Technology and the Labor Market:
What We Know and How We Can Know More

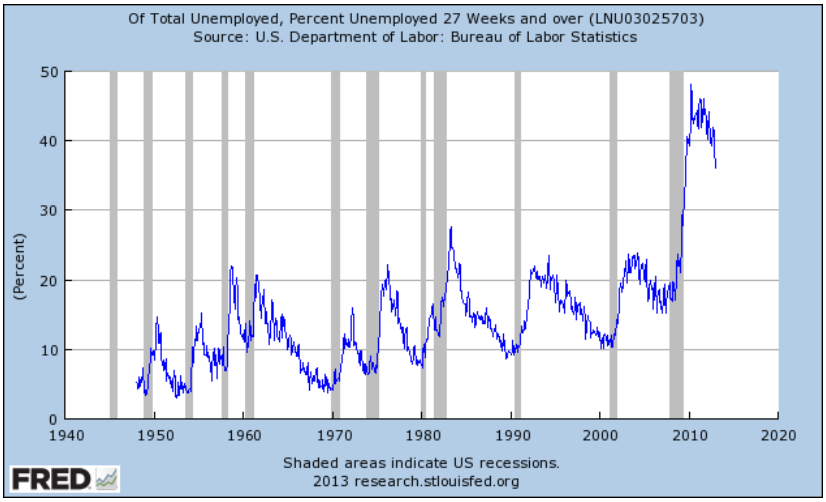
John M. Abowd, Michael R. Strain, & Lars Vilhuber

There is compelling evidence to support the theory that technology is changing the types of jobs in the U.S. labor market.

To understand how, think of three kinds of jobs: low-skill (a custodian, say), middle-skill (a bookkeeper), and high-skill (a technical engineer). A computer is not going to replace a custodian — computers cannot yet make the rounds in an office after hours, emptying trash cans, vacuuming carpets, and cleaning restrooms. And a computer is not going to replace a technical engineer. The engineer engages in creativity and problem-solving which computers are not capable of — computers are good at taking orders, but not at giving them. But a computer could replace a bookkeeper. In fact, accounting software for personal and small-business finances has replaced many bookkeepers. Similarly, ATMs have replaced bank tellers.

Bookkeepers and bank tellers were never the lowest paid workers. Their jobs required trustworthiness, professionalism, and accuracy — the bank teller has to add up the deposit correctly each time; the bookkeeper has to balance the accounts each month. There is more at stake if a bookkeeper or bank teller makes an error than if a custodian does. But precisely because their jobs required repetition and rule following, they were able to be replaced by a (much) cheaper computer.[1]

Does the effect of computers on employment have anything to do with our current labor market troubles? Is current unemployment being driven by structural, technology-related changes in the labor market? The labor market is in abysmal shape. Perhaps most troubling are the approximately five million workers who have been unemployed for 27 weeks or longer — the long-term unemployed. Prior to the Great Recession, the previous post-war record was a little under three million long-term unemployed. Today, over one-third of unemployed workers are long-term unemployed — a staggeringly high share, again a post-war record outside the current downturn.



One of the most urgent policy questions facing the United States today concerns the fate of the long-term unemployed. What will become of them? Will they be able to find good jobs and resume productive employment? To answer that question, it would be helpful to know the cause of the changing face of unemployment: Why are so many workers experiencing long-term unemployment? Why is it so hard for them to find employers who want to hire them?

Of course, we won't provide a definitive answer to that question here. But to begin thinking about that question, take a step back and ask whether our current unemployment is structural or cyclical.

While there are a number of ways to define the difference between these two types of unemployment[2], cyclical unemployment can be thought of as being a consequence of temporary responses to a lull in demand. A worker's unemployment is cyclical, for example, if she is on a temporary layoff and will return to her firm when the economy picks back up.

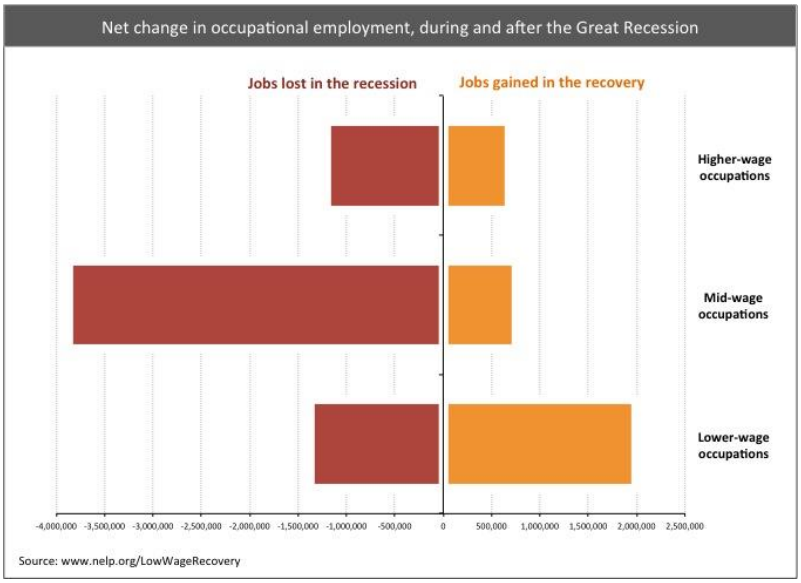
Structural unemployment is a different story altogether. A worker's unemployment is structural if, for example, he no longer has the skills or demographic characteristics the labor market demands, or if he lives in a geographic location where there are no jobs.

The weight of the evidence from the economics literature seems to be that current unemployment is mostly cyclical — it is largely a response to the significant decline in aggregate demand associated with the Great Recession — not structural.[3] But both types of unemployment are surely present in today's labor market — it is likely that many of the workers who have been unemployed for longer than six months don't have the skills (or demographics) required to find and maintain employment in the current labor market.

What would cause this structural unemployment? One answer is that firms may be using the Great Recession as an opportunity to reorganize their production functions. Structural unemployment could result if firms are using the current downturn to move their operations to a better location, to change the types of workers they employ, to reorganize their occupational distribution. And, very importantly, to change the way they use technology in their production of goods and services.

During the current downturn, are firms changing the way they use technology in a way that affects employment?

The National Employment Law Project calculates that from 2008 to 2010 low-skill and high-skill occupations each constituted about one-fifth of Great Recession jobs losses. The remaining 60 percent of job losses were from middle-skill occupations.



In the recovery, nearly 60 percent of job gains have been from low-skill occupations, with roughly one-fifth of gains coming from the remaining two skill groups.[4]

These summary statistics are far from conclusive. But they are consistent with the theory that firms are using the recession to reorganize the composition of their workforce. And that this reorganization is consistent with what's been happening in the United States for several decades — middle-skill jobs are disappearing and are being replaced by technology. This could mean that middle-skill workers will have an increasingly hard time finding a job, being forced to spend months on unemployment.

HOW CAN WE learn more about the way that technology is transforming the labor market? There are many answers to this question, but here's a very important one: Through better data.

Many of the ways in which technology affects workers, jobs, and the labor market are inferred rather than measured. Firms, of course, have a pretty good idea of what their own workers are doing. Community colleges have a pretty good idea of what they are training workers to do. Workers typically know their own skills well.

But putting it all together — even from the point of view of a sophisticated economic analyst trying to make sense of the labor market — is an exercise frustrated by the incompleteness of current data. Much of what each party knows is not reflected in any dataset, and the information that is recorded is hard to put together across different data sources.

Consider the evolution of manufacturing. Back in the old days, many workers, manhandling machines, would produce a widget. Then the machine was replaced with a semi-automated machine, and the operator — now only one — needed to be trained in the particular electronic

interface of the machine. Then, sometime in the hypothetical future, the machine will again be replaced, this time with an industrial 3D printer, requiring computer knowledge to program it, but requiring little training to operate. The worker now operating the machine has a degree in computer science.

This is, of course, a simplification of the dramatic changes in manufacturing, but it illustrates the underlying reality that the types and number of workers being employed by the manufacturing firm changed dramatically over time. The output of the firm, on the other hand, may not have changed at all: the machine still produces widgets. Traditional output-based measures would completely miss the complicated, technology-driven changes to the firm's workforce.

To understand and fully measure these phenomena, we need a way to pry open the black box of production. What is it that workers do at work? What do workers contribute to the production of goods and services? How do workers make these contributions? We need to be able to answer these basic questions through the direct measurement of economic data.

Currently, representative statistics only have a small window on what workers do and how much they earn doing it and cannot reliably be linked to what firms produced. We only have a periodic small glimpse into that world, and it is incomplete.[5] We need better data to get better measurement.

HOW CAN WE get better data? By building on data the government already gathers.

Did you know that every quarter, the U.S. government currently surveys over nine million establishments employing 98 percent of all workers? Every quarter. And every year, over 142 million survey forms are filled out by businesses — more than the number of households in the United States.

We don't normally consider them surveys: the above numbers relate to the Quarterly Census of Employment and Wages (QCEW) and the individual 1040 tax returns reported to IRS.[6]

We don't treat them as surveys because they are administrative data collected for the purposes of computing firm and person tax liabilities. But using them as surveys would represent a major improvement in our attempts to understand how technology affects employment.

How? Consider the possibilities present with income tax records. Many folks use electronic methods to fill out the form. One could use sampling methods to randomize some components — for instance, you could ask some percentage to provide information on their occupation using such survey tools as activity calendars, consistency checks, and default values (sometimes providing last year's value, sometimes not) — without increasing overall respondent burden. Overall, this might lead to more reliable measures of occupation, available for statistical purposes.

Consider also the unemployment insurance wage records that underlie the unemployment insurance administration and that are also the source for the QCEW. Currently, those wage records have only limited information: the worker's earnings, her Social Security number, and a code that identifies her employing firm. Supplementing this with additional information about the worker's job would tell us a lot. Analyzing millions of worker-level records over the course of several years would tell us much of what we want to know about how technology affects employment.

A simple improvement would be to include the occupation of each worker on her wage record. Many European countries already do this. Ideally, in addition to occupation, the form would indicate the types of tasks done by the worker. The form could indicate whether the workers' tasks are repetitive and automatable, or whether they involve creativity and judgment. Another simple improvement would be to list the number of hours the worker worked in, say, the week before the record was completed. Also, including the precise date on which the job ended would be a big help towards understanding the nature of the job.[7]

Data on gross labor market flows — on new hires, quits, layoffs, job creation, job destruction — are currently available in the United States,[8] but they do not contain information on worker occupation or worker tasks. We could significantly increase our understanding of technology's impact on employment if they did.

WILL A ROBOT take your job? With better data, we could fill in the blanks in this answer: Well, the economy destroyed X million jobs over the last three years, and Y percent of those jobs were characterized by routine tasks. The economy created Z million jobs over the same time period, and W percent of those jobs were characterized by tasks which require creativity and judgment. There were Q million "stable jobs" — jobs which last a reasonably long time with steady earnings — in the economy over the last five years. Of those Q million, R percent of the stable jobs that were destroyed this year were characterized by routine tasks.

Today we have crude measures of skills: median wage in an occupation and average years of schooling are two commonly used measures to define an occupation's skill level. But if we knew the actual tasks a worker performs, then we could have a much more precise measure of the skill level of the job and of the worker.

Today our best data on workers is not readily linked to our best data on firms. The Current Population Survey, Decennial Census, and American Community Survey are great data sources from which you can learn about the occupation of American workers, but they are not readily linked to information about the firms in which the workers are employed.[9]

Today we know stock measures: Relative to 1980, low- and high-skill occupations are more prevalent than middle-skill occupations. We know a series of snapshots, but we don't know the underlying dynamics. We can calculate net flows, but not detailed gross flows by occupation, skill, or tasks.

Better data could tell us more about skills than occupation and wages. Better data could let us link workers and firms to see how tasks result in output. Better data could let us measure dynamics — the rate at which middle-skill jobs are being destroyed, the types of low-skill jobs that are being created, and more.

The United States already has the infrastructure necessary to move our understanding of the effect of technology on employment from inference to measurement. The task ahead is building on that infrastructure. Adding more detail to IRS tax records and unemployment insurance wage records would be a great place to start.

John M. Abowd is the Edmund Ezra Day Professor of Economics, Professor of Information Science, and member of the Department of Statistical Science at Cornell University, where he directs the Labor Dynamics Institute. Professor Abowd is also Chair of the Business and Economic Statistics Section and Fellow of the American Statistical Association, serves as a Distinguished Senior Research Fellow at the U.S. Census Bureau, and serves on the National Academies' Committee on National Statistics.

Michael R. Strain is a Research Fellow in Economic Policy Studies at the American Enterprise Institute for Public Policy Research, a Fellow at the Labor Dynamics Institute, and was previously the manager of the New York Census Research Data Center and a member of the research team of the Longitudinal Employer-Household Dynamics program at the U.S. Census Bureau.

Lars Vilhuber is Executive Director of the Labor Dynamics Institute, Senior Research Associate in the School of Industrial and Labor Relations, and affiliated faculty in the Department of Economics, at Cornell University. He is also a Senior Economist at the Center for Economic Studies at the U.S. Census Bureau, and a Visiting Scholar at HEC Montréal.

The views expressed in this paper are those of the authors only and do not necessarily reflect the views of their affiliated and employing institutions.

NOTES

[1] David Autor, Frank Levy, and Richard Murnane attempt to discover how computers change the labor market. The rapid adoption of computing technology changed the labor market in two ways: (1) by replacing some jobs with computing technology, and (2) by making other jobs more productive. The types of jobs that computers are good at performing were replaced: jobs that are characterized by routine tasks — tasks which can be accomplished by following explicit, programmable rules. Computers made workers in high-skill occupations more productive. ("The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, Vol. 118, No. 4, November 2003.) In a separate paper, Autor and David Dorn calculate the change between 1980 and 2005 in the share of employment accounted for by different occupations. When occupations are ranked by skill level, a U-shaped pattern emerges. The share of employment accounted for by low-skill occupations is increased significantly over the time period. The same is true for the employment share of high-skill occupations. Middle-skill occupations account for a significantly smaller share, suggesting that those jobs are being disproportionately replaced by computers. ("The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, forthcoming.)

[2] Other prominent definitions not discussed above are (1) that cyclical unemployment can be affected by fiscal and monetary policy whereas structural cannot; (2) that cyclical shifts are short-lived whereas structural shifts are "permanent"; and (3) that cyclical unemployment is caused by changes in labor demand whereas structural unemployment is caused by changes in labor supply.

[3] A number of economists have studied this question. Atif Mian and Amir Sufi find that between 2007 and 2009 counties that experienced a large negative shock to consumer demand lost relatively more jobs in the non-tradeable goods sector (e.g., restaurant jobs), whereas job losses in the tradeable sector were distributed evenly across counties. They find that around two-thirds of job losses were caused by a drop in demand. ("What Explains High Unemployment? The Aggregate Demand Channel," NBER Working Paper No. 17830, February 2012.) Edward Lazear and James Spletzer finds that losses and gains in the labor market have been symmetric: those industries which initially had the largest increase in unemployment and the most "mismatch" between vacancies and unemployment later had the largest decreases. also show the largest decreases. ("The United States Labor Market: Status Quo or A New Normal?"

NBER Working Paper No. 18386, September 2012.) See also Jessse Rothstein, "The Labor Market Four Years into the Crisis: Assessing Structural Explanations," *Industrial and Labor Relations Review*, Vol. 65, No. 3, June 2012.

[4] National Employment Law Project, "The Low-Wage Recovery and Growing Inequality," Data Brief, August 2012.

[5] For example, the Occupational Employment Statistics (OES) program at the Bureau of Labor Statistics (BLS) produces employment and wage estimates for over 800 occupations, but it only covers 62.2 percent of total national employment, and only surveys a given establishment once every three years. The Confidential Information Protection and Statistical Efficiency Act of 2002 should help with linking data on workers to data on firm output, but our experience is that data sharing between federal agencies could be improved.

[6] The Quarterly Census of Employment and Wages (QCEW) program — also referred to as the ES-202 program, its old name — at the BLS produces county- and MSA-level quarterly counts of employment and wages. The counts are comprehensive for workers covered by state unemployment insurance laws and for federal workers covered by the UCFE program. The 2010 Statistics of Income (SOI) Complete Report estimates 142,892,050 individual returns filed.

[7] Data on U.S. federal workers already has the date on which the job ended (the Enterprise Human Resources Integration-Statistical Data Mart). The Canadian government also collects this information for all workers, not just government workers (called the Record of Employment).

[8] To learn more, see John M. Abowd and Lars Vilhuber, "National Estimates of Gross Employment and Job Flows from the Quarterly Workforce Indicators with Demographic and Industry Detail," *Journal of Econometrics*, Vol. 161, March 2011.

[9] The Longitudinal Employer-Household Dynamics (LEHD) program of the U.S. Census Bureau creates and analyzes confidential, longitudinally linked employer-household microdata.