top one percent to the share of the 90-99 percentile. Two findings are consistent and robust across specifications. The first is that “growth is pro-rich,” i.e., when a country has grown faster than average, top income earners benefit more than proportionately. Second, the level of financial development (measured by the ratios of stock market capitalization and bank deposits to income) has been pro-rich. Neither of these findings differs between the Anglo-Saxon countries and the others, and the financial development effect seems to be strongest at low levels of economic development. They find one difference in the Anglo-Saxon countries, that they have a substantially higher slope coefficient on openness to trade (the ratio of exports plus imports to GDP), and that this is sufficient to explain half of the differential rise in top-share inequality in the Anglo-Saxon countries compared to the others. They do not offer any explanation of this finding, although it is consistent with the discussion in section 4.3 above in which several authors attribute a substantial effect of increased trade in reducing US wages at the bottom relative to the top.

Koeniger and Leonardi (2007) draw an institutionally-oriented contrast between Germany and the US based on the existence of wage floors for unskilled workers in Germany but not in the US. They argue that institutions like the minimum wage that compress wages result in less complementarity between capital investment and skill, so that capital investment decreases rather than increases the skill premium. Their results are not convincing as they stand, because they fail to determine whether minimum wages simply exclude workers from employment or whether they encourage firms to invest more in raising the training of low-skilled workers.

At the very top level of incomes plotted in Figure 6 for the top one percent, Piketty and Saez (2006) point out that the divergence between the English-speaking countries and the others occurs in labor income, not capital income, as the “working rich” have replaced the “rentiers.” They propose three broad classes of explanations. First is SBTC favoring people at the top, but they object that technological changes have been similar everywhere while changes in top income shares have not. On the other hand, the relative supply of different skill classes has not been identical across countries, so SBTC need not be dismissed out of hand.

27. Chong and Gradstein provide a concise review of the somewhat small literature on institutions and inequality.
A second set of explanations class includes regulations, unions, and social norms, a view that they claim implies that “the surge in executive compensation actually represents valuable efficiency gains.” The third class is that the U. S. managerial power explanation that we have associated above with Bebchuk and co-authors, “the increased ability of executives to set their own pay and extract rents at the expense of shareholders.”

We favor a blend of all three explanations. The market is at work in the increase of market capitalization in the U. S. that spilled over into executive compensation through the greater use of stock options than in other countries. We have supported the managerial power view in our summary of Bebchuk’s work in part 6.4 above. And we have summarized several institutional differences previously in this section.

The greater use of stock options to reward executives in the US than in other countries itself reflects institutional differences. Pfanner (2003) reports that in Germany, only half of the companies in the DAX stock market index have any stock option program at all. He quotes a European compensation expert as saying that “There’s obviously a cultural difference” between the US and Europe regarding stock options. And more than cultural differences are at work. According to Buerkle (2000), Belgium taxes stock options when they are granted, while France and the UK impose the tax when the stock is sold. Rules vary so much across European countries that a given gain “could be taxed two or three times on the same option” if a worker moves across national boundaries. As noted above, even if stock options are awarded efficiently, their use will raise average salaries since the value to a CEO of a risky salary is lower than a riskless one. This could explain some of the difference in average pay that is correlated with stock option use.

There is an easy explanation of greater equality at the top in Japan—until 1997 stock options were illegal, except at small start-up companies. Japan loosened its restrictions and also introduced defined-contribution pension plans in 2001. But there is a long lag in the adoption of stock options by major companies after decades of tradition in which executive pay is many multiples less relative to average worker pay than in the US (Bremner, 1999).

Overall we see no point in trying to find a monocausal explanation of the increase in CEO pay in the US relative to other developed countries. Price-earnings ratios increased more in the US than elsewhere, at least through 2000, causing stock market gains to spill over into CEO pay due to the widespread and growing use of stock options. To some extent the lesser use of stock options represents a catch-up phenomenon, with European companies adopting US practices after a lag of one or two decades.
But there is still room for a complementary institutional explanation based on different policies and regulations. Different laws (e.g., the illegality of stock options in Japan before 1997), different customs (the role of labor on corporate boards in Germany), and different institutions (consensual bargaining in the Netherlands and other countries) all play a role in explaining why corporations outside the US have been constrained from offering to their top officers the types of pay packages typical in the US.

9. Conclusion

This paper has provided a comprehensive survey of the increase in American inequality since 1970. Our discussion treats the evolution of labor’s share, the change in 90-50-10 ratios of incomes “at the bottom,” hypotheses about the evolution of the 90-50-10 ratios, nuances in the hypothesis of skill-biased technical change (SBTC), the causes of increased inequality within the top 10, 1, 0.1, and 0.01 percent, the distinction between consumption and income inequality, questions about inequality in the evolution of price deflators, life expectancy, and geographical location decisions, and finally international differences in the evolution of inequality, especially at the top.

We argued in section 2 that there have been no interesting changes in labor’s share of national income over the last two decades, once a consistent cyclical chronology is applied. Over the full period 1950–2006 labor’s share has risen, not fallen, but once the labor portion of proprietor’s income is added in, labor’s share has been almost exactly flat for more than 50 years. Further, we point out that labor’s share in national income is not related to the current debate about increased inequality. If the labor income of the highest-paid workers increased enough, we could observe simultaneously an increase in labor’s share and a decline in the real income of the median worker.

Section 3 documents the evolution since the late 1970s of the 90-50-10 ratios from CPS data for men, for women, and for both together. Our most important finding is that all discussions of income by percentile below the 90th must distinguish carefully between men and women. We were surprised to learn that the 90-10 income ratio for women has increased by fully double the increase for men. While the 90-50 ratio for both men and women increased slowly and steadily from 1979 to 2005, the 50-10 ratio showed a sharp jump in 1979–86 that was twice as large for women as for men. Then the 50-10 ratio remained on a high plateau for women about 20 percent above its 1979 value, while for men the 50-10 ratio gradually slipped back to its 1979 value.

In examining causes for these changes, we focus in section 4 on five elements, the decline of unionization, the increase of trade, the increase of immigration, the decline in the real minimum wage, and the drop in top-bracket income tax rates. The
sharp concentration of the increase in the 50-10 ratio for both men and women on the
1979–86 interval provides strong circumstantial evidence for declining unionization as
a cause for men and the declining real minimum wage as a cause for women. The
timing of the subsequent post-1986 evolution of the real minimum wage is also
consistent with the stable 50-10 ratio for women. Our examination of quantitative
evidence in the academic literature finds a small role for the decline in unionization,
but only for men. The evidence on trade suggests that low-paid foreign workers do
compete with domestic workers, and that the increased penetration of imports pushed
wages downward in the middle by an amount that is difficult to quantify.

The immigration literature is contentious, but we were convinced by a recent
paper showing negligible impact of increased immigration on domestic workers but
rather a big downward impact on foreign-born workers who specialize in particular
occupations traditionally dominated by immigrants.

Section 5 reviews the SBTC hypothesis and potential objections to it,
particularly the slow wage increases of apparently skilled occupations like engineers
and computer programmers, compared to the rapid income gains of managers. We
endorse the effort by Autor and co-authors to broaden the skill distinction to three or
more categories; their polarization hypothesis makes a lot of sense in explaining the
facts about rising inequality and also the occupations most prone to outsourcing. The
key distinction is between interactive work at the top, whether lawyers in courtrooms
or investment bankers making deals in person, and interactive work at the bottom,
whether attendants in nursing homes or immigrant workers mowing the lawns of
well-off people. These jobs at the top and bottom cannot be outsourced. But jobs in
the middle can be outsourced in the broad middle where people do routine, easily
duplicated jobs such as airline reservations agents or workers at technical call centers.

Section 6 finds ample evidence that SBTC is a major explanation of increased
skewness of labor incomes at the top. We distinguish three different types of top
incomes. Superstars include the top members of any occupation that provides
disproportionate rewards to the first-best as contrasted with the second-best. The pure
superstar phenomenon has at its core the magnification of audiences, the fact that a
single performance can be witnessed by an audience of one person or ten million
people, depending on the perceived attraction and talent. A second category of top
incomes is market-driven and includes law partnerships, investment bankers, and
hedge fund managers, where there is no obvious analogy to audience magnification.

The most contentious question regards the third category, that is, the sources of
enormous increases in the ratio of top executive compensation to that of average
workers, both over time and between the United States and other developed nations.
The core distinction is that superstars and other market-driven occupations have their
incomes chosen by the market, whereas CEO compensation is chosen by their peers in
a system that gives CEOs and their hand-picked boards of directors, rather than the market, control over top incomes. This idea that managers have power over stockholders is nothing new; the idea that managers control stockholders rather than vice versa goes back to Berle and Means (1932) and R. A. Gordon (1945). This idea that the principal-agent control of stockholders should be reversed has been applied fruitfully by such authors as Bebchuk and Fried (2004). We endorse their idea that managerial power lies behind some of the outsized gains in CEO pay, while also recognizing that stock options created an automatic spillover from the stock market gains of the 1990s directly into executive pay.

Has consumption inequality also risen as much as income inequality? If increased cross-sectional income inequality is simply the result of larger transitory shocks to income, and if financial markets are sufficiently well developed (assuming, against substantial evidence to the contrary, that liquidity constraints are not a major impediment), then consumption and welfare inequality could have stayed constant. In reviewing the evidence, it is clear that consumption data in the US does not measure exactly what we might hope for. While authors have found parts of the Consumer Expenditure Survey (CEX) revealing little increase in consumption inequality, other more believable parts of the CEX show consumption inequality to rise at roughly the same rate as income inequality. This evidence is consistent with that of Kopczuk, Saez, and Song (2007) who find that there has been no increase in income mobility associated with the rise in income inequality.

Our survey introduces two new issues into the discussion of consumption inequality. First, we connect with the literature on price index bias and argue that the price deflator for goods consumed by the poor has increased less rapidly than for goods consumed by the rich. This difference reflects the slower inflation rates of food, clothing, and electronics as well as the “Wal-Mart effect” that is not captured by official price data such as the CPI. Part is a compositional effect, that the relative prices of services increase over time due to relatively slow productivity increases, and the rich consume relatively more of the high-inflation services.

While the poor may do better when price indexes are corrected, they do much worse when their health outcomes are considered. Recent evidence suggests that between 1980 and 2000 the life expectancy of the bottom 10 percent increased at only half the rate of the top 10 percent. This translates into an increase in health welfare (using a method developed by Nordhaus) that is roughly 1.5 percent per year faster for the rich than the poor, expressed as a ratio to initial market consumption. This may be the most important single source of the increase in inequality in the United States, and it combines not only unequal access to medical care services and insurance, but also to differences in personal habits and environment related to education and income.
An aspect of inequality rarely discussed in the literature is the divergent evolution of relative per-capita incomes (compared to the US average) across US major metropolitan areas. We present two new results for the 37 largest broadly-defined US metropolitan areas. The first is that most of the areas have converged toward the national mean as the southern states have risen from sub-par toward the average, and the old Midwest industrial cities have regressed from above-average toward the mean. The striking exception is the set of bicoastal metropolitan areas that have been called “superstar cities,” with both levels and rates of change of per-capita income that significantly exceed the US average. We link this rise in geographical inequality both to constraints in the supply of land for housing and to the desire by superstar scientists and other top-income professionals to live near each other in these bicoastal superstar cities.

Some of the most interesting remaining issues in the area of increased inequality involve cross-country differences. The post-1970 upsurge in US inequality, particularly the relative rise of top incomes, is much greater than in continental Europe or Japan, with the UK and Canada somewhere in between. We propose a mix of institutional and market-driven explanations. Institutional differences between the US and Europe include the earlier and more pervasive introduction of stock options in the US, the tradition of corporatism and cooperative bargaining in Europe that creates constraints on management compensation excess, and the larger role of unions and a higher real minimum wage in some European countries. But the market matters also; gains in profits and price-earnings ratios in the US stock market in the 1990s spilled over to executive compensation, interacting with the large increase in the share of executive compensation taking the form of stock options.

The study of income inequality is of fundamental importance to economics. The most obvious reason is that if economics is at all concerned with understanding the development of the economy over time, we must understand not only changes in means, but also changes in distributions. Second, changes in inequality can be indicative of changes in the structure of the economy that may favor one group or another, e.g. skill-biased technical change. Third, variation in inequality can tell us how well our theories about risk sharing and consumption smoothing actually fit with peoples’ experiences. Fourth, we can learn about the effects of various institutions on inequality by studying the experiences of different countries. This allows informed policy choices to be made in the future. What these policy choices should be, if any, are beyond the reach of this paper. We have attempted to link facts and hypotheses, and some of these links are clearly robust. These facts should be taken into account in policy discussions, and some, simply by being aired, may improve outcomes in the economy.
REFERENCES


