

RESEARCH STATEMENT

GIOVANNI L. VIOLANTE

In modern economies, income inequality among households is largely determined by the effectiveness of markets in rewarding individual skills, and it is significantly limited by the extent to which policies and institutions redistribute resources across agents. Furthermore, an additional crucial factor shapes the income distribution: risk, or luck. As Lucas (1992) puts it: “*the study of distribution is [...] the study of social mobility, and one cannot discuss social mobility without reference to uninsured individual risk.*”

Most of my research has focused on three questions concerning the role of skills, risk, and policies for the individual distribution of income. First, what are the economic determinants of inequality? Second, what are its implications for macroeconomic efficiency and for households’ welfare, given the presence of uninsurable risk? Third, how is the distribution of resources in the economy affected by government policies?

To address these questions, I have mostly used the set of tools and methodologies known as *quantitative macroeconomics*, i.e., dynamic stochastic equilibrium theory combined with calibration/estimation techniques to achieve an economic model that can be used as a measurement instrument of the phenomenon of interest (Kydland and Prescott, 1996). My research lies at the intersection between macroeconomics and labor economics. Labor economics has long recognized the pervasiveness of individual *heterogeneity* in micro data and, therefore, its importance to interpret key economic outcomes such as earnings and unemployment. In the macroeconomic models I have developed, individual heterogeneity is a chief feature. I have made large use of several different micro-data sets to measure heterogeneity and properly incorporate it into the model.

The end-results of this methodology are frameworks of analysis rich enough to capture salient aspects of the micro data and yet, being structural, they allow to understand the determinants of inequality through counterfactual experiments; they permit to measure its welfare consequences thanks to their sound microfoundation; and they are suitable for large-scale policy evaluation that takes into account general equilibrium effects in the macroeconomy, as advocated for example, by Heckman (2001).

I now move to a detailed analysis of my three main research questions.

UNDERSTANDING THE DETERMINANTS OF LABOR MARKET INEQUALITIES

The labor market is the source of much inequality, since the economic fortunes of individuals are largely shaped by their labor market experience. Data for the U.S. economy reveal that the cross-sectional inequality of labor earnings in the population is, today, at its postwar peak. In particular, in the last three decades the U.S. have experienced a sharp surge in labor income dispersion. In 1970, the ratio between the earnings of a worker at the 90th percentile of the cross-sectional distribution and a worker at the 10th percentile was just over 3. This ratio has

now reached 4.5. The empirical literature has established that about one half of this increase can be attributed to a rise in the market returns to observable skills, especially education and experience, while the other half is “residual”, i.e., due to unobserved attributes of workers belonging to the same educational or demographic group.

In my first line of research—fully articulated within the broader context of the related literature in *The Effects of Technical Change on Labor Market Inequalities (Handbook of Economic Growth, 2005)*—I have argued that the recent wave of technological transformation associated to the diffusion of new information and communication equipment in the economy is, through a variety of channels, an important determinant of the observed changes in the wage structure.

Measurement of technical change A key step in the above argument is the proper measurement of investment-specific productivity improvements, i.e., the speed of technical change. Under general assumptions, one can back out investment-specific technical change from the relative price of capital. However, this approach is complicated by the fact that the quality of certain investment goods (e.g., computers, software) changes very rapidly, so official price indexes are bound to partly miss these changes. Much of my work in this area is summarized in the paper *Investment-Specific Technical Change in the U.S. (1947-2000): Measurement and Macroeconomic Consequences (RED, 2002)*. By extrapolating Gordon’s (1990) measures of the quality-bias in the official NIPA price indexes, we construct quality-adjusted prices for 24 types of equipment and software (E&S), and use them to measure technical change at the aggregate and at the industry level. We find that aggregate productivity growth embodied in E&S averaged 4% annually in the postwar period, and over 6% annually in the 1990s. The acceleration of the 1990s occurred in every industry, consistent with the view that information capital represents a new “general purpose technology”. Our time series are now widely used in macroeconomic research. For example, several papers written on the “Gali-Fisher” controversy about the response of aggregate hours worked to a technological shock use our data.

Education premium The rise in the relative wages of college graduates observed since the mid-1970s has occurred even though, at the same time, the relative supply of college graduates has grown speedily. Simple neoclassical analysis suggests that the economy has witnessed a relative labor demand shift in favor of skilled workers, a concept labelled as skill-biased technical change (SBTC). Albeit popular, a substantial drawback of the SBTC hypothesis is that SBTC has been modelled as an unobserved residual measured from the aggregate production function. So, very much like TFP, it is a “black box”. In *Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis (ECA, 2000)*, we give a deeper economic content to the notion of skill-biased technical change. This paper argues that the substantial drop in the relative price of equipment capital is the force behind the rise in the skill premium. The decline in the price of equipment due to productivity improvements, especially those embodied in the information and communication technologies rapidly adopted in the 1980s and 1990s, led to an increased use of capital in production. Since skilled labor is more complementary with equipment capital than is unskilled labor, the higher capital stock increases skilled wages relatively more than unskilled

wages. Consequently, the skill premium rises. With plausible elasticities of substitution, the model can explain the entire surge in the wage gap between college and high-school graduates in the U.S. economy since the late 1960s.

Residual inequality In an influential paper, Nelson and Phelps (1966) wrote that “*educated people make good innovators, so that education speeds the process of technological diffusion.*” This conjecture can be generalized by arguing that more “able” workers tend to fare better in time of rapid technological change because it is less costly for them to acquire the additional skills needed to use the new technology and adapt to the new production environment.

By postulating this role for “innate ability”, which is inherently unobservable in the data, one can arrive at a coherent theory, based once again on technological change, to rationalize the rise in residual wage inequality. In *General Purpose Technology and Wage Inequality (JEG, 2002)*, we lay out a formal theoretical framework to study, analytically, how the Nelson-Phelps hypothesis can improve our understanding of the recent changes in the wage structure. Among the main results, the model qualitatively explains the increase in the experience premium detected in the U.S. data.

By relying on ex-ante fixed differences in ability across individuals, this explanation has a stark empirical implication: the rise in residual inequality should be extremely persistent because workers would tend to become more stratified in the wage distribution on the basis of their *innate skill* dimension. However, the empirical literature has often challenged this implication: a sizeable part of the rise in residual inequality (at least one third) is attributable to more pronounced individual wage volatility which, in turn, reflects factors that are of a very *transitory* nature.

In *Skill Transferability, Technological Acceleration and the Rise in Residual Wage Dispersion (QJE, 2002)* I offer a quantitative theory of the rise in the transitory component of residual wage inequality. An acceleration in the growth rate of quality-improvement in equipment capital, like the one measured in the data since the 1980s, increases the productivity differentials across successive vintages of machines and, therefore, across jobs. In a frictional labor market, where bargained wages are proportional to the productivity of capital used by the worker, this force translates into higher wage dispersion even among ex-ante equal workers. In presence of some vintage-specificity of human capital, this acceleration reduces workers’ capacity to transfer skills from old to new machines, generating a rise in the cross-sectional variance of skills, and therefore of wages. Through calibration of the model, I show that this mechanism can account for most of the rise in the transitory component of wage dispersion in the U.S. economy. Two key implications of the theory, faster within job wage growth and larger wage losses upon displacement, find empirical support in the Panel Study of Income Dynamics (PSID) data.

Luck vs. Skills One of the most fundamental questions in the study of distribution is: how much inequality among workers is due to different skills (e.g., ability, education, occupation-specific human capital) vs. different luck (e.g., in finding the “right” job). One can call this latter type of wage inequality, inherently associated to informational frictions and luck in the

search and matching process, *frictional wage dispersion*. The canonical search model provides a natural starting point for thinking about this issue. The paper *Frictional Wage Dispersion in Search Models: A Quantitative Assessment* demonstrates that standard search models of equilibrium unemployment, once properly calibrated, can generate only a small amount of wage dispersion. We derive this result for a new measure of wage dispersion—the ratio between the average wage and the lowest wage paid (the “mean-min ratio”). We show that in a large class of search and matching models this statistic arises very naturally from the reservation wage equation, and can be expressed, in closed form, as a function of observables.

Exploring various independent micro data sources (PSID, Census, and Occupational Employment Survey) suggests that, empirically, residual wage dispersion (i.e., inequality among observationally similar workers) exceeds the model’s prediction by a staggering factor of 20. Extending the baseline model to incorporate risk aversion and on-the-job search can improve its performance, but only modestly. We conclude that either frictions account for a small fraction of residual wage dispersion, or the standard model needs to be augmented to confront the data. Since the search models is a workhorse of macroeconomics, we expect that, in the near future, further empirical and theoretical research will be devoted to resolve this puzzle.

MEASURING THE ECONOMIC IMPLICATIONS OF RISING INEQUALITY

My second line of research has analyzed the consequences of the sharp change in the earnings distribution for the survival of institutions, like labor unions, and for the welfare of U.S. households.

Unions Historically, unions have been key players in the determination of wages and of other important labor market outcomes. However, over the past 30 years, the U.S. experienced very rapid deunionization. To date, the chief explanation of this phenomenon is that the legal and political environment supporting union membership deteriorated in the 1970s and 1980s. Moreover, there is a variety of evidence that unions compress the structure of wages, and thus many suspect that their decline may have been a decisive factor in the increase of inequality.

In *Deunionization, Technical Change and Inequality (CRCSP, 2001)* we outline the new hypothesis that deunionization has been caused by the technological developments of the past three decades. The argument rests on the view that unions are coalitions of heterogeneous workers (i.e., some skilled, others unskilled) which extract rents from employers and only exist insofar as both types of members have an incentive to participate in the coalition and continue bargaining in a centralized fashion. Centralized bargaining raises the wage level for all workers, but compresses wage differentials relative to marginal product differentials.

Skill-biased technical change (SBTC) can dramatically alter union participation incentives by increasing the productivity gap between workers. As SBTC progresses, the wage compression imposed by unions becomes unsustainable, and skilled workers break away from unions. Therefore, our new theory discards deunionization as a parallel source of inequality, independent of technical change. Rather, it posits that deunionization *amplifies* the original impact of techno-

logical forces on inequality, by removing the wage compression constraint imposed by collective bargaining.

Households In a series of four papers with J. Heathcote and K. Storesletten, I have applied the dynamic macroeconomic theory of incomplete markets to advancing our understanding of how households can mitigate the impact of labor market risk through a series of “insurance channels”, such as labor supply, within-family risk-sharing, self-insurance via precautionary saving, and the acquisition of education.

It is known that the slope of the empirical life-cycle profile for consumption dispersion contains precious information about the ability of households to insure against labor market risk. In *Two Views of Inequality over the Life-Cycle (JEEA, 2005)*, we argue that the choice of whether to control for cohort effects or for time effects, when plotting such age profile, has a drastic impact on the conclusion. We gather strong evidence, in thirty years of U.S. data, that time effects are required to account for observed trends in inequality, while cohort effects are not. Given this conclusion, we find that the life-cycle profile for consumption dispersion is much less steep than previously assumed in the literature pointing towards the conclusion that labor market risk is *insurable* to a larger extent than previously thought.

In the paper *The Macroeconomic Implications of Rising Wage Inequality in the United States*, we explore the consequences of the recent sharp rise in wage dispersion for the cross-sectional household distributions of hours worked, consumption and earnings. Using panel data from the PSID, we decompose the rise in wage inequality over 1967-1996 into changes in the variances of persistent and transitory shocks. We build an overlapping-generations model with incomplete markets à la Huggett-Aiyagari, and show that given the estimated wage process, the model can account for a large set of facts relating to the evolution of the joint distribution over earnings, hours, consumption and wealth. The richness of our framework allows to quantitatively investigate the question that underlies and motivates much of the research on the recent rise in U.S. wage inequality: how large are the welfare losses of this phenomenon for U.S. households? We find that the answer depends on the demographic group of interest. For high-school graduates entered in the labor force in the early 1980s—the worst-affected cohort—the welfare loss is equivalent to a 15% decline in lifetime income, whereas college graduates of the same cohort experienced, on average, a gain of 12%. We are currently extending the setup to incorporate an education choice, and a model of two-earner families.

The twin papers *Insurance and Opportunities: The Welfare Implications of Rising Wage Dispersion*, and *Consumption and Labor Supply with Partial Insurance: An Analytical Framework* lay out a tractable setting where individuals choose consumption and labor supply while facing a rich stochastic process for wage and preference shocks (incorporating both random walk and transitory components). Agents can perfectly pool certain risks within their “group” (e.g., families, sectors, occupations), but they can only insure against group-level risks by trading non-contingent bonds between groups. The model also allows for the size of both insurable and

uninsurable risk to be correlated with aggregate fluctuations. The key to maintain analytical tractability is a generalization of the Constantinides and Duffie (1996) methodology that allows for labor supply, group-insurance, transitory wage shocks, and preference heterogeneity.

This is, to our knowledge, the first equilibrium framework of imperfect insurance that can be entirely solved and characterized analytically: one can obtain exact closed-form expressions for cross-sectional equilibrium means, variances and covariances of wages, hours, earnings and consumption. Its potential applicability, in my opinion, is extensive.

Our first paper offers an analytical characterization of the welfare effects of changes in labor market risk. Welfare effects are expressed both in terms of changes in the observable joint distribution over individual wages, consumption and hours, and in terms of the underlying parameters defining preferences and wage risk. Our closed-form expressions reveal an important trade-off for welfare calculations. On the one hand, as wage uncertainty rises, so does the cost associated with missing insurance markets. On the other hand, greater wage dispersion can be welfare improving, since it presents opportunities to increase aggregate productivity by concentrating market work among more productive workers. Interestingly, when we calculate the welfare gains from “completing the markets” in the economy, we find that they stem primarily not from reduced consumption dispersion, but precisely from a more efficient allocation of labor effort.

In our second paper, we use the closed-form expressions defining the evolution of inequality through time and across the life-cycle to identify and estimate all the model’s structural parameters. Our panel data comes from the PSID and the Consumer Expenditure Survey (CEX). We implement a hybrid estimation strategy, by Minimum Distance, that exploits information contained both in changes in variances at the aggregate level (the cross-sectional facts on dispersion familiar to macroeconomists) and in the variances of changes at the individual level (the micro facts familiar to labor economists). One important outcome of the estimation is that we compute that almost 2/3 of the recent rise in wage risk was insurable, from the standpoint of U.S. households.

QUANTIFYING THE EFFECTS OF POLICIES ON THE MACROECONOMY

My third line of research studies the impact of numerous labor market and social insurance policies on the household distribution of income and on some important macroeconomic aggregates, such as unemployment, output, and factor prices. I also address the question of how existing policies can be improved upon in order to gain efficiency.

Welfare-to-Work While the literature has studied in detail the efficient design of unemployment insurance within a dynamic moral hazard framework, there has been no attempt to analyze the optimal design of more complex Welfare-to-Work (WTW) programs, where the government can use a wide mix of policy instruments (e.g., unemployment insurance, job search monitoring, social assistance, wage subsidies) targeted to the unemployed. The paper *Optimal Welfare-to-Work Programs (REStud, forthcoming)* lays out a dynamic principal-agent frame-

work (in recursive form) suitable for analyzing the chief features of an optimal WTW program. The economic environment takes into account relevant aspects of the labor market neglected by the previous literature, notably that during an unemployment spell the wage depreciates and the re-employment probability falls.

The optimal WTW program endogenously generates an absorbing policy of last resort, resembling to social assistance policies like Food Stamps, characterized by a constant lifetime payment and no active labor force participation by the agent. The typical sequence of policies is quite simple: the program starts with unemployment insurance, then switches into monitored search and, finally, into social assistance. While wage depreciation is necessary for “policy transitions” within an optimal WTW program, a declining unemployment hazard is not. Benefits are decreasing during unemployment insurance and constant during both job search monitoring and social assistance, whereas wage taxes (subsidies) must decrease (increase) during a phase of job search monitoring.

In a calibration exercise, we use our model to analyze quantitatively the features of the optimal program for the U.S. economy. We conclude that switching from the existing U.S. system to the optimal WTW scheme can deliver sizeable welfare gains to workers at the low-end of the skill distribution (up to 3% of lifetime income), mainly by providing more insurance across employment states. We are currently extending this framework to study the optimality of training programs for workers on welfare.

Employment protection Employment protection legislation (EPL) has two separate components: a transfer from the firm to the laid-off worker (e.g., severance payments) and a tax paid outside the firm-worker pair (e.g., legal expenses in case of a trial). It is well established that, in the absence of contractual and market frictions, a government-mandated pure transfer from the firm to the dismissed worker is neutral on the labor market equilibrium. This powerful theoretical result has led the vast majority of macroeconomic researchers to conceptualize EPL as a tax only—a useful shortcut to describe a world in which also severance payments, coupled with some form of market imperfection, would have real effects on the economy.

This mainstream approach can, in principle, be justified on two grounds: (i) if quantitatively the tax component of EPL is substantially larger than the transfer component, or (ii) if the existence of contractual imperfections in actual labour markets induces the transfer component to act exactly as a tax on equilibrium (un)employment. In *The Employment Effects of Severance Payments with Wage Rigidities* (EJ 2005), we show that both presumptions are likely to be misplaced. First, we show that, in European economies, quantitatively the transfer component of EPL appears sizeable, and may even be considerably larger than the tax component. Second, through a set of new equilibrium comparative statics results obtained within a search model, we conclude that in presence of wage “rigidities” (arguably the most relevant contractual imperfection in actual labour markets) severance payments have, in general, different employment effect from firing taxes. This finding remains true when the degree of wage rigidity is endogenized in equilibrium.

U.S. vs. European institutions Labor market outcomes in Europe and the U.S. have differed systematically over the past thirty years. In particular, the unemployment rate has increased in Europe, while it has remained roughly constant in the U.S. economy. The dominant view in the profession is that a common macroeconomic shock and different policies across the two regions account for such diverging labor market experiences. Ljungqvist and Sargent (1998) offered the first rigorous model of this technology-policy interaction. Their argument is that unemployment in Europe surged in the 1980s because collecting generous unemployment insurance was, for the jobless workers, more beneficial than working at the low wage induced by the fact that technological change made their skills obsolete. Note that this mechanism takes place on the *labor-supply* side of the market.

In *Vintage Capital in Frictional Labor Markets*, we combine elements of standard vintage capital models with elements of frictional labor markets models into a new framework where this technology-policy interaction can be studied, theoretically and quantitatively, from the *labor-demand* side. In the model, the impact of an acceleration in capital-embodied productivity on equilibrium unemployment depends on the size of the policy variables. More rigid European-style institutions (generous welfare benefits, high taxes, and strict employment protection) exacerbate the long-run creative-destruction effect of a capital-embodied productivity shock on unemployment, particularly along the unemployment duration margin, as seen in the data. The mechanism we propose is also quantitatively consistent with the decline in the labor share and in the job-vacancy rate observed in Europe in the 1980s.

The search literature investigating the effects of technological change on equilibrium unemployment distinguishes between two polar models of capital replacement: a “creative destruction” economy where buying new machines requires destroying the existing match and searching for a new idle worker, and an “upgrading” economy where old machines in existing matches are replaced by new machines and the current worker is retained. These two approaches may have opposite comparative statics of technological change on equilibrium unemployment. This “replacement controversy” has been mostly qualitative and there has been no attempt at a careful quantitative comparison: once the models are calibrated to reproduce some standard facts about the macroeconomy, are their predictions truly so different? In *The Replacement Problem in Frictional Economies: A Near Equivalence Result (JEEA, 2005)* we address this question. Our main result is that these two models produce very similar quantitative outcomes. We explain this finding in light of the fact that, when the models are calibrated to the U.S. economy, both unemployment and vacancy durations are very short, i.e., the matching frictions are quantitatively minor. Hence, the equilibrium allocations of the model are remarkably close to those of a frictionless version of our economy where firms are indifferent between upgrading and creative destruction, since they face an infinite arrival rate of idle workers.

Social security A vast literature studies the issue of social security reform in the U.S. in light of the projected demographic trends which may soon put the sustainability of the current system at risk. The existing quantitative analyses are performed under the assumption that the

U.S. is a closed economy. In the paper *Global Demographic Trends and Social Security Reform (JME, forthcoming)* we offer a new benchmark for quantitative analysis of social security reform represented by a two-region model of U.S. and Europe (the North) and the less-developed world (the South), where capital is allowed to flow across regions. The timing and the extent of the demographic transition—and the associated economic forces shaping capital accumulation and equilibrium factor prices—are very different in the two regions. Therefore, the projected paths of factor prices in the North diverge substantially between closed and open economy benchmarks. We perform a wide range of policy experiments under the two scenarios, and compare results.

The main conclusion of our exercise is that if one is interested in quantifying the path of the fiscal variables (e.g., the value of the payroll tax) needed to keep the current social security system viable or to finance a transition towards a fully-funded system, then these two benchmarks yield similar results because of offsetting forces on the government budget constraint, in equilibrium. However, if the focus is on quantifying the path of factor prices, aggregate variables and, ultimately, welfare, then the two approaches can diverge significantly, and the choice of the benchmark is crucial. We argue that the open economy benchmark is becoming progressively more relevant. Our model, as well as several other demographics-based model of current account dynamics, predicts that capital will soon start flowing systematically from less to more developed countries. Then, the typical criticism moved to neoclassical open-economy models, i.e., that they overpredict the size of external wealth, will become irrelevant since, arguably, sovereign risks are negligible in the developed world thanks to better institutions.

In *Quantifying the Effects of the Demographic Transition in Developing Economies (BEJM, 2006)*, we lay out a similar two-region framework of the world economy, but this time we take the point of view of the less developed countries (the South), and highlight that the effects of the demographic trends for this region depend on the degree of international capital mobility, and on the extent to which the large social security systems in place in the more developed world will be reformed. In particular, these economic forces will affect the future distribution of income between labor and capital in the South. For example, under a full social security privatization scenario in the North, life-cycle saving expand in the North and capital flows towards the South, where wages grow faster, but rates of return to capital fall more rapidly, inducing a sizeable gain for labor income earners relative to capital holders.

Education subsidies Empirical evidence suggests a tight link between educational attainment and earnings, thus policies affecting education outcomes will also influence inequality, productivity and welfare. In *Education Decisions and Policy Interventions: A General Equilibrium Evaluation*, we build a life-cycle model where individuals make endogenous choices of education, consumption/saving, and labor supply, while facing incomplete insurance and financial markets. The model allows for heterogeneity among individuals in innate ability, which is transmitted across generations, and in productivity shocks during work-life. Moreover, the model incorporates inter-vivos transfers between parents and children as well as borrowing limits to address the issue of whether individuals are financially constrained in their education choice.

Various micro-data sources are used to calibrate and estimate the structural parameters of the model. Our policy experiments support the view that long-term, general equilibrium outcomes of education subsidies may drastically differ from their short-term partial equilibrium (i.e., with fixed factor prices) counterparts. The key channel is how the subsidy acts on the *selection* of those acquiring education. In partial equilibrium, subsidization increases education levels among the less able and reduces earnings inequality. In general equilibrium, instead, it is mostly the more able, but liquidity constrained, individuals who take up extra education, while the education levels of the less able can decrease, leading to a rise in earnings inequality.

References

- [1] Costantinides, G. and D. Duffie (1996). Asset Pricing with Heterogeneous Consumers, *Journal of Political Economy*, vol. 104, pp. 219-240.
- [2] Gordon, R. (1990). The Measurement of Durable Goods Prices, Chicago, IL: University of Chicago Press.
- [3] Heckman, J. (2001). Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture, *Journal of Political Economy*, vol. 109, pp. 673-748.
- [4] Kydland, F. and E. Prescott (1996). The Computational Experiment: An Econometric Tool, *Journal of Economic Perspectives*, vol. 10, pp. 69-85.
- [5] Ljungqvist, L. and T. Sargent (1998). The European Unemployment Dilemma, *Journal of Political Economy*, vol. 106, pp. 514-550.
- [6] Lucas, R.E. (1992). On Efficiency and Distribution, *Economic Journal*, vol. 102, pp. 233-247.
- [7] Nelson, R. and E. Phelps (1966). Investment in Humans, Technology Diffusion, and Economic Growth, *American Economic Review*, vol. 56, pp. 69-75.